Evaluation Technical Assistance Brief

AUGUST 2024 • NUMBER 10

Subuhi Asheer and Betsy Keating

Designing an Implementation Evaluation: A Guide for Regional Partnership Grant Recipients

Introduction

Parents and caregivers affected by substance use disorders have complex service needs that span medical, behavioral health, and social service systems. Through projects funded by Regional Partnership Grants (RPG), organizations help families navigate these systems to improve child safety and well-being, family functioning, and adult recovery outcomes. RPG-funded projects often operate by integrating services across child welfare, substance use treatment, mental health, and judicial systems to improve participant outcomes. This requires extensive coordination and communication across systems.

As part of the RPG program, grant recipients and local evaluators collaborate to evaluate whether and how these programs benefit families. Although grant recipients are required to assess participant outcomes, implementation or process evaluations also can offer emerging evidence and knowledge about RPG interventions. An implementation evaluation (sometimes called a process evaluation) is a "scientifically valid analysis of how providers implement programs and how participants experience them. It can assess successes, challenges, and lessons learned from implementation, which can be used to identify facilitators of and barriers to program delivery" (Werner 2004).

Implementation evaluations can help RPG grant recipients understand how a funded program or set of services were delivered. Implementation evaluations typically focus on the process of program delivery and lessons learned rather than examining participant outcomes, such as changes in family functioning, substance use, or mental health metrics. For instance, implementation evaluations can assess how well the frequency and type of counseling services fit the needs of the intended participants, the role and benefits of key partnerships, and the potential for reaching additional families in the community.

Findings from implementation evaluations also help interpret the quantitative results of impact or outcome evaluations and help explain the reasons why an intervention was able to meet its intended goals (Exhibit 1). For example, an RPG program delivering home visiting services in three different counties may find that participants' mental health and substance use outcomes improved in one county but not in the others. Grant recipients can use an implementation evaluation to explore the reasons for the differences in these outcomes by collecting data on what services looked like in practice in the different communities, the characteristics of the participants served in each county, the variation in participant experiences, and key challenges and strengths of program delivery.

Grant recipients can use implementation evaluations to gather data on a variety of topics to meet their needs and goals (Exhibit 2).

Who should read this brief?

This brief provides information on how an RPG grant recipient might plan for an implementation evaluation. It was written with RPG project staff, their local evaluators, and other partners in mind, but it could also be relevant to other programs and organizations. This work was funded by the Children's Bureau in the Administration for Children and Families at the U.S. Department of Health and Human Services.

Exhibit 1. Uses for implementation evaluations

- **Document key information** about the program and how it was delivered, such as the number of sessions, frequency, and specific content or approaches used.
- **Contextualize findings** from outcome and impact evaluations by identifying key drivers or factors, such as the fit of the program for intended participants, the services that were most connected with outcomes, and the benefits of the program for participants.
- **Make evidence-informed decisions** about the program, including strategies for improvement, sustainability, and meeting participant needs.

RPG Regional Partnership Grants
Children's Bureau
Mathematica Progress Together

Exhibit 2. Key elements of an implementation evaluation

Dosage: The amount and frequency of sessions or services that were received by participants relative to what was planned.	**Adherence:** The degree to which the program (including all intended services or components) was offered as planned.	**Staff or Partner Experience:** Provider staff or partners' perceptions of the intervention being offered, the methods used, or how the intervention met the needs of participants.	**Quality:** The quality of staff-participant interactions and/or other aspects of program delivery.
Participant Needs and Satisfaction: Aspects of the intervention or services participants liked the most, and/or whether the participants were satisfied with their overall experience.	**Lessons Learned:** Key barriers to program implementing as intended as well as catalysts for program successes.	**Areas for Improvement or Adaptation:** Suggestions for strengthening program delivery to ensure better program fit for participants, and improve outcomes as well as program sustainability.	**Context:** The local context that may affect delivery of services in the community, such as other available services in the community, local child welfare or substance use policies and community support for services for families affected by parental substance use.

RPG project teams can begin by outlining a clear and realistic implementation evaluation plan that documents all the elements involved—from developing research questions, identifying data sources, and defining the data collection approach to describing the methods for analysis and disseminating findings. In developing a plan, RPG project teams should define the evaluation's purpose and scope, including how the implementation evaluation will align with their impact (or in some cases, outcome) evaluation; how the evaluation will help teams answer questions that participating communities have about the program; and what resources (such as budget, staff, and time) are available for the evaluation. Typically, this will be a formal document that project teams and evaluators develop in collaboration with relevant staff, partners, and program participants.

This brief outlines the key steps grant recipients can use for developing their plans and offers additional resources and tips they can use to guide them.

Step 1: Develop research questions for the evaluation plan.

Research questions form the backbone of any evaluation design. Defining targeted and meaningful research questions that are aligned with the RPG project's goals is an important first step for the implementation evaluation. Well-framed research questions are usually focused on one element and are framed in a concrete way. For example, if the study is evaluating how well partners collaborated to provide services, researchers might ask, What role do partner organizations play in service provision? Research questions should also be feasible to answer using data that are available and likely to produce findings that meet the study's goals. Consider the number of research questions necessary for the implementation evaluation and whether they address each of the key elements of the implementation evaluation (Exhibit 1). The project team should tailor the focus of the implementation evaluation to their specific context and needs; the implementation evaluation does not need to address all elements.

Backward mapping

In defining the research questions of interest, RPG project teams and local evaluators should start by engaging staff and the community to make sure that the evaluation's goals and approach are credible to those who are delivering and benefiting from the RPG program. For example, the evaluators and project team could begin by having discussions with frontline staff, partner staff, potential participants, and community leaders to understand what they are hoping to learn and achieve through the project and how the implementation evaluation can help. This process, known as backward mapping, would ideally take place during the planning period or early stages of implementation so that these considerations can be integrated into the evaluation design.

Based on the goals defined through these collaborative discussions, the team can map the goals to specific research questions (Exhibit 3). For example, if partners want to understand whether participants prefer in-person counseling rather than virtual counseling, researchers might ask, What mode of counseling do participants like better and why? This approach also advances the goal of conducting research equitably so that staff, community partners, and participants all have a voice and ownership. Using participatory methods, including diverse perspectives, and integrating equitable approaches into the evaluation design will support meaningful findings for the program as well as the communities being served (Shelton et al. 2020; Metz et al. 2021).

Exhibit 3. Illustrative research questions.

- How well did staff feel the training prepared them for program delivery? What additional trainings would be helpful?
- How did participation vary across the different service components offered by the program?
- On average, what percentage of planned sessions were staff able to complete?
- What strategies did service providers use to deliver the sessions?
- What state or local policies or practices influenced the implementation of RPG services?
- What components of the RPG intervention did participants find most useful and engaging?

Logic Model Review

Evaluators should also review the logic model for the intended program and consider the goals and research questions being addressed in their outcome or impact studies. Doing so helps to identify the specific program activities and possible outcomes that may form the basis for implementation evaluation research questions. For example, a program's logic model may show how a bundle of services and activities (such as case management, transportation services, and peer support groups) will come together to change participant outcomes. A sample research question based on this logic model might ask, What is the actual dosage received by participants for each service type? Understanding the amount of programming (dosage) received by participants for each type of service in the bundle would help (1) contextualize how variations in the services received affected participant outcomes and (2) explain impact evaluation findings.

Step 2: Identify and map data sources to address the research questions.

Data sources for implementation evaluations typically fall under two broad categories: (1) quantitative and (2) qualitative. Researchers often use quantitative methods like surveys to gather data from a wide variety of respondents and examine changes or trends over time. Qualitative data, such as interviews with provider staff or participants, is useful for digging deeper and collecting nuanced data that help unravel why and how participants and staff may experience programming and what factors may be driving their perspectives. The two types of data complement each other and can provide a well-rounded picture of implementation when used together in a mixed-method approach. For example, asking what service components program participants found useful could be addressed through a participant satisfaction survey (quantitative) and participant focus groups (qualitative). Participant surveys delivered after sessions can help evaluators understand the level of satisfaction with a particular component, while focus groups with participants can provide details about why the level of satisfaction is low or high.

Project teams and evaluators should identify the data sources that can help answer each of the specified research questions. Exhibit 4 outlines data sources, the topics they can address, sample research questions, and key considerations associated with each. The process of selecting research questions and mapping data sources is iterative, so it is important to approach it with flexibility. For example, grant recipients may need to revise a research question to align with available data or drop a research question if the data source to answer it is unavailable or too expensive to pursue. It is important that teams consider their capacity to collect and analyze data and only propose the data sources that will be used. Grant recipients should also think about how they will manage the data that are collected, particularly if they are using a new data source not previously collected through the grant.

As part of the cross-site evaluation, RPG grant recipients already collect data on various components of service delivery that could be helpful for an implementation evaluation. For each data source, teams should define the limitations and benefits and balance them accordingly. For example, observations are a useful source of data on the quality of interactions between service providers and participants, but limitations include the subjectivity of the observer and the risk that having an observer present during service delivery might change the behavior of the service provider and program participant.

Enrollment. Data on the participants who enroll in the RPG program include their specific characteristics and demographics. Evaluators can use these data to answer whether the program is reaching the intended population and participants with specific characteristics or to identify variation in participant enrollment and areas where enrollment may be lower or higher than expected.

Service logs. Service logs consist of data on the length of service interaction, number of family members present, date and frequency of interaction, the type of service received, name of specific model or intervention (if any), and participant's level of engagement in the service. These data can help grant recipients address research questions on dosage and adherence or understand whether the intervention's timing and frequency aligned with what was planned.

Closure. Closure data provide information on program completion and possible reasons for participants dropping out of the program.

Semi-annual progress reports (SAPR). Grant recipients document their activities, accomplishments, and challenges every six months in the SAPR. These reports can be a useful source of data to address research questions on program reach and dosage; fidelity (for example, changes to program delivery compared with what was intended); types of barriers to service delivery; facilitators for success; and lessons learned.

Other data that RPG grant recipients could consider collecting for an implementation evaluation include:

Fidelity logs. RPG project teams may also consider using fidelity logs to measure adherence (whether and how their program is implemented as intended), which can be useful for a rigorous impact or implementation evaluation. Grant recipients often already report in their SAPR on changes to program delivery compared with what was intended. RPG project staff and evaluators should work together to decide which data to collect through fidelity logs, based on the RPG program's structure, content, and theory of change along with the goals of the research questions. Project teams implementing an evidence-based program may have fidelity logs that have already been developed and could be used or adapted as needed. Typically, program staff complete fidelity logs after each service encounter.

Participant Surveys. Surveys of participants can provide data on participant perceptions of the program and its key components as well their needs and areas for improvement. Surveys could include both multiple-choice and open-ended response options to gather quantitative and qualitative measures. Project teams can also choose to administer them online or in pencil-and-paper format.

Participant interviews or focus groups. Open-ended discussions with program participants (as well as potentially their family members who may be involved in programming) are useful for gathering nuanced qualitative data on topics such as participant needs, community context, perceptions of services received, and suggestions for improvement. In planning focus group data collection, consider logistics and participant comfort as well as the feasibility of offering food, transportation reimbursement, and child care for in-person groups.

Staff surveys. Surveys of staff can help RPG project teams gather systematic data on staff characteristics (such as education level, educational background, professional experience, and so on) as well as feedback on staff perceptions about training, the program, and participant needs. To ensure candid and honest perspectives, project teams should ensure that survey data are anonymized to protect staff identities.

Staff and partner interviews. Discussions with staff can offer helpful qualitative data from the perspectives of providers, supervisors, and partner agency staff on perceived needs, successes, and challenges in delivering services. This is also an important source of data on the policy and service delivery context of the communities served.

Independent observations of service delivery and/or training. Structured observations conducted by independent observers can help assess program fidelity, quality of service interactions, engagement of participants, and successes and challenges. Structured observations of the training offered to service delivery and partner staff can also assess adherence (for instance, if the developer requires that all staff receive specific training before beginning implementation, an observation can determine the type of training received and the benchmarks or expectations for implementation).

Review of other materials – Materials may include organizational policy documents, curriculum materials, clinical or facilitator manuals, protocols or assessment tools, and outreach flyers, among many others. These can help evaluators gain an in-depth understanding of the intended intervention, plans for service delivery, the organization's policy and service delivery environment, and any impending challenges or roadblocks.

Exhibit 4. Data sources and their uses

Data Source	Key topics	Sample research question	Considerations
Enrollment	• Adherence • Context • Participant perspectives	What were the demographic characteristics of the program's participants?	These are data that RPG projects already collect, so they likely would have minimal additional burden to collect them for an implementation evaluation.
Service logs	• Adherence • Dosage • Quality	What percentage of planned home visiting sessions were staff able to complete?	These are data RPG projects already collect, so they likely would have minimal additional burden to collect them for an implementation evaluation.
Closure	• Program completion • Reasons for dropout	To what extent did participants complete the program?	These are data RPG projects already collect, so they likely would have minimal additional burden to collect them for an implementation evaluation.
Fidelity logs	• Adherence • Dosage • Quality • Staff perspectives	What, if any, adaptations did staff make to the planned service components?	Fidelity is typically easier to assess if the intervention (for example, an evidence-based program) has defined and structured components but more challenging to measure with flexible components such as peer supports.
Participant surveys	• Engagement • Overall program quality and satisfaction • Participant perspectives • Participant satisfaction • Session quality	What were the emotional and mental health needs of participants? What did participants find most useful and engaging?	Consider the appropriate timing and purpose of surveys based on the research questions and outcomes being assessed. For example, post-session surveys are very brief, quick ways of gathering immediate reactions from participants, while follow-up surveys are used after the program ends and may be longer. It may also be useful to coordinate the timing of surveys, if they are being used to gather data on both implementation and impact studies.
Staff surveys	• Areas for program improvement • Barriers • Context • Staff perspectives • Strengths	How well did staff feel the training prepared them for program delivery?	Consider timing the survey to align with research questions. For instance, teams could administer the staff survey after staff complete training as well as at the end of programming, to get their feedback at multiple points in time.
Staff and partner interviews	• Areas for program improvement • Barriers • Context • Staff perspectives • Strengths	What do staff think are the main barriers to providing services? What additional trainings do staff think would be helpful?	Consider timing the interviews to align with the research questions. For instance, teams could conduct interviews at program start and end or at other intervals depending upon what they are hoping to learn.
Participants interviews and focus groups	• Areas for program improvement • Barriers • Context • Participant perspectives • Quality • Strengths	What were the aspects of programming that participants found most useful? What would they suggest changing?	Use a systematic recruitment approach to include a diverse mix of participant characteristics and perspectives.
Independent observations of service delivery and/or training	• Adherence • Context • Quality	To what extent did staff use trauma-informed strategies in their sessions?	Consider the range of staff, settings, and type of service components to ensure observations accurately represent diverse aspects of programming. There is also a risk that observer presence may change the quality of program delivery or participant response.
SAPR and other program materials	• Adherence • Context • Quality	What are the main goals of the program? What are the main service components of the program?	In addition to program-specific documents, this could also include press releases, media coverage, and so on.

Step 3: Define a timeline and approach for collecting data.

In defining the data collection approach, grant recipients and evaluators should consider the mode of the data collection, the staff used to collect the data, the timing, and the needs and context of participants and their families (Exhibit 5). In the evaluation plan, evaluators should also factor in the time needed to obtain institutional review board approval when determining the timeline for data collection. In the evaluation plan, RPG projects can consider outlining the following characteristics of data collection:

Exhibit 5. Key factors that could affect data collection plans

- Length of intervention and of grant period
- Frequency of data collection
- Respondent burden
- Institutional review board process
- Staff availability
- Community needs and context
- Grant or other funding requirements

Mode. Evaluators and program staff should discuss what format and method is optimal for collecting each type of data, given what they know about their program and community's context and setting, the resources available, and the participant characteristics. For example, if most participants live in rural communities with limited or unreliable internet access, then conducting frequent online surveys or virtual interviews may not be the most viable option.

Role of data collectors. Teams should also define who will be responsible for collecting each type of data. For example, evaluators may administer interviews or focus groups to allow participants to speak more freely, while program staff may collect fidelity or service logs because of their knowledge of service provision. In making these decisions, it is important to consider strategies for building trust and credibility to ensure that participants feel comfortable sharing their candid perspectives.

Timing and frequency. The timing of data collection depends upon the types of data included in the study, the research questions, as well as the analysis plans. For some data sources, it makes sense to collect data on an ongoing basis, while others are best collected at specific points in time. Having a combination of ongoing and point-in-time data collection across data sources can give program staff and evaluators a more well-rounded assessment of implementation activities and potential participant outcomes. To assess participant reaction to the services received, evaluators could collect satisfaction surveys after each session and conduct focus groups once at the end of programming.

Community needs and context. In planning and considering all of the elements above, RPG projects should keep the needs of the community and the local context in mind. Teams should work collaboratively with program participants and others involved in programming to define an approach that will be successful and effective in gathering meaningful data. Evaluators should think carefully about what will work in the program and community's context in terms of data collection mode, frequency, timing, and so on. For instance, a project team may initially plan to survey its partners every six months on their perceptions of the collaboration and partnership but later realize, based on planning discussions with partners, that this would be too burdensome as well as costly. To ensure that the data collection approach is aligned with partner needs, project needs, and the budget, the team could reduce the frequency of the partner survey to two or three times in the grant period.

Step 4: Discuss and outline the approach for analyzing the data and reporting the findings.

As part of the evaluation plan, RPG grant recipients can describe how they will analyze the data to address the research questions. Depending upon the type of data collected, evaluators may need to use both quantitative and qualitative analysis methods.

For quantitative data, the plan can identify methods such as calculating the average, range, or percentage that met specific metrics. For qualitative data, evaluators can describe the methods for identifying themes across respondents and data sources (see this helpful how-to guide for more information on coding and analyzing qualitative data).

In line with the best practices for conducting an equitable evaluation, RPG projects can consider including ways for staff, program participants, and partners to weigh in as part of the analysis process. For example, the local evaluation team could document plans to convene discussions to gather feedback from program staff, community representatives, and program participants on emerging findings (also known as sense-making).

> "Equitable implementation occurs when strong equity components—including, explicit attention to the culture, history, values, assets, and needs of the community—are integrated into the principles, strategies, frameworks, and tools of implementation science." (Metz et al. 2021)

Finally, evaluators can describe the planned reporting and dissemination activities for the implementation evaluation, such as a final report, presentations, blog posts, or journal articles. Some grant recipients may develop videos for social media or podcasts to share findings in creative ways with their audience. It may be beneficial to rely on multiple dissemination strategies designed to reach different audiences in a more equitable way. For instance, grant recipients may publish journal articles or conference presentations to inform the broader research field and tailor more focused presentations for partner organizations or community groups.

Summary and additional resources

Implementation evaluations can offer important insights on key lessons related to delivering RPG-funded programs and can identify strengths, areas for improvement, and barriers to service delivery for participants. Developing a detailed and tailored plan that engages staff, partners, and participants in the process; defining meaningful goals and research questions; and using well-aligned data sources will forge an important foundation for learning and facilitate analysis and reporting of findings. This plan can become the basis for continuing evaluation and program improvement as well as other research activities that can inform grant recipients, the Children's Bureau, and the wider community.

For more information on implementation evaluations, please visit the following resources:

- Measuring implementation fidelity brief
- Evaluation of implementation and outcomes brief
- Tip sheet on developing an evaluation plan
- Tip sheet on conducting focus groups
- Tip sheet on conducting observations
- Data collection methods for evaluation: Document review
- Formative evaluation toolkit
- Toolkits with resources on implementation evaluations
- Office of Planning, Research and Evaluation (OPRE) program manager's guide to evaluation, third edition
- Step by step guide to conducting qualitative analysis (PDF) and video
- Brief on implementation evaluation planning, including guidance on backward mapping

References

Metz, A., Woo, B., and Loper, A. "Equitable Implementation at Work." *Stanford Social Innovation Review*, vol. 19, no. 3, 2021, pp. A29–A31. https://ssir.org/articles/entry/equitable_implementation_at_work.

Shelton, R. C., Chambers, D. A., and Glasgow, R. E. "An Extension of RE-AIM to Enhance Sustainability: Addressing Dynamic Context and Promoting Health Equity Over Time." *Frontiers in Public Health*, vol. 8, 2020. https://www.frontiersin.org/journals/public-health/articles/10.3389/fpubh.2020.00134.

Werner, A. *A Guide to Implementation Research*. Washington, DC: The Urban Institute Press and Rowman & Littlefield, 2004.