



Developing a Program Performance Evaluation Plan

Brief | January 2021

Kathleen Hebbeler, SRI Education; Sarah Nixon Gerard, SRI Education; Erika Gaylor, SRI Education; Georgene Weisenfeld, National Institute for Early Education Research

Preschool Development Grant Birth through Five (PDG B-5) Grantees are required to develop a Program Performance Evaluation Plan (PPEP) and conduct a program performance evaluation. A well-designed PPEP is key to conducting a meaningful evaluation of the state's PDG B-5 activities. The Administration for Children and Families (ACF) intends for Grantees to regularly assess progress toward the goals and objectives in their strategic plans and adjust the plan, as needed, based on what they learn. PDG B-5 Grantees will use the program performance evaluation to examine the implementation of activities and determine the extent to which the state achieved the intended outcomes for children and families.

This brief is designed to assist states with developing or revising their PPEPs. It provides guidance and recommendations for states on three topics related to the PPEP:

1. *Showing the logic behind the state's strategic plan*
2. *Making decisions about measurement*
3. *Setting priorities to make the evaluation manageable*

We present numerous examples throughout and provide additional resources at the end.

Show the Logic of the Strategic Plan

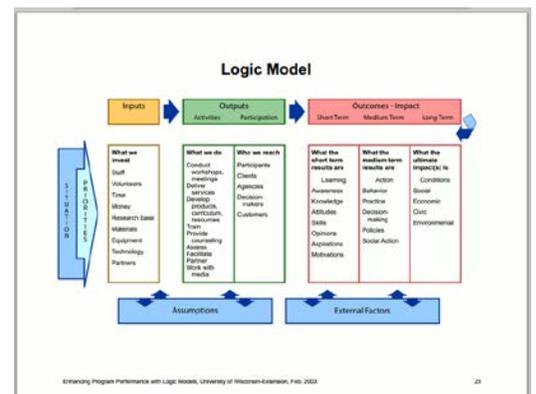
The state's PDG B-5 strategic plan identifies what the state wants to achieve for children and families and what activities it will undertake to achieve these outcomes. A logic model is a graphic representation that shows the connections between activities and intended outcomes, and a well-developed logic model provides the foundation for developing the PPEP.

The critical components of a logic model for the PPEP are the activities, outputs, and outcomes.

- *Activities* are actions in the strategic plan that the state will be carrying out.
- *Outputs* are direct products or results of those activities. Outputs lead to the outcomes, but they are not outcomes.

What is a PDG B-5 program performance evaluation?

- Ongoing self-assessment of progress toward goals in the state's strategic plan.
- Source of information to support continuous quality improvement.



Source: betterevaluation.org

- *Outcomes* are changes in systems or people achieved through activities, such as improved state infrastructure, quality of the service, access to services, or child or family skills and capacities. Logic models typically show short-, intermediate-, and long-term outcomes when one outcome leads to another. The time frame to achieve “short-” or “long-term” outcomes can be highly variable depending on the activity. The key concept is that the outcomes are linked in an if-then chain that shows that the short-term outcomes must occur for the intermediate- and long-term outcomes to occur.

Below we show the outputs and outcomes that a state might expect as a result of implementing this set of example activities. Notice the logical connections across the three columns. Are there intermediate- or long-term outcomes for children and families that the state would hope to achieve as a result of these short-term outcomes?

Activities:	Outputs:	Short Term Outcomes:
<ul style="list-style-type: none"> • Improve the quality of early childhood care and education (ECE) environments and workforce. <ul style="list-style-type: none"> – (1) Study, revise, and implement a Quality Rating and Improvement System (QRIS) that is outcome-driven. – (2) Provide the workforce with specialized training. – (3) Implement a pilot for an incentive model to encourage retention and a robust recruitment model. 	<ul style="list-style-type: none"> • A new QRIS and its associated standards • Number of trainings on QRIS standards provided • Number of ECE providers and teachers who attend the training • Incentive model pilot implemented in 5 sites 	<ul style="list-style-type: none"> • ECE providers know what high-quality, developmentally appropriate, responsive, and inclusive care and learning looks like. • ECE providers and teachers implement high-quality, developmentally appropriate, responsive, and inclusive care and learning. • ECE providers and teachers understand what supports are available to improve their practice and well-being. • ECE providers stay in the field longer.

For a logic model to guide the development of the PPEP, it must be sufficiently detailed to show the logical pathway between each major activity and its corresponding outputs and outcomes. The one-page summaries of logic models that Grantees included in their PDG B-5 renewal proposals communicate the high points of the plan’s logic, but do not have enough detail to inform the PPEP. States may need separate logic models for each major area in the strategic plan to show the connections between each activity and its expected outcomes. Given the comprehensiveness of the strategic plans, displaying pathways between each activity and its outcomes could result in many pages of logic models. We provide some guidance for how to make your PPEP manageable later in the brief by clustering similar activities and setting priorities for what you measure.

Developing a comprehensive logic model or series of logic models can be time-consuming, but it is well worth the investment. The finished product will show *how* the state expects each of its activities to lead to long-term outcomes for children and families. As the state moves forward with implementing activities, the

evaluation data might show that the activities were not implemented as planned, that they did not produce the expected outputs, or that they did not achieve the short-term outcomes. Reviewing the assumptions displayed in the logic model alongside the evaluation data will assist the states in deciding what, if any, adjustments are needed to increase the likelihood of achieving the long-term outcomes.

Make Informed Decisions about Measurement

Having developed a clear, comprehensive presentation of the strategic plan’s logic, the state must then decide what evidence will demonstrate that the activities were implemented, the expected outputs produced, and the intended outcomes achieved.

Develop evaluation questions

Formulating evaluation questions about the components of the logic model can be helpful to guide decisions about measurement. Evaluation questions about activities often take the form of “Was [activity] implemented?” States also might opt to ask questions about the quality or process of implementation (e.g., “How well was [activity] implemented?” or “What factors contributed to successful implementation?”). Questions about outputs focus on the products of the activities (e.g., “What did the activity produce?” or “To what extent did the activity produce what was expected?”). Questions about outcomes ask whether the activities resulted in changes in systems, programs, knowledge, attitudes, or behaviors (e.g., “Did more children show improved [outcome]?”). Below we provide examples of evaluation questions based on the sample logic model components in the previous section.

Questions about Activities:

- Was the QRIS revised and implemented? Was the new QRIS outcome-driven?
- Was the ECE workforce provided with specialized training?
- Was the pilot for the incentive model implemented as planned? What implementation challenges were encountered?

Questions about Outputs:

- How many trainings were provided on the revised QRIS?
- How many trainings were provided on how to provide high-quality inclusive care and learning?
- How many teachers were trained?
- How many sites implemented the pilot model?

Questions about Outcomes:

- How many ECE providers and teachers are knowledgeable about the critical features of high-quality, developmentally appropriate, responsive, and inclusive care and learning?
- How many teachers are providing high-quality, developmentally appropriate, responsive, and inclusive care and learning?
- How many teachers know how to access supports to improve their professional practices and personal well-being?
- Has retention of ECE providers and teachers improved in the pilot sites?

Describe activities at the right level of detail

To measure performance, the state must clearly describe each of the activities, outputs, and outcomes. Our review of the PDG B-5 Grantees' strategic plans showed that states describe activities with varying levels of specificity. For example, "Facilitate coordination between school systems and local community-based programs" could be listed as an activity; however, this broad description does not provide the level of detail needed to determine whether the state implemented it. By contrast, "Provide grants to local communities to establish community councils" describes an activity with enough detail to identify what evidence is needed to show the activity was carried out. States also need to avoid describing activities at too fine a level. The "Revise and implement a QRIS" activity discussed previously is likely to require multiple tasks (e.g., convene stakeholders, circulate drafts, revise drafts). Although this much granularity is helpful to the person managing the activity, documenting each step in the activity process is too much detail for the evaluation. The state needs to describe activities at the appropriate level of specificity to be both clear and measurable.

Develop indicators

To measure outcomes, the state must first develop one or more indicators for each outcome. An *indicator* is the evidence that shows whether the intended outcome was achieved. Outcomes are commonly written as global statements of what the state wants to achieve for the system or children and families (e.g., "Families will thrive and be healthy" or "Children will be ready for school"). Indicators operationalize these broad concepts and make them measurable. Good indicators begin with "number of" or "percentage of" and include the amount of change expected (how much), the intended audience (who), and the target timeline (when). A target or benchmark helps a state define what success looks like. For example, "90% of kindergarten teachers will report that they have received information from preschool teachers about entering kindergarteners' knowledge and skills by 2022" provides the target that allows the state to know if the activities focused on supporting a successful transition to kindergarten have been effective. This example also includes the elements of a good indicator: how much, who, what, and when.

Features of a good indicator

- Meaningful
- Direct
- Useful
- Practical to collect

Elements to include

- How much
- Who
- What
- When

Source: Innovation Network (n.d.)

Decide on data sources to use

Data provide the answers to states' evaluation questions. Two major categories of data available for the PPEP are (1) administrative data that are already being collected regularly for other purposes, and (2) new one-time or time-limited data collected for the program performance evaluation. Both administrative data and new data provide valuable information, but both types of data also have limitations.

Using existing administrative data can be an extremely cost-effective and a low burden strategy for obtaining data for the evaluation. Programs such as subsidized child care, Head Start and Early Head Start, and early intervention and preschool special education must regularly collect and report data.

Other potential data sources include the state's QRIS, licensing database, and workforce registry. However, administrative data may lack necessary data elements, and the quality of the data may be questionable. States should always interpret administrative data with caution, especially if the accuracy of the data has not been verified. Increased use of the data will work to improve its quality as data reporters come to understand the value of accurate information.

Some Components of a PPEP

- Comprehensive logic model aligned with strategic plan
 - Activities, outputs, short-, intermediate-, and long-term outcomes
- Evaluation questions
- Indicators and measures
- Data sources
 - Maximize use of existing data
 - Determine how to collect and manage new data
- Plans for analysis and reporting

States can augment administrative data by adding new data fields or by integrating data across systems to provide information that no single program's data can provide. Using PDG B-5 funds to enhance the state data capacity can provide a legacy of useful information long after the grant ends. For example, a state wants to use its workforce registry to provide data for this indicator: "50% of ECE providers and teachers will participate in evidence-based professional development (PD) each year", but the registry does not contain data on whether PD offerings are evidence-based. The state could address this by adding a new data field to the registry data collection.

One of the most powerful ways for a state to obtain new information without collecting additional data is by linking data across programs through the development of an early childhood integrated data system (ECIDS). Developing an ECIDS is a long-term undertaking, but is a worthwhile investment because it provides information that no single state data system can provide. Maximizing the use of current administrative data, working to improve data quality, and linking across relevant data systems can position the state to be able to monitor outcomes for programs and children and families for years into the future.

States also are likely to need to collect new data to answer some of their evaluation questions. Data can be collected through any number of techniques such as surveys, focus groups, interviews, or observations. New data can be collected once or repeated over several years. Depending on the nature of the data collected, new data collections may require the development of new tools and processes for data collection, management, and analysis. Some of these will be of relatively low cost (e.g., online surveys) while others will require substantial resources, if done on a large scale (e.g., classroom observations). For all data, but especially for the indicators of child and family well-being, the state should identify what information is regularly needed to examine the effectiveness of the early childhood system. As with all data collections, decisions about what kind of new data to collect need to be weighed carefully against the cost of collecting and sustaining the data collection. As part of developing its PPEP, each state needs to decide the appropriate balance of evaluation resources to allocate to time-limited data collections and to enhancing the state's administrative data capacity that will produce data into the future.

Make the Evaluation Manageable

States may notice that many of the good evaluation resources to support the development of the PPEP discuss how to develop a logic model for “your program” or how to develop indicators for “your program’s outcomes.” PDG B-5 Grantees are not developing a program; rather, they are engaged in systems building, requiring multi-agency activities to impact numerous short- and long-term outcomes. The comprehensiveness of their strategic plan requires that each state intentionally address this complexity in the PPEP, or it can easily become overwhelming and unmanageable. We provide guidance on how states can do this.

Cluster activities that are related

As part of developing a logic model and communicating with stakeholders about the evaluation, the state can group activities that are conceptually related or intended to lead to the same outcomes. For example, the state might group similar activities into larger systems efforts such as governance, access, or quality; or it might cluster a group of activities that focus on professional development needs that are expected to improve the quality of ECE environments. Clustering is a communication technique to help stakeholders see relationships among key activities in the strategic plan. Clusters can help provide an overview of the data collection and a structure for how to review the evaluation findings. The state will need to evaluate each of the activities in the cluster separately, but meaningful clusters can help stakeholders get a high-level view of what has and has not been accomplished and how these activities have worked together to impact the outcomes.

Set priorities for measurement

The primary purpose of the program performance evaluation is to support each state’s learning about the implementation of the strategic plan. The state is not obligated to collect data about every output and outcome in the logic model if it does not have the resources to do so. The state needs to determine which parts of the logic model are most important and feasible to learn about. Also, there is no need to collect two kinds of data or have three indicators, if one will suffice. States will need to collect data on each of the activities because this information is critical for reporting to stakeholders. Also, PDG B-5 Grantees must report on progress of activities in the Annual Program Performance Report. Much of the data for documenting the implementation of the activities can be straightforward. For example, the evidence that the state developed a new website for families is its address (i.e., URL). A state might choose to collect more information about a subset of activities because the state wants to understand more about how these activities were carried out and what factors contributed to their success.

The evaluation is meant to support the state’s learning. If resources are not sufficient to look at everything, the state must make informed choices about what is most important to know. Criteria to consider in deciding how to allocate evaluation resources include:

- Effort and cost of the data collection.
- Importance of the data for continuous improvement and informed decision-making.
- Sustainability of data collection and management methods.

By setting priorities for evaluation and designing data collection and analysis methods to reflect those priorities, states will have the information that is most valuable to them for continuous learning.

Use tables to show connections across components of the PPEP

Tables showing the connections between the elements of the PPEP (e.g., evaluation questions and data sources) help communicate the focus of the evaluation. They also help with managing the evaluation and guiding the analysis and reporting of the findings after the data are collected. We suggest that states organize the information around key conceptual areas and use multiple tables if a single table gets too complicated. The table on the next pages provides one example of how to show information about the evaluation questions and data sources associated with one set of activities, outputs, and outcomes.

Conclusion

The state's PPEP should provide a roadmap for how state leaders will use data to assess the effectiveness of the activities undertaken to build a high-quality early childhood mixed delivery system and engage in ongoing improvement. Equally important, thoughtful choices about data collection, integration, management, and analysis will allow the state to sustain continuous quality improvement for years to come.

Resources to support PPEP development

- [PDG B-5 TA Center Program Performance Evaluation 10/6/2020 Webinar slides](#)
- [PDG B-5 TA Center Resources on PPEP](#)
- [Preschool Development Grant Birth Through Five Initial Program Performance Guidance](#)
- [Preschool Development Grant Birth Through Five Program Performance Secondary Guidance](#)
- Innovation Network (n.d.) [Evaluation Plan Workbook](#)
- Innovation Network (n.d.) [Logic Model Workbook](#)

Topic: Improving Quality of Early Childhood Education Experiences

Activities	Evaluation Questions	Data Sources
Develop training modules for teachers and directors	Were the training modules developed?	Documentation of training modules (see outputs)
	Were the training modules easily accessible?	Documentation of how directors and teachers access the training modules Survey of directors and teachers <ul style="list-style-type: none"> • awareness of modules • ease of access
	Were the training modules high quality?	Review panel <ul style="list-style-type: none"> • indicators of quality
Provide coaching through the Child Care Resource and Referral (CCR&R) agencies	Was coaching provided?	Documentation that coaching was provided (see outputs)
Outputs	Evaluation Questions	Data Sources
# of training modules developed	How many training modules were developed?	Count of training modules
# of teachers and directors trained <ul style="list-style-type: none"> • # of teachers and directors using the training modules will increase by 10% each year 	How many teachers and directors completed the training modules? Are more teachers and directors participating each year?	Documentation from the online training system
# of coaches hired	How many coaches were hired?	CCR&R tracking
# of teachers who received coaching	How many teachers received coaching?	Coaches' logs
# of coaching sessions provided in a 12-month period	How many coaching sessions were provided?	Coaches' logs
Outcomes and Indicators	Evaluation Questions	Data Sources
Classroom quality will improve <ul style="list-style-type: none"> • # of ECE teachers with EC certificates increases by 10% annually • 80% of the sites that received coaching will show improved CLASS scores on their next observation 	Did classroom quality improve?	Teacher certification database Site Performance Reports

Suggested Citation:

Hebbeler, K., Gerard, S. N., Gaylor, E., & Weisenfeld, G. (2021). *Program Performance Evaluation Plan*. SRI International.

This Center is funded by the U.S. Department of Health and Human Services, Administration for Children and Families, Office of Child Care. Contract Number: HHSP2332015000411



PDG B-5 TA Center

A Service of the Office of Child Care
1100 Wilson Boulevard Suite 2800 (28th floor)
Arlington, VA 22209

Email: PDGB5TA@sri.com

Subscribe to Updates

[https://www.occ-cmc.org/
occ_announcements/signup.aspx](https://www.occ-cmc.org/occ_announcements/signup.aspx)