Refining Evaluation Tools for Evidence-Based Professional Development

Prepared for the Office of Special Education and Rehabilitative Services (OSERS) in support of the State Personnel Development Grants (SPDG) program.

By:

Cecilia Ball, Kate Gladson Harbaugh, Mariana Romero Chong, and Kassie Haan

Morgridge College of Education, University of Denver

CFSP 4363: Program Development and Evaluation

Dr. David Furjanic

June 11, 2025

Table of Contents

Executive Summary

Overview

This project focused on refining the United States Department of Education's State Personnel Development Grant (SPDG) rubric, which is used to evaluate the quality of professional development applications for funding made by school districts and state education agencies across the country. Our team's goal was to streamline and strengthen the rubric so that it more clearly communicated expectations, reduced redundancy, and aligned more closely with evidence-based practices in adult learning and implementation science. Our work included conducting a literature review, analyzing past rubric versions and real grantee submissions, drafting edits to the rubric, and preparing the final recommendations and materials for SPDG leadership.

Purpose

The purpose of revising the rubric was to address challenges that school districts and state education agencies were facing in understanding and meeting the rubric's expectations. Our team found that the previous version often included overlapping requirements across sections, used technical or ambiguous language, and could be improved in alignment with current research on the elements that make professional development effective. In order to support both grantees and reviewers, our revisions aimed to make the rubric more straightforward, user-friendly, and grounded in research.

Sources to Support Improvements

The updated rubric draws on several key sources. The Standards for Professional Learning from Learning Forward (2022) emphasize that effective professional development is job-embedded, collaborative, and aligned with system-wide goals. Dunst and Trivette (2012)

identify participatory strategies such as modeling, coaching, and reflection as critical for sustained learning, particularly when used over extended periods of time. Fixsen and colleagues (2005) make a case for implementation science principles, including the importance of coaching, feedback loops, and systems-level support. These sources helped us revise the rubric to be more reflective of not only what professional development should look like, but also how it should be implemented and supported within real-world educational systems.

Main Areas of Improvement

During our analysis, we identified four major areas for improvement. First, many domains of the rubric repeated ideas, potentially leading to over-reporting and confusion among grantees. Second, some rubric items were vague or relied on jargon. Third, the rubric did not consistently reflect research on how adults learn or how high-quality professional development is implemented. Finally, scoring inconsistencies appeared to be common due to a lack of clear examples or guidance.

In response to these findings, we made several key changes. We:

- replaced dense paragraphs with plain-language bullet points and clarified what grantees should submit for each section.
- added tools for verification, such as resumes, memorandums of understanding, and fidelity rubrics.
- created “Score 4” examples to model strong responses.

Brief Recommendations

To support long-term implementation, we recommend providing annotated examples, developing training for reviewers and grantees, piloting the rubric before full rollout, and including supports to reduce self-scoring bias. These changes aim to improve the clarity and

reliability of the rubric while helping state and local education agencies more effectively demonstrate the quality of their professional development systems.

REFINING EVALUATION TOOLS FOR EVIDENCE-BASED PD: FULL REPORT

Introduction and Purpose

Improving the quality of professional development systems begins with the tools used to evaluate them. This project was conducted in collaboration with the Colorado Department of Education's State Personnel Development Grant (SPDG) team, which invited our graduate evaluation group to revise Rubric A, a key instrument used to assess the alignment and quality of grantee-submitted professional development plans. Rubric A, along with its worksheet, is designed to assess whether professional development aligns with evidence-based practices, adult learning theory, and implementation science. Over time, concerns emerged about the rubric's clarity, consistency, and usefulness, particularly in domains that showed wide variability between self-assessments and reviewer scoring.

In response, our team, consisting of four students in the CFSP 4363: Program Development and Evaluation course at the University of Denver, was asked to serve as external reviewers. Rather than developing new content, we were tasked with refining the existing rubric to enhance clarity, measurability, and alignment with research findings. We analyzed the current rubric's structure and language, compared independent edits, and collaboratively identified areas where grantees might encounter difficulties in interpretation or implementation.

A key finding that emerged across our analysis was that the current rubric's structure, particularly in high-variance domains, lacked clarity and measurability, making it difficult for both applicants and reviewers to use it consistently and effectively. The purpose of this project was to improve the effectiveness and usability of Rubric A. We aimed to clarify expectations within each domain, reduce redundancy, and strengthen connections to research on adult learning and systems-level change. By refining the rubric's language and structure, we sought to support

more accurate grantee self-assessment and greater consistency among reviewers, ultimately helping SPDG ensure its funded initiatives are both rigorous and sustainable.

Our process involved an iterative cycle of literature review, rubric analysis, revision, and group synthesis. Final deliverables include a revised version of Rubric A, a rationale document outlining key edits, a logic model, and a written report summarizing our work. While these materials represent our current recommendations, we recognize the value of continued collaboration and feedback. Our goal is that these tools will make a meaningful contribution to SPDG's efforts to strengthen personnel development systems and improve outcomes for students.

Background

The SPDG program, established under the Individuals with Disabilities Education Act (IDEA) and most recently reauthorized in 2004, is designed to support state educational agencies in reforming and enhancing their systems of personnel preparation and professional development. Its ultimate goal is to improve educational, early intervention, and transitional outcomes for children with disabilities. At the center of the program's mission is the recognition that high-quality, evidence-based professional development is a key driver of effective practice implementation, which, in turn, leads to improved child outcomes. To support this objective, the SPDG rubric functions as both a guidance and evaluation tool for grantees, helping them design, assess, and refine their professional development initiatives.

The rubric serves as a way to ensure that SPDG-funded efforts align with established standards for effective professional development. Its content and structure are meant to reflect core principles from implementation science, adult learning theory, and education policy. Implementation science, as described by Fixsen et al. (2005), identifies critical components of

effective practice adoption, including coaching, training, leadership, and data-informed decision-making. These components offer a framework for sustainable change. Furthermore, legislation such as the Every Student Succeeds Act (ESSA) emphasizes the importance of collaborative, data-driven professional development tied to student learning, highlighting the need for clear structures to support continuous improvement in educational practice.

In revising the SPDG rubric, our team drew from current research on professional and adult learning. Guskey's (2000) framework for evaluating professional development outcomes stresses the importance of aligning training with both organizational and educator needs. Dunst and Trivette's (2012) meta-analysis offers additional insight, showing that professional learning is most impactful when it involves active engagement, feedback, and real-world application across multiple sessions. Their Participatory Adult Learning Strategy (PALS) model illustrates how adult learners benefit from job-embedded learning experiences that include coaching, modeling, and reflection. Similarly, Learning Forward's 2022 Standards for Professional Learning provide a contemporary framework for designing effective professional development, emphasizing alignment with curriculum goals.

Together, each piece of literature informed our rubric revisions by reinforcing the need for clarity, alignment, and actionable guidance. The edits aim to remove redundancy, reduce ambiguity, and better align rubric components with best practices in the field. By integrating these research-based principles, the revised rubric is more likely to effectively support grantees in designing professional development that leads to improved educator practice and, ultimately, better outcomes for students with disabilities.

Our team's flow of logic for this project is as follows. We received input from the Office of Special Education and Rehabilitative Services at the U.S. Department of Education in the

form of multiple versions Rubric A, worksheets, a slide deck with information on SPDG, previous grantee submissions, and several sources of research to review. We analyzed each document, made edits to Rubric A and the worksheets based on our analysis of the provided literature, constructed our logic model to monitor alignment with our goals, and prepared to present our findings and revisions. As outputs, we provided the revised version of Rubric A and worksheets, change comparisons, recommendations, and a presentation that summarized our activity throughout this process. There are various expected outcomes from this project. In the short-term, Rubric A will present with improved clarity and reduced confusion for grantees. In the medium-term, OSERS will receive more consistent data and less redundancy from applicants. In the long-term, professional development effectiveness will increase, leading to greater support for educators, and ultimately, students across the country.

Figure 1 *Logic Model*	
Inputs:	Rubric versions, worksheet, SPDG guidance, grantee submissions, research.
Activities:	Document analysis, rubric editing, literature integration, logic modeling, presentation prep.
Outputs:	Revised rubric & worksheet, change comparison, recommendations report, presentation.
Outcomes:	Short-term: Improved clarity & reduced confusion for grantees. Medium-term: More consistent data & reduced redundancy. Long-term: Stronger EB-PD implementation & improved support for educators

Method

Our original proposal aimed to revise Rubric A to improve clarity, reduce redundancy, and better align scoring criteria with evidence-based practices in adult learning and implementation science. We planned to review the rubric and worksheet, make independent edit recommendations, and synthesize those revisions collaboratively. While this structure remained intact, our work evolved from proposing general language edits to engaging in deeper conceptual revisions, such as aligning rubric content with implementation science, refining scoring language to reduce ambiguity, and identifying patterns of self-scoring variance based on domain analysis and team discussion. We began by individually reviewing Rubric A, selected grantee submissions, and relevant literature. Each team member developed proposed edits and supporting rationale, with particular attention to vague, redundant, or misaligned sections. We

then met as a team to compare and analyze these changes using a structured comparison process. The comparison process included identifying shared themes and areas of disagreement, which prompted further discussion and refinement.

Our comparison revealed several common priorities: the need for clearer documentation expectations (e.g., resumes, fidelity tools), the removal of undefined or jargon-heavy terms, and a stronger emphasis on practices that support long-term sustainability, such as regular coaching, data-informed decision-making, and participant feedback loops. We placed particular emphasis on high-priority domains, such as B5 (Sustainability of Implementation Supports), which showed notable variance in scoring across reviewers. B5 focuses on how well applicants plan for the ongoing use and improvement of implementation supports, such as coaching structures and feedback systems, beyond the initial funding period. While some edits initially differed in tone or structure, we worked collaboratively to synthesize the strongest elements into a cohesive final rubric. Throughout this process, we aimed to maintain the rubric's original intent while enhancing its usability and fairness. Our goal was to ensure consistent interpretation among reviewers and applicants and to reflect better adult learning practices that support lasting change. Edits were grounded in implementation science literature and SPDG's internal guidance, and we prioritized language that supports capacity-building over compliance-based scoring.

Limitations

This project had several limitations. Most notably, it was outside of the scope of this work to pilot the revised rubric with grantees or reviewers, so the practical effectiveness of our revisions remains untested. Additionally, our scope was limited to Rubric A and the worksheet, excluding broader implementation tools or training protocols that may influence how the rubric is applied. Finally, our ten-week timeline limited opportunities for field testing and stakeholder

engagement. In the future, these limitations could be mitigated by piloting the rubric with a small group of grantees, gathering calibration data from reviewers, and incorporating direct feedback from SPDG stakeholders throughout the revision process.

Timeline

The process unfolded over several weeks. We finalized our project scope with SPDG in April and conducted literature reviews and individual rubric analyses in early May. By mid-May, each team member submitted proposed edits and rationale, which we synthesized into shared revisions. In late May, we created the change comparison chart, finalized edits, and developed presentation materials. We shared a draft with SPDG in the final week of May and presented our recommendations on June 4. This formal written report will be submitted on June 25. Each phase of the project reflected our commitment to clarity, collaboration, and alignment with SPDG's goals for equitable and sustainable professional development systems.

Findings

Throughout our revision of Rubric A, three overarching themes emerged that hindered the rubric's clarity, consistency, and effectiveness as both an evaluation tool and a capacity-building resource: redundancy across criteria, ambiguous or jargon-heavy language, and limited alignment with best practices in professional development and implementation science. These findings, drawn from both qualitative review and SPDG workbook analysis, shaped the direction and rationale for the changes our team proposed.

Redundancy

Redundancy across domains led to confusion and inefficiencies for both applicants and reviewers. Domains D2 (Data Systems) and D3 (Data Use), in particular, included overlapping prompts around data practices, with insufficient distinction between the structures for data

collection and the active use of data for implementation refinement. This overlap resulted in repetitive and/or vague applicant responses and unclear scoring outcomes. Our revisions sought to clarify the unique intent of each domain; D2 now focuses on the infrastructure and coherence of multi-level data systems, while D3 centers on how those data are interpreted, shared, and used to guide adjustments to training and support. These revisions draw from Fixsen and colleagues (2005), who distinguish between organizational and competency drivers in high-fidelity implementation systems (see Appendix A, p. 18 for more information).

Terminology

Ambiguous or jargon-laden language was another barrier to consistent, fair evaluation. Reviewers frequently noted difficulty in scoring responses due to vague terms like "ongoing support," "coaching is provided," or "data are used." This was particularly evident in domains B3 (Skill-Building Activities) and B5 (Trainer Feedback and Fidelity), where applicants tended to list professional development activities without explaining how those activities were delivered, evaluated, or aligned to specific skill-building goals. The revised language across these and other domains now prompts applicants to describe implementation processes with greater specificity, such as outlining roles, timelines, and tools, while also reinforcing expectations around evidence of practice rather than intention. These changes were grounded in Dunst and Trivette's (2012) findings that effective adult learning must involve multiple active strategies (e.g., modeling, feedback, reflection) over time and apply learning in real world professional settings.

Research Alignment

Strengthening alignment with current research and implementation frameworks emerged as a key opportunity to improve the rubric's overall effectiveness. In several domains, the

original language prioritized compliance-oriented descriptions over performance-based evidence, reducing the rubric's ability to capture meaningful implementation. To address this, our revisions shifted the focus from whether an activity occurred to how it was designed, delivered, and evaluated for impact. Domains such as B4 (Training Outcome Data), D5 (Data Collection Procedures), and E2 (Leadership Feedback Loops) now better reflect implementation science principles by asking applicants to describe feedback systems, decision-making processes, and how changes are informed by participant outcomes. These revisions draw on Fixsen et al.'s (2005) implementation drivers, Guskey's (2000) five levels of professional development evaluation, and Learning Forward's standards for professional learning.

Anchors

In addition to these three major themes, our team also identified repeated scoring discrepancies in the SPDG workbook data. In response to these discrepancies, we introduced an anchoring feature in six high-variance domains: A2, B3, B4, B5, D2, and D3 (See Appendix B). These domains showed the greatest discrepancies between applicant self-ratings and panel reviewer scores, often due to vague or minimally elaborated responses that failed to meet the threshold for a score of 4. To increase scoring alignment and support applicant understanding, each of these domains now includes the prompt: "Applicants who score a 4 typically provide..." followed by concrete examples. These additions increase transparency and make expectations more accessible to all applicants. In alignment with Guskey's (2000) work on formative evaluation and Learning Forward's call for defined success criteria, these changes reframe the rubric as both an evaluative tool and a learning resource.

Miscellaneous Revisions

Additional edits were also made across other domains to address inconsistencies, improve alignment with SPDG priorities, and reinforce sustainable implementation systems. These edits included clarifying communication of expectations (A1), strengthening coaching models and accountability (C1, C2), specifying roles in fidelity monitoring (D1), and integrating reinforcement and recognition (D4). These changes are outlined in Appendix A, which includes a domain-by-domain breakdown of original versus final language.

Summary of Findings

These revisions and additions ultimately strengthen the rubric's value not only as a scoring tool but as a support mechanism for building sustainable, equity-focused professional development systems. By clarifying expectations, eliminating redundancy, and grounding revisions in implementation science and adult learning research, the revised rubric promotes more consistent application across reviewers and applicants alike. In doing so, it advances SPDG's broader goal of supporting high-quality, scalable initiatives that drive lasting improvement in educator practice and student outcomes.

Recommendations for Improving Rubric Implementation and Scoring Accuracy

To improve the transparency, consistency, and accuracy of rubric-based evaluation processes, several key recommendations are proposed. First, it is recommended that annotated examples be provided. Developing clear "score 4" exemplars presented in side-by-side comparisons will assist grantees in accurately self-evaluating their submissions and foster more consistent scoring across reviewers. Second, comprehensive training for both reviewers and grantees could be developed and implemented. This training, delivered through webinars and detailed how-to documents, would serve to clarify common misinterpretations of rubric criteria and address prevalent rating challenges.

Third, before full-scale implementation, it is advisable to conduct a pilot of the revised rubric with a small cohort. This pilot phase will enable testing of the rubric's usability and scoring reliability, and the feedback gathered will inform necessary refinements to ensure the rubric's effectiveness when deployed at the statewide or national level. Finally, it is essential to address potential self-scoring bias proactively. Recognizing the risks of both over-rating and under-rating—whether due to misunderstandings of rubric criteria or organizational culture—it is important to implement validation strategies. These strategies can help ensure more accurate and consistent scoring across different contexts. Such strategies may include reflection checklists, calibration sessions, and embedded prompts within the rubric itself, all designed to improve scoring accuracy. Collectively, these recommendations aim to foster a more equitable, transparent, and reliable evaluation process.

References

Dunst, C. J., & Trivette, C. M. (2012). Moderators of the effectiveness of adult learning method practices. *Journal of Social Sciences, 8*(2), 143–148.

Fixsen, D. L., Naoom, S. F., Blase, K. A., Friedman, R. M., & Wallace, F. (2005). *Implementation research: A synthesis of the literature* (FMHI Publication #231). University of South Florida, Louis de la Parte Florida Mental Health Institute, The National Implementation Research Network.

Learning Forward. (2022). *Standards for professional learning.*

Learning Forward. (n.d.). *Definition of professional development.* Powered by Title II. Retrieved May 28, 2025.

APPENDICES

Appendix A: PD Components Column Edits and Revision

A1	
Original	Edited
Required elements: □Description of expectations for PD participants (e.g., attendance in training, data reporting) □Identification of what schools, districts, or other agencies agreed to provide (e.g., necessary resources, supports, facilitative administration for the participants) □Description of how schools, districts, or other agencies were informed of their responsibilities	Required elements: □ Description of expectations for PD participants (e.g., attendance in training, data reporting) □ Identification of what schools, districts, or other agencies agreed to provide (e.g., necessary resources, supports, facilitative administration for the participants) □ **Explain how expectations were reinforced over time (MOUs, regular reminders, kickoff meetings) not just communicated at the start.** □ Description of how schools, districts, or other agencies were informed of their responsibilities
A2	
Original	Edited
Clear expectations are provided for SPDG trainers and SPDG coaches/ mentors. Required elements: □ Expectations for trainers' qualifications and experience and how these qualifications will be ascertained. o Description of role and responsibilities for trainers (the people who trained PD participants). □ Expectations for coaches'/ mentors' qualifications and experience and how these qualifications will be ascertained. Description of role or responsibilities for coaches or mentors (the people who provided follow-up to training).	Clear expectations are **described and verified** for SPDG trainers and SPDG coaches/mentors. Required elements: • **Describe expectations for trainers' qualifications and experience, and how those qualifications were verified (e.g., resumes, interviews, performance tasks).** • Describe the roles and responsibilities of trainers (individuals delivering the PD) • **Describe expectations for coaches' or mentors' qualifications and experience, and how those qualifications were verified.** • Describe the roles and responsibilities of coaches or mentors. (individuals providing follow-up support after training). *Note: Descriptions that rely solely on attachments (e.g., "see appendix") will not receive full credit.*
B1	
Original	Edited
Accountability for the delivery and quality of training. Required elements: • Identification of the lead person(s) accountable for training. • Description of the role and responsibilities of the lead person(s) accountable for training.	Accountability for the delivery and quality of training. Required elements: □ Identification of the lead person(s) accountable for training. □ Description of the role and responsibilities of the lead person(s) accountable for training. □ **Include how lead trainer(s) will evaluate the effectiveness of training delivery, observe trainers, and adjust implementation supports based on fidelity or participant feedback.**

B2	
Original	Edited
Effective research-based adult learning strategies are used Required elements: • Identification of adult learning strategies used, including the source (e.g., citation). • Description of how adult learning strategies were used. • Description of how data are gathered to assess how well adult learning strategies were implemented.	Effective research-based adult learning strategies are used. Required elements: • Identification of adult learning strategies used, including the source (e.g., citation). • Description of how adult learning strategies were used. **• Description of how implementation of adult learning strategies was assessed (e.g., participant surveys, fidelity checklists, structured reflection).**
B3	
Original	Edited
Training is skill-based (e.g., participant behavior rehearsals to criterion with an expert observing) Required elements: • Description of skills that participants were expected to acquire as a result of the training. • Description of activities conducted to build skills. • Description of how participants' use of new skills was measured.	Training includes structured, active skill-building with opportunities to practice and receive feedback. Required elements: **• Clearly describe the specific, observable skills participants were expected to acquire.** **• Describe the specific activities used to build those skills (e.g., role-playing, modeling, guided practice, or feedback session).** **• Describe how and when skill acquisition was measured (e.g., observation checklists, fidelity rubrics, during/post training assessment).**
B4	
Original	Edited
Training outcome data are collected and analyzed to assess participant knowledge and skills. Required elements: • Identification of training outcome measure(s). • Description of procedures to collect pre- and post-training data or another kind of assessment of knowledge and skills gained from training. • Description of how training outcome data were reported. • Description of how training outcome data were used to make appropriate changes to the training and to provide further supports through coaching.	**Training outcome data are collected, analyzed, and used to improve training and follow up supports.** Required elements: **• Identify the specific training outcome measure(s) used (e.g., pre/post test, skill rubric, exit ticket).** **• Describe how and when outcome data were collected (e.g., timing, tools, procedures before/after training or during follow up).** **• Describe how training outcome data were summarized or reported (e.g., individual feedback, group trends, dashboards).** **Describe how outcome data were used to revise future training and/or provide differentiated follow-up support (e.g., targeted coaching, new training content, pacing changes).**

<table>
<tr><th colspan="2">B5</th></tr>
<tr><th>Original</th><th>Edited</th></tr>
<tr><td>Trainers (the people who trained PD participants) are trained, coached, and observed.
Required elements:
• Description of training provided to trainers.
• Description of coaching provided to trainers.
• Description of procedures for observing trainers.
• Identification of training fidelity instrument used (measures the extent to which the training is implemented as intended).
• Description of procedures to obtain participant feedback.
Description of how observation and training fidelity data were used (e.g., to determine if changes should be made to the content or structure of trainings, such as schedule, processes; to ensure that trainers are qualified).</td><td>Trainers (the people who trained PD participants) are trained, coached, observed, and provided feedback using a fidelity process.

Required elements:
• Description of the training provided to trainers, including duration and content focus.
• Describe the structure and frequency of coaching provided to trainers (e.g., individual, group session).
• Describe the procedures used to observe trainers including who observed and how often.
• Identify the fidelity instrument used to assess trainer delivery.
• Describe how participant feedback was collected (e.g., surveys, exit tickets, feedback forms).
• Describe how observation and fidelity data were used (e.g., to revise content, delivery, or ensure trainers met expectations).</td></tr>
<tr><th colspan="2">C2</th></tr>
<tr><th>Original</th><th>Edited</th></tr>
<tr><td>SPDG coaches use multiple sources of information in order to provide assistive feedback to those being coached and also provide appropriate instruction or modeling.

Required elements:
• Should describe the coaching strategy used and the appropriateness for use with adults (i.e., evidence provided for coaching strategies).
• Describe how SPDG coaches monitored implementation progress.
• Describe how the data from the monitoring are used to provide feedback to implementers.</td><td>SPDG coaches use multiple sources of information in order to provide data-driven feedback, support skill development, and adjust coaching strategies.

Required elements:
• Describe the coaching strategy or model used and explain why it is appropriate for adult learners, including citations or references.
• Describe how SPDG coaches monitored implementation progress.
• Describe how the data from the monitoring are used to provide feedback to implementers.</td></tr>
</table>

D1	
Original	Edited
Accountability for fidelity measurement and reporting system is clear (e.g., lead person designated) Required elements: • Provide a description of the role/responsibilities of the lead person and who this person is.	**Accountability for fidelity measurement and reporting system is clear (e.g., lead person designated).[10]** Required elements: • **Identify the individual or team responsible for overseeing fidelity measurement and reporting.** • **Describe the responsibilities of the identified lead(s), including how they ensure fidelity data is collected, reviewed, and reported accurately and on time.**
D2	
Original	Edited
Coherent data systems are used to make decisions at all education levels (SEA, regional, LEA, school). Required elements: • Describe data systems that are in place for various education levels. • Describe how alignment or coherence is achieved between various data systems or sources of data. Describe how multiple sources of information are used to guide improvement and demonstrate impact	**Coherent data systems are used to align decisions and support improvement across education levels (SEA, regional, LEA, school).** Required elements: • **Describe the data systems used at the school, district, and state level** • **Explain how data systems are aligned across levels (e.g., shared tools, timelines, data definition).** • **Describe how multiple types of data are analyzed together (e.g., fidelity, student outcomes, coaching logs) to guide decisions and demonstrate program impact**
D3	
Original	Edited
Implementation fidelity and student outcome data are shared regularly with stakeholders at multiple levels (SEA, regional, local, individual, community, other agencies) Required elements: • Describe the feedback loop for each level of the system the SPDG works with. ○ Describe how these data are used for decision-making to ensure improvements are made in the targeted outcome areas. Describe how fidelity data inform modifications to implementation drivers (e.g., how can Selection, Training, and Coaching better support high fidelity).	**Implementation fidelity and student outcome data are regularly shared across levels of the system to guide decision-making and improve implementation fidelity.** Required elements: • **Describe how implementation and outcome data are shared across system levels (e.g., school to district, district to state).** • **Describe who reviews the data, how often they review it, and for what purpose.** • **Describe how the data are used to make decisions at each level (e.g., identifying additional support needs, revising coaching intensity or training content).** • **Describe how fidelity data were used to modify or refine implementation drivers (e.g., selection, training, coaching) to support more effective and sustained implementation**

D4	
Original	Edited
Goals are created with benchmarks for implementation and student outcome data, and successes are shared and celebrated.[10] Required elements: • Describe how benchmarks are created and shared. • Describe positive recognition processes for achievements. Describe how data are used to "market" the initiative.	**Goals are created with benchmarks for implementation and student outcome data, and successes are shared and celebrated.** Required elements: • Describe how benchmarks are created and shared. • **Describe how achievements are recognized and celebrated.** Describe how data are used to "market" the initiative.
E1	
Original	Edited
Administrators are trained appropriately on the SPDG-supported practices and have knowledge of how to support its implementation. Required elements: • Role/job description of administrators relative to program implementation provided. • Describe how the SPDG trains and supports administrators so that they may in turn support implementers.	**Administrators are trained appropriately on the SPDG-supported practices and have knowledge of how to support its implementation.** Required elements: • **Clear description of administrators' roles and responsibilities in supporting implementation of SPDG practices.** • Describe how the SPDG trains and supports administrators so that they may in turn support implementers.
E2	
Original	Edited
Leadership at various education levels (SEA, regional, LEA, school, as appropriate) analyzes feedback regarding barriers and successes and makes the necessary decisions and changes, including revising policies and procedures to alleviate barriers and facilitate implementation. Required elements: • Describe processes for collecting, analyzing, and utilizing student and teacher data to recognize barriers to implementation success. • Describe processes for revising policies and procedures to support a new way of work.	**Leadership at various education levels (SEA, regional, LEA, school, as appropriate) analyzes feedback regarding barriers and successes and makes the necessary decisions and changes, including revising policies and procedures to alleviate barriers and facilitate implementation.** Required elements: • Describe processes for collecting, analyzing, and utilizing student and teacher data to recognize barriers to implementation success. • **Describe how this information is used to revise systems, such as policies, procedures, staffing, resource allocation, or timelines, to remove barriers and support implementation.**

Appendix B: Exemplar (4) Description Additions

Domain	Addition
A2	Applicants who score a 4 typically provide: • Specific quantifiable expectations (e.g., observe 5 sessions) • Clear qualification verification methods • Role-specific responsibilities • Balanced detail between trainers and coaches
B3	Applicants who score a 4 typically provide: • Clearly defined observable skills participants are expected to acquire • Description of interactive training methods used to build those skills • A specific assessment method for skill acquisition • Details on how and when skills were assessed • Evidence that activities were completed
B5	Applicants who score a 4 typically provide • A clear defined training plan for trainers • Structured and frequent coaching • Description of how trainers are observed using fidelity tool • A process for collecting and using participant feedback • Clear expectations of how data were used to revise delivery and provide targeted trainer support
D2	Applicants who score a 4 typically provide: • A clear description of the data systems used at each level • An explanation of how data are aligned across systems to support consistency and shared decision-making • Examples of multiple data sources being used together • A regular review process with evidence of data-informed action • Defined reporting structures
D3	Applicants who score a 4 typically provide: • Evidence that data are shared regularly across levels • Clear explanation of who reviews the data, how often, and the purpose • A feedback loop that results in documented action such as coaching, targeted assistance, or revised implementation • Description of how fidelity data were used to modify implementation drivers