Westat®

Improving lives through research™

Center to Improve Project Performance

Guidelines for Working with Third-Party Evaluators

Authors

Sarah Heinemeier, Compass Evaluation & Research
Anne D'Agostino, Compass Evaluation & Research
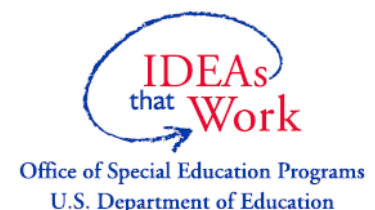Jill Lammert, Westat
Thomas A. Fiore, Westat

June 24, 2014

About these Guidelines

These Guidelines were developed as part of the Center to Improve Project Performance (CIPP) operated by Westat for the U.S. Department of Education, Office of Special Education Programs (OSEP). The authors thank our external reviewers Deborah Daniels-Smith, Amy Gaumer Erickson, and Shavaun Wall, and the OSEP Project Officers who provided input.

Suggested Citation:

Heinemeier, S., D'Agostino, A., Lammert, J.D., & Fiore, T.A. (2014). *Guidelines for Working with Third-Party Evaluators*. Rockville, MD: Westat.

IDEAs
that Work
Office of Special Education Programs
U.S. Department of Education

Overview of the Center to Improve Project Performance

First formed in 2008, CIPP's overall mission is to advance the rigor and objectivity of evaluations conducted by or for OSEP-funded projects so that the results of these evaluations can be used by projects to improve their performance and used by OSEP for future funding decisions, strategic planning, and program performance measurement. CIPP is operating under its second five-year contract.

The first CIPP contract provided summative evaluation support and oversight to 11 projects, selected by OSEP, in planning and executing their summative evaluations. CIPP staff worked with project and OSEP staff to refine each project's logic model and develop its summative evaluation design. Based on the evaluation design and plan, CIPP staff oversaw project summative evaluation activities and provided technical assistance (TA), as needed, to the grantees by selecting samples; developing draft instruments; monitoring data collection and performing reliability checks; analyzing study data; providing accurate descriptions of the methods and valid interpretations of findings; and organizing, reviewing, and editing project evaluation reports.

The second CIPP contract continues the work with the selected projects from the prior contract. Additionally, beginning in 2014, CIPP will provide intensive TA to 16 of OSEP's largest grantees in the development of their logic models and their formative evaluation plans. Related to the work on project evaluations, CIPP staff will work with OSEP staff to improve the consistency, objectivity, and rigor of OSEP's 3+2 evaluations, a formal process applied to projects funded in excess of $500,000 to evaluate their implementation and early outcomes following Year 2 of their grant. Also, CIPP will continue to provide TA in evaluation to OSEP-funded projects on request, prepare a variety of TA products focused on evaluation issues, and provide presentations on evaluation through Webinars and conferences.

Contact Information:

Thomas Fiore, CIPP Project Director
Westat
ThomasFiore@westat.com

Jill Lammert, CIPP Assistant Project Director
Westat
JillLammert@westat.com

Patricia Gonzalez, Project Officer
Office of Special Education Programs
U.S. Department of Education
Patricia.Gonzalez@ed.gov

GUIDELINES FOR WORKING WITH THIRD-PARTY EVALUATORS

Contents

Exhibits in Main Text

Exhibits within Appendices

Introduction

Generally, the purpose of an evaluation is to provide information on a project's implementation and outcomes. This includes providing qualitative and quantitative information on how well the project components have been implemented and analyzing the extent to which the project's objectives and outcomes have been achieved. The results of such evaluations provide project implementers with evidence to make decisions about project improvements, expansion, and sustainability; assess efficiency and guide cost-containment strategies; and facilitate replication in other settings. More importantly, evaluation results can provide information on a project's impact—information that can be used by the funder, the Office of Special Education Programs (OSEP), and by other key stakeholders to make an assessment of the nature and scope of project achievements. OSEP grantees are required to report on their project's accomplishments using tools such as the Annual Performance Report. Grantees use evaluations to identify what to measure for this reporting and to plan for and track the process of collecting, analyzing, and reporting on each desired accomplishment or evaluation metric. This document is written to assist grantees and their OSEP Project Officers in planning for, finding and hiring, and working with third-party evaluators to design, implement, and complete a project evaluation.

Evaluations typically feature three components targeted at three distinct lines of inquiry: progress monitoring, formative evaluation, and summative evaluation.

- A *progress monitoring component* examines the extent to which the project is *progressing* toward attaining its objectives and yearly benchmarks. Methods used often rely on administrative records and on descriptive (e.g., frequency of responses, measures of central tendencies) and correlational (i.e., exploring relationships among variables) statistical techniques.
- A *formative component* addresses questions related to *how well* the project's components and strategies are being implemented. Methods commonly include qualitative techniques such as interviews and observations and quantitative techniques such as surveys, and descriptive and correlational statistics.
- A *summative component* addresses the question of the *effectiveness* of the project in achieving its goals and desired impact (including impact on students) and identifies features/components of the project that were unique and/or effective (or ineffective). Summative methods often focus on quantitative methods such as descriptive, correlational, and advanced statistics, but also can include qualitative analysis of observational, interview, and open-ended survey data.

Appendix A contains more information on each of these components. Some readers may find it helpful to review Appendix A before proceeding. Lammert, Heinemeier and Fiore (2013) is another good resource.

Grantees often choose to work with a third-party evaluator—a qualified professional trained and experienced in the techniques to be used in the evaluation—who can help the project conduct any or all of the evaluation components listed above. This product is prepared under the assumption that grantees have already decided to hire a third-party evaluator, although the tasks that may be assigned to the third-party evaluator may vary considerably from one project to the next. Throughout, we provide ideas, tips, strategies, and suggestions grantees may find useful to make the most of an evaluation that incorporates a third-party evaluator. The document presents a discussion of the

benefits, drawbacks, and limitations of using a third-party evaluator and practical guidelines for creating a third-party evaluation scope of work, developing a Request for Proposals, soliciting bids for and contracting with a third-party evaluator, and monitoring and managing the work of the third-party evaluator.

Part 1 of this document provides general considerations for grantees who want to make the most of a third-party evaluation. Following this, Part 2 discusses the steps involved in finding and hiring a third-party evaluator, and provides guidance on creating a Request for Proposals, completing the solicitation process, and preparing and executing a contract for services.

Part 3 discusses specific guidance on monitoring and managing the work of the evaluation overall, including how to create a successful working relationship with the third-party evaluator and how to know when the evaluation is—and is not—proceeding as planned or meeting project needs. Part 4 focuses on wrapping up the evaluation project. Finally, the Appendices include a primer on evaluation design and planning (for those readers who may want some additional information on this topic), sample documents and templates that may provide further ideas and guidance for working with a third-party evaluator, information on practices for protecting confidentiality, and recommended readings on research and evaluation methodology.

Part 1. Making the Most of a Third-Party Evaluation

Typically, a third-party evaluator can be thought of as a "critical friend" who provides support, assistance, and feedback to the project through the formative and summative methods of evaluation. To this end, this document is written to help ensure a project makes the most of its investment in a third-party evaluator—which requires grantees have some exposure to and familiarity with evaluation basics. Appendix A includes a brief primer on evaluation basics; grantees may find reviewing that information useful before proceeding with the rest of this document.[1]

> **Top 5 Tips for Working with Third-Party Evaluators**
>
> 1. Hire as early as possible (such as during the application development or planning stages) even if only to conceptualize and design your evaluation;
> 2. Expect to devote time to the evaluation—even if only in the form of communication and monitoring the evaluation's progress;
> 3. Conduct an evaluation needs assessment—use the findings to create a contracted scope of work for the third-party evaluator;
> 4. Communicate regularly—keep regular track of evaluation activities and any implementation issues that arise; and
> 5. Receive interim reports and work products at regular intervals—monitor implementation of activities and use feedback to make project improvements.

1.1 Determining what is needed from the third-party evaluation

At the start of a project, grantees may (a) have a complete evaluation plan, (b) need to revise or update the project's evaluation plan, or (c) need to develop an evaluation plan for the project. Grantees that have a complete evaluation plan may elect to work with a third-party evaluator to complete specific tasks. Grantees that need to revise, update, or develop a plan may choose to work with a third-party evaluator to complete these design tasks. The third-party evaluator may then continue to work with the project to conduct the evaluation or the grantee may elect either to do the evaluation work internally—especially if the evaluation will be primarily formative—or to search for and hire a different third-party evaluator.

The first step in developing an evaluation plan is to identify the project's goals, strategies, outputs, outcomes, and the evaluation questions (see Appendix A). This information can then be used to complete an evaluation needs assessment, like in the example presented in Exhibit 1. Grantees are encouraged to review Appendix A or other evaluation resources (see Appendix F) if any of the items or terms in the needs assessment are unfamiliar. **An evaluation needs assessment can help grantees identify the specific tasks that need to be conducted for the evaluation, including those that will be contracted to a third-party evaluator.** Ideally, the needs assessment will be conducted as part of the proposal process or as soon as possible after the project receives its "green light" from OSEP.
The sample needs assessment presented in Exhibit 1 is for the fictional Anywhere State Speech and Language Pathologist (SLP) Support Project, which is designed to respond to the need for highly qualified SLPs who are proficient in evidence-based practices and who can work with bilingual secondary students with disabilities (more examples featuring the Anywhere State SLP Support project are presented in Appendix A).

---

[1] See also Lammert, Heinemeier & Fiore (2013).

Exhibit 1. Sample Evaluation Needs Assessment for the Anywhere State SLP Support Project

Question	Check the best option…	Possible Third-Party Evaluator Tasks

(1) Does your program already have an evaluation plan (a description of the evaluation questions, data collection tools and methods, analysis approach, and reporting requirements)?
 	
Note: Very often projects have some or all of an evaluation plan in place but the plan requires review or revision after a project is funded.
	☐ Yes, there is a complete evaluation plan in place, which responds in full to the evaluation requirements—**proceed to question 2. If you want to double check your answer, complete the checklist at right to identify possible third-party evaluator tasks**
	
	☒ There is a plan, but I'm not sure if it is complete or if it responds to requirements in full— **complete the checklist at right to identify possible third-party evaluator tasks**
	
	☐ No— **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Create or review the comprehensive evaluation plan
		OR
		☒ Review, develop, or refine formative evaluation questions
		☒ Review, develop, or refine summative evaluation questions
		☒ Identify or review data collection sources
		☒ Identify or review data collection instruments
		☒ Create/pilot test data collection instrument(s)
		☒ Design data collection procedures
		    ☐ Implementation progress monitoring
		    ☒ Service statistics (e.g., numbers served; numbers of services provided)
		    ☐ Fidelity of implementation
		    ☒ Outcomes/impact data
		☒ Design data entry/ management procedures
		☒ Create data analysis plan
		☐ Design or review evaluation budget
		☐ Design or review report template(s)

(2) Are there internal staff with skills necessary to conduct the evaluation?
 	
Note: Very often projects will ensure statisticians and qualitative specialists (team members who specialize in qualitative research) are available to work on or support the evaluation.
	☐ Yes, internal staff are qualified for the types of evaluation required— **check off the applicable and needed skills below and proceed to question 3**
	
	    ☐ Formative evaluation—the evaluation will collect data on implementation progress and provide periodic feedback to project implementers to support project improvement
	
	    ☐ Measuring Fidelity of Implementation—the evaluation will collect data on implementation of the core components of the project, measure fidelity to the proposed theory of change, create and assign fidelity scores, and determine the level of component-level and overall fidelity of implementation
	
	    ☐ Experimental design—the evaluation will collect data on individuals randomly assigned into treatment and control groups; the evaluation will rigorously monitor treatment and control group conditions over the duration of the project
	
	    ☐ Quasi-experimental design—the evaluation will collect data on individuals placed into treatment and comparison groups; the evaluation will rigorously monitor treatment and comparison group conditions over the duration of the project
	
	    ☐ Non-experimental—the evaluation will collect data on the treatment group; a comparison group may be created post hoc (the evaluation will not track comparison group conditions over the duration of the project)
	
	    ☐ Design and implementation of a sampling plan—the evaluation will design a sample that is sufficient for the evaluation's approach, methodology, and analysis framework. The evaluation will identify how to treat sampled data (e.g., establish sample weights and limitations on interpretation of data, if any.)
	
	☒ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Conduct formative evaluation activities
		
		☐ Conduct study of fidelity of implementation
		
		☒ Implement experimental or quasi-experimental design study (evaluator should have advanced background and expertise or training in sampling, research methodology)
		
		☐ Implement non-experimental study (evaluator should have basic background and expertise or training in research methodology)
		
		☒ Design and implement a sampling plan

Westat®

Question	Check the best option…	Possible Third-Party Evaluator Tasks

(3) Can internal staff be sufficiently allocated to perform all evaluation tasks and responsibilities?	☐ Yes—**proceed to question 4**
	
	☒ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Create/pilot test data collection instruments
		☒ Collect data on
		   ☐ Implementation progress
		   ☐ Service Statistics (e.g., numbers served; numbers of services provided)
		   ☐ Fidelity of implementation
		   ☒ Outcomes/impact
		☒ Perform data entry/management
		☒ Conduct data analysis
		☐ Provide performance feedback to project team
		☒ Write reports
		☐ Other: _____

(4) Can internal staff perform all evaluation tasks and responsibilities objectively and without jeopardizing the credibility of evaluation findings?	☐ Yes—**proceed to item 5**
	
	☒ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☒ Collect data on
		   ☐ Implementation progress
		   ☐ Service Statistics (e.g., numbers served; numbers of services provided)
		   ☐ Fidelity of implementation
		   ☒ Outcomes/impact
		☐ Perform data entry/management
		☒ Conduct data analysis
		☐ Provide performance feedback to project team
		☒ Write reports
		☐ Other: _____

(5) NEEDS ASSESSMENT COMPLETED
- **If the answer to all questions is "yes", the project may not need a third-party evaluator.**
- **If the answer to one or more questions is "unsure or no", the project may benefit from hiring a third-party evaluator to perform specific tasks, as identified in this assessment.**

As illustrated in Exhibit 1, once the needs assessment is completed, grantees may find that the project already has qualified and available staff who can perform a number of evaluation tasks. Similarly, the needs assessment can help the project team to identify the areas where additional support may be needed for the evaluation. Grantees can use the items identified in the "Possible Third-Party Evaluator Tasks" column above to develop a list of third-party evaluator responsibilities and tasks.

In the example presented above, the needs assessment indicated the following:
- The project had an evaluation plan that was submitted with its proposal. The evaluation plan received comments from the OSEP review team and requires revisions.
- The plan identified several quasi-experimental elements to the summative evaluation. However, none of the internal project staff have experience in implementing quasi-experimental studies.
- The plan identified two sampling opportunities. However, none of the internal project staff have experience in designing or implementing sampling plans.
- The project needed assistance collecting outcome data, especially observation data. The project also needed assistance with data entry, data quality reviews, and data analysis and reporting.
- The project's internal staff could not provide sufficient objectivity and credibility, especially with regard to outcomes data collection, analysis, and reporting.
- The project needs assessment identified the following tasks that could benefit from third-party evaluator support:
  - Evaluation design with specific attention to:
    - Review, development, or refinement of evaluation questions
    - Identification or review of data collection sources
    - Identification or review of data collection instruments
    - Creation/ pilot testing of data collection instrument(s)

- Design of data collection procedures (Service Statistics and Outcomes/ impact data)
- Design of data entry/ management procedures
- Development of a data analysis framework
  - Guidance and expertise in designing and implementing quasi-experimental studies
  - Guidance and expertise in designing and implementing sampling plans
  - Data collection (outcomes/impact data)
  - Data entry /management
  - Data analysis
  - Report writing

This information can be used to create a scope of work for the third-party evaluator (and for the overall project evaluation), prepare a Request for Proposals (RFP), define contract terms, and establish project management milestones, as discussed later in this document. A blank needs assessment form is located in Appendix B.

1.2   Benefits and limitations of working with a third-party evaluator

Grantees who work with a third-party evaluator should be aware of the potential benefits and limitations of this working relationship, as shown in Exhibit 2. Benefits include the needed skills or objectivity brought to the project by the third-party evaluator while limitations refer to the, often unforeseen or unplanned, tasks or costs associated with monitoring and managing the work of the third-party evaluator.

Exhibit 2. Benefits and Limitations of Working with a Third-Party Evaluator

Benefits	Limitations
Third-party evaluators can:	Third-party evaluators may:
• Bring technical expertise in research methodology, statistics, or related topics to the project team • Provide credibility and objectivity by acting as an external "critical friend" • Take on responsibility for completing some or all of the (formative and summative) evaluation tasks, allowing project staff to focus on project implementation	• Add unanticipated or additional cost to the project • Add to project monitoring and management tasks focused on the work of contractors • Not know the project background or content area as well as project staff • Be less available or accessible, as compared to project staff

It is important to keep in mind that even when the third-party evaluator has a significant role in the project, the Project Director (or Principal Investigator) bears ultimate responsibility for ensuring that the project and its evaluation are carried out as planned and that all OSEP project implementation and reporting requirements are met.

1.3   Determining when to bring a third-party evaluator on board

The decision of when to hire the third-party evaluator affects what the evaluator can and cannot provide to the project. If the grantee engages the evaluator at the start of the project, the third-party evaluator can provide guidance and assistance on foundational aspects of the evaluation—its design and methods. A third-party evaluator brought on board towards the end of the project may have limited ability to revise, modify, or correct decisions made earlier in the evaluation. Exhibit 3 provides an overview of how the timing of hiring can influence what a third-party evaluator can contribute to a project.

Exhibit 3. The Influence of Timing on Expectations for Third-Party Evaluations

	Third-Party Evaluator is Hired Early in the Evaluation	Third-Party Evaluator is Hired Late in the Evaluation
Evaluation Questions	The third-party evaluator can contribute to the development of formative and summative evaluation questions.	The third-party evaluator will be limited in his or her ability to provide recommendations or guidance on formative and summative questions. The evaluator may be able to identify the limitations of existing evaluation questions.
Evaluation Design	The third-party evaluator can contribute to and provide significant guidance on the evaluation's design.	The third-party evaluator will be limited in his or her ability to provide guidance. The evaluator likely will not be able to revise, modify, or correct design features. Evaluation questions may not be addressed at all or answered incompletely by a poor design.
Data Collection	The evaluator can provide guidance and assistance in determining the logistics and methods of collecting data. The evaluator can ensure the data collected are consistent with the evaluation design.	The evaluator may not be able to collect or correct for data that have been collected and are missing, incomplete, or inconsistent. Poor quality data may need to be eliminated from the evaluation.
Data Entry and Management	The evaluator can provide guidance and assistance in creating a data entry and management system that streamlines the movement of raw data into analysis—saving the project team time and money.	The evaluator may have to re-enter, re-code, or re-align data that have already been entered into a [faulty] system—this can be a timely and expensive process.
Data Analysis Framework	The evaluator can provide guidance and assistance in creating a data analysis framework and ensure the framework is consistent with the evaluation design.	The evaluator may be able to review the data collected and establish an analysis framework that works for the data that have been collected, but the framework may or may not be consistent with the evaluation design if it is constructed after all data collections are completed.
Reporting	The evaluator can provide guidance and assistance in the creation of report templates that are aligned with the project's theory of change and funder expectations. Knowing ahead of time what the report will look like may help focus the evaluation over the course of its implementation.	The evaluator can use the data that are available to respond as appropriately as possible to the report template or questions. There is a chance that data may not have been collected correctly, appropriate data may not have been collected, or the evaluation may not be well-aligned with the report template, questions, or requirements of OSEP.

1.4   Developing a third-party evaluator scope of work

Grantees that use a third-party evaluator may create (a) a scope of work for the internal project evaluator and (b) a scope of work for those aspects of the evaluation the third-party evaluator will complete. As outlined above, completing an evaluation needs assessment can give the project team important information that can be used to develop a scope of work. Ideally, a scope of work is written with sufficient detail to clearly convey the specific tasks to be completed, the duration of the work involved, and the required deliverables.

Generally speaking, the third-party evaluator's scope of work contains the same types of elements as the overall project evaluation's scope of work, with a focus on the third-party evaluator's contributions. Creating a comprehensive third-party evaluator scope of work that contains details such as specific tasks, expectations for meetings and communication, expectations for submission of draft and final products, travel and lodging requirements, use of respondent incentives in data collection, etc., will help grantees accurately and adequately budget for third-party expenses. Importantly, the third-party evaluator scope of work also should contain performance management expectations and milestones, so that the grantee can conduct effective oversight of the third-party evaluator. The scope of work can be incorporated into the Request for Proposals (RFP), either as part of the main RFP document, or as an attachment (Section 2.1 discusses development of RFPs).

> **TIP: The evaluation scope of work may need to change over time.** New or unplanned evaluation questions or objectives may arise due to mid-stream changes in project implementation or as a result of evaluation activities. Thus, it may be necessary to change the evaluation's scope of work at some point during project implementation. It is important to keep in mind that these changes may affect the third-party evaluation as well as the overall project budget.

1.5   Creating an evaluation budget

The evaluation budget is one piece of the overall project budget and identifies the resources (e.g., personnel, instruments, incentives, travel) that will be necessary to complete the evaluation. Accurately and adequately budgeting for the overall project evaluation and third-party evaluator tasks can be challenging since the budget is informed by the nature and scope of the evaluation. The logistics of participant recruitment, data collection (e.g., number of sites, number of individuals, location of study sites, and travel requirements), and data management and maintenance (e.g., staffing, software, systems development) also will contribute to the cost of an evaluation. For example, a complex, long-term evaluation with an experimental design may require project staff with advanced training (or a third-party evaluator with the appropriate qualifications), multiple sites, multiple data collections, and complex data analysis—all of which add to the cost of the evaluation. This being said, other budget constraints may affect the total resources available for the evaluation. A third-party evaluator can work with projects to get the best return on the available investment in evaluation.

Following are some of the most common evaluation budget items:
- **Personnel**—including internal and third-party evaluation staff (e.g., evaluation project managers; support or logistics staff; personnel working on the design, deployment, and maintenance of data entry and management systems, such as databases; and staff assigned to evaluation planning, data collection, coding, data quality review, data analysis, and report writing).

- **Non-personnel costs of data collection**—including (but not limited to) costs of purchasing commercial data collection instruments, mailings, and respondent incentives.
- **Travel**—including costs related to recruitment of study sites (if needed) and data collection, such as airfare, mileage, lodging, per diem, etc.
- **Training**—including costs related to appropriate and adequate training of evaluation staff (e.g., in the use of specific observation protocols or assessment procedures).
- **Software and data systems**—including hardware and software costs related to data entry, coding, and management.
- **Security of hard copy and electronic data**—including the costs of secure filing or storage, anti-virus and anti-hacking software, and encryption software.

Given the potential for a great deal of variation in the cost of an evaluation, grantees (and evaluators) may find budgeting to be an iterative process, involving design and review of the evaluation plan, completion of the evaluation needs assessment, and review and revision of the project and evaluation budgets. It is most often better to hire the third-party evaluator early in the process so as to establish a solid foundation for the evaluation and to more effectively manage the costs.

It is relatively common for grantees to underestimate the time and resources that are necessary for a rigorous evaluation, including data collection, entry, and management. Grantees thus may fail to allocate sufficient staff time or resources for the type of evaluation that is desired or required. In these cases, the grantee may consider working with OSEP to reallocate funds so as to provide greater resources for the evaluation, while still adhering to OSEP requirements regarding the use of grant funds. Alternately, the grantee may need to re-group and re-think the evaluation budget if there are insufficient resources for necessary (internal or third-party evaluation) staff. If resources are limited, a third-party evaluator may be able to provide overall guidance to the evaluation and limited assistance to data collection, entry, and management tasks that are performed by internal project staff.

Appendix C presents specific staff and time considerations that may influence the evaluation budget, as well as tips grantees may find helpful when creating an evaluation budget. We recommend that grantees use the table included in the appendix to review common evaluation staffing needs (more specifically, the availability of qualified staff) and to determine whether or not the project has allocated sufficient time to specific evaluation tasks. Appendix D provides time estimates for two common data collection activities: focus groups/interviews and web-based surveys.

> **TIP: Evaluation design drives costs.** Evaluations have costs that reflect the evaluation's methodology and comprehensiveness. Evaluations that are more methodologically intricate (e.g., evaluations that incorporate experimental or control-group designs) often require greater resources. The evaluation budget should include all costs necessary to perform a high-quality evaluation, including—as appropriate—the costs of a third-party evaluator.

In the next section we discuss the process of finding and hiring a third-party evaluator.

Westat®

Part 2. Finding and Hiring a Third-Party Evaluator

After completing a review of the project's evaluation, determining the project's evaluation's needs, and identifying the tasks to be completed by a third-party evaluator, grantees can set about finding and hiring a third-party evaluator. [2] This section describes the three basic steps for finding and hiring a third-party evaluator:

1. Developing a request for proposals
2. Navigating the solicitation and review process
3. Preparing the contract

2.1 Developing a Request for Proposals

The request for proposals (RFP) is a formal invitation to potential third-party evaluators to submit a proposal in response to an identified need—the needs assessment presented in Appendix B can help grantees develop a list of project evaluation needs. The RFP also is a procurement process for assessing which evaluator's experience, qualifications, and approach will best meet the needs of the project.

The structure of the RFP will guide the evaluator's response. Generally speaking, evaluators who respond to an RFP should address all of the RFP's stated requirements. If, however, a grantee doesn't ask for a specific piece of information, the evaluator probably will not provide it in their RFP responses. Thus, it is important for grantees to carefully consider the project's evaluation needs and the expected scope of work (see Part 1) when creating the RFP. Grantees will want to ask specific questions about how the third-party evaluator will respond to the project's evaluation needs and the level of training and experience the evaluator can bring to the project.

A project may have a RFP template or guidelines endorsed by their agency for use in finding and hiring contractors such as third-party evaluators—the project may want to consult with its grants and contracts office before developing a new RFP. When developing a new RFP (in the absence of a template or guidelines), it is common to include the following types of components:

Description of the Project or Program. This section should establish the context for the work to be performed and include a statement of purpose describing the overall objectives of the evaluation contract and the extent of the services desired. Grantees also can use this section to present a brief overview of the grantee organization and of the project to be evaluated. Generally, the more information a grantee can provide about its project or program, the better or more focused the third-party evaluator proposal will be. Grantees may consider sharing the approved project proposal or application or other written materials that discuss the project's goals, theory of change, primary strategies, timeline, etc.

---

[2] Grantees may want to reference the Joint Committee on Standards for Educational Evaluation's *Program Evaluation Standards* for information on what to expect from a high-quality evaluation. The standards can be accessed at: http://www.jcsee.org/program-evaluation-standards-statements

Description of Services Required. This is the heart of the RFP document. This section contains the third-party evaluator scope of work or the list of tasks the grantee identified in the evaluation needs assessment. This section also should clearly convey the specific work needed and the duration of the work involved.

Deliverables. This section provides a complete list of all products, reports, and plans to be delivered by the third-party evaluator and the projected deadlines for the deliverables. If the grantee has an evaluation plan with overall timelines and data collection schedule, they can be used to establish the deliverables schedule. Otherwise, grantees can identify specific products to be delivered and give general guidelines of the expected timing of delivery (e.g., 20 weeks after the contract signing date). It may be helpful to have the evaluation timeline precede the OSEP reporting schedule so that evaluation findings can be used to inform grantee reporting.

Evaluation budget. Grantees may or may not include in the RFP the amount of funding set aside for the third-party evaluator (see Section 1.5 and Appendix C for information on creating an evaluation budget). Whenever possible, it is helpful for third-party evaluators to have access to the proposed evaluation budget or the permissible budget range in order to better understand the scope of the project and to develop a proposal that reflects both project needs and available resources.

Regardless of whether the estimated funding amount is included in the RFP, grantees will want to receive a projected budget from the applicant. It is very helpful for grantees to provide clear and complete instructions regarding how applicants are to break down the budget in table and narrative form. Details will likely include hourly or daily rates for all personnel (and benefits if they are not included in the rates) and equipment and a detailed estimate of expenses, if they are to be reimbursed under the contract. Overhead (indirect) rates also should be expected, if overhead is not rolled into personnel rates. Grantees should identify any budget constraints, such as a cap on indirect rates or restrictions on the type and maximum amount of reimbursable expenses.

Contract Terms and Forms. This section specifies the length, start date and end date of the contract, and any options for renewal. Grantees at larger organizations or agencies, such as universities, may have access to (and be required to use) standard contracting forms, certifications, and assurances that can be attached to the RFP. In these cases, we recommend grantees make inquiries within their agency as to staff or departments that may be able to provide guidance, standard forms and templates, required certifications and assurances, etc.

Proposal Requirements. This section outlines all of the information the proposal should contain. It is a good idea to require applicants to use a consistent structure and format for proposals in terms of line spacing, font size, proposal length,[3] content, and information and documents required. This will help simplify the proposal review process and enable grantees to develop a contract for the selected third-party evaluator, and manage the third-party evaluation once awarded.

> **TIP: Give applicants clear instructions on proposal formatting.** Specify expected font size, line spacing, and document or page limits in the RFP to limit the size of proposals received from applicants. This will force the applicants to focus more specifically on the grantee's evaluation requirements and may help expedite the review process.

---

[3] In particular, grantees should allow a sufficient number of pages, such that the applicant can adequately address all required RFP elements.

For example, grantees may require that proposals contain some or all of the following elements:
- Summary of the applicant's background, history, capabilities, and experience
- Explanation or description of the applicant's philosophy of evaluation—including statements regarding the nature and frequency of interactions or collaboration between the third-party evaluator and the project
- The applicant's proposed approach, which should include a description of how the third-party evaluator will respond to the evaluation needs or scope of work. This section of the applicant's response may contain:
  o Design considerations (e.g., the use of control or comparison groups)
  o Anticipated instruments and data collection methods (including a data collection plan)
  o Plans for data entry and management, data analysis, data security, and reporting
  o Plans for project management and communication
  o Specification of all reports and deliverables, with draft and final submission timelines
- List of tasks to be performed, linked to deliverables and a timeline
- Descriptions of recent and relevant evaluations conducted
- Proposed evaluation team staff, with full-time equivalent (FTE) and qualifications
- Client references
- Proposed budget for the evaluation

Additional information required of applicants may include resumes of key personnel to be involved in the project, previous work samples or reports, and any other information deemed to be important or useful for evaluating applicants' abilities and capacities for completing the work. Grantees may request that evaluators submit the technical approach and budget separately so that the initial review process focuses primarily on the technical quality of the approach and not cost. As noted earlier, grantees may consult with their contracting office or department for agency requirements on how technical and budget proposals should be received.

Proposal Evaluation Criteria. It is helpful for the grantee to clearly describe for applicants how their proposal will be evaluated by the proposal reviewers. A description of the proposal review and selection process should include, at a minimum, a timeline for the proposal review; the method to be used in evaluating proposals including specific criteria and their associated scores; and details of the interview process or other follow-up, if there is to be any. Further, it is helpful for the grantee to provide the estimated or exact date of the final selection and contract award, method of notification, period of negotiation, and any special contract terms and conditions.

RFP Schedule. This section often is used to convey schedule or timeline information such as bidders or pre-proposal conferences (if applicable); date for applicants' letter of intent, if necessary; deadlines for questions and contact information for submitting questions; information about how and when responses to questions will be disseminated; proposal due date; number and type (e.g., hard copy or electronic) of copies of the proposal required; and address for the proposal submission.

> **TIP: Allow time for evaluators to develop high-quality proposals.** There is no rule for expected "turnaround time," or the amount of time between when a grantee issues the RFP and the deadline for submission of proposals. We encourage grantees to allow applicants at least four weeks to prepare and submit their proposals.

A sample RFP schedule might be as follows:
- – March 1: RFP Released
- – March 5-7: Conference calls with potential vendors to discuss RFP
- – March 10: Deadline for questions
- – March 15: Responses to questions posted
- – April 5: Deadline for proposal submission
- – April 7-14: Proposal reviews
- – April 15: Final selection and notification
- – April 25: Contract signed
- – May 1: Targeted project start date

2.2   Navigating the solicitation and review process

While recommendations of colleagues are probably the most common source grantees use for identifying potential evaluators, grantees may also need to post or list the RFP in a variety of places to attract qualified and interested third-party evaluators. Some possible places for posting the RFP include local and national meetings (such as the annual conference of the American Evaluation Association or the Canadian Evaluation Society), websites (such as SIGnetwork), and listservs (such as those linked to the American Educational Research Association or the Council for Exceptional Children), and professional publications and newsletters (such as Education Week).

Other specific sources that may be useful in finding a potential third-party evaluator include the following:

- **Research and consulting firms.** Many experienced evaluators are part of their own or others' research and consulting firms. Suggestions from colleagues and other agencies, listservs, and the Internet can be valuable resources for locating professional evaluation firms. Many federal agencies and national trade or advocacy groups also have technical assistance websites with lists of evaluators.

> **Online resource:** The American Evaluation Association maintains an RFP page for the purposes of linking programs projects needing evaluators to available and interested evaluation professionals:
> http://www.eval.org/p/cm/ld/fid=88.

- **The American Evaluation Association (AEA).** AEA is an international professional organization of more than 7,500 evaluators located across the U.S. and in over 60 foreign countries. Although the AEA does not endorse individual evaluation consultants or firms, a listing of evaluators is provided as a free public service on their website at http://www.eval.org/p/cm/ld/fid=108.
- **Local colleges and universities.** Departments of education, sociology, psychology, social work, public health, and public administration, as well as university-based research centers, are all possible sources of professionals offering evaluation services. If they are unable to personally assist you, they may be able to refer you to others who can.

Grantees will need to identify reviewers to evaluate the proposals after they are received—reviewers may be internal to a project or organization or external, depending upon an organization's rules for contracting with external service providers. While this can be done after proposals are received, grantees may find it helpful to compile a list of reviewers during the RFP development process. At least

one of the reviewers should be knowledgeable in standard evaluation and research methodologies in order to adequately assess the techniques and methods suggested for the evaluation. The list of sources for third-party evaluators also may be used to find external reviewers, if needed, as these sources have access to individuals with sufficient training and expertise to reliably assess a proposal.

We recommend that grantees develop a list of criteria or a rubric for reviewers to use when they read the proposals (and include the criteria in the RFP as suggested earlier). This will help grantees establish a standard scoring system and may help reviewers identify how they should read different aspects of the proposal. Examples of possible focus areas for the scoring system include the following:

- Applicant's qualifications, background, and experience
- Demonstration of understanding about the project context
- Proposed technical approach to the evaluation
- Articulation of how the evaluation addresses the project's goals, objectives, and expected outcomes
- Quality and utility of proposed deliverables
- Adequacy of timeline and evaluator capacity to achieve the proposed work in the specified time frame
- Proposed budget for the evaluation

Grantees may find it helpful to apply point values to these different focus areas. If the grantee requested that the technical proposal and budget were submitted separately, the first round of review may assign points to the technical aspects of the evaluations and then the second round of review would look at the proposed budgets of the two or three top-rated proposals. The proposal with the lowest budget and an adequate technical proposal score would be selected for award.

2.2.1 Assessing the applicant's qualifications, background, and experience

An important part of assessing proposals is to consider applicants' qualifications, background, and experience as they relate to the project's evaluation and evaluation needs. This assessment process also may include contact with or information from evaluator-supplied references (i.e., current or previous clients).

Grantees may find it helpful to use the technical specifications of their evaluation to identify necessary qualifications, provided an evaluation plan exists.[4] Many projects, in fact, work with external evaluators just to understand and complete this review and ensure their RFPs will solicit and receive proposals from qualified professionals.

With regard to basic qualifications, evaluators should possess formal training in evaluation and previous performance of evaluation tasks. It is common for third-party evaluators to provide proposals that include resumes or curriculum vitae that identify their education, previous experience, and a list of technical reports and publications. This information, possibly along with work samples provided in the

---

[4] As noted earlier, grantees that do not have an evaluation plan may find it helpful to work with a third-party evaluator to develop an evaluation plan first and then develop an RFP for the services described in the evaluation plan. The third-party evaluator that developed the plan may or may not be the one who carries out the evaluation. This may add extra costs to the overall evaluation budget, but it will help to ensure that the scope of work and evaluation budget are aligned.

proposal, will be useful in determining whether this individual or group will be a good match for the project. Specific considerations may include the following:

Educational background. There are no licensing or certification requirements and few degree programs in program evaluation. Most evaluators have formal training in research methods but it is usually in a social science discipline such as psychology, education, or social work. Further, some evaluators have expertise in qualitative methods such as interviewing and focus groups, while others are more skilled or experienced with quantitative or statistical methods required to analyze surveys and other numeric data. Grantees should review applicant resumes for information indicating that they have the skills necessary for evaluating their particular project; the evaluation needs assessment found in Exhibit 1 indicates the skills that might be necessary.

Specific Content Knowledge. While it is important for the evaluator to have specific, technical research skills in his or her educational background, grantees also may value an evaluator who is well versed in the project's research or evidence base and theory of change. Does the evaluator need specific experience, knowledge, or information that can only come from working on projects with similar content or a similar technical approach? If so, this should be included in the RFP. If this is not required, it still may be useful for grantees to review the resources and sample work products to determine whether the third-party evaluator will need to spend considerable time getting "up to speed."

Experience. In addition to the third-party evaluator's educational background and training, and any specific content knowledge, the grantee may consider information such as the length and relevance of the applicant's previous evaluation experience. What types of programs have they evaluated? Does their experience include work in similar settings? Another way to check this is to ask for samples of previous work such as reports, published articles, briefs, bulletins, PowerPoint presentations, etc. Are documents written for audiences similar to the grantee's? Is the writing understandable and easy to follow? Are tables and graphs clear and unambiguous? Is the format user friendly?

Evaluation philosophy or approach. Grantees may find it important for the evaluator's philosophical orientation to match the approach required for evaluating their project. Is the evaluator open to collaborative evaluation or an approach that focuses on the intended use of the evaluation's findings? While it is important that the evaluator be independent enough to render an objective assessment of the project, it also is essential that the grantee provide sufficient oversight of the evaluation and its contracted elements. Will the third-party evaluator share drafts of instruments and reports? Will the third-party evaluator discuss with key stakeholders or program staff the proposed design, methods, and data collection procedures, or will he or she dictate the course of the evaluation without the grantee's input?

Location. There are many highly skilled, qualified, and professional evaluators working across the United States, Canada, and internationally. One aspect of the review process may be a consideration of location of the third-party evaluator and its implications for the logistics of conducting the evaluation. An evaluation design and methodology that require the third-party evaluator to be on-site or in close proximity to the project or its intervention sites (e.g., because multiple in-person observations are required) may influence the choice of evaluators and the total amount budgeted in the evaluation for travel.

2.2.2 Checking references

As in any procurement process, the grantee will want to check the client references provided by the applicant. Applicants will understandably provide as references the names of clients with whom they have worked successfully, so the grantee should approach reference checks with an expectation that they will receive positive feedback about the applicant. Even with the bias, however, the grantee can gain important insights into a potential evaluator's strengths, limitations, experience, and philosophy. Therefore, specific questions are more fruitful than general ones. Here are some questions grantees might ask when talking with references:

- Did the third-party evaluator meet the work and project deadlines?
- How responsive and available was the third-party evaluator?
- What is the third-party evaluator's philosophy of evaluation—is he or she relatively collaborative or distanced? Did he or she tend to favor more quantitative or qualitative methods?
- Did any problems arise? Were they satisfactorily resolved?
- Was the third-party evaluator easy to work with? Knowledgeable? Good at anticipating and resolving unexpected problems?
- What was the quality of the third-party evaluator's work and experience?
- Was the third-party evaluator instrumental in completing the evaluation project successfully?
- Were there any billing disputes? If so, how were they resolved?
- Would you use the third-party evaluator again?
- Was the evaluator capable of explaining the evaluation and its results to a range of stakeholders?

2.3 Preparing the Third-Party Evaluation Contract

The next step in hiring the third-party evaluator is to prepare a written contract specifying the roles and responsibilities that will be expected. The contract is a legally binding document that details the activities to be performed, the amount of time to complete the evaluation, and the cost for services. At a minimum the evaluation contract should contain the following:

- **Scope of Work**—This section explains the goals and objectives of the evaluation, outlines what the contracted evaluation will and will not include, lists the required deliverables, and states the responsibilities of the third-party evaluator. The scope of work identifies the evaluation tasks, which include information such as who will collect and enter data, transcribe interviews, analyze data, write up reports, and similar tasks.[5]

> **Online resource:** Evaluation contracting resources may be found at sites such as The Evaluation Center's Evaluation Checklist project:
> http://www.wmich.edu/evalctr/checklists/evaluation-checklists/

- **Definition of grantee responsibilities**—This section specifies the grantee's responsibilities for supporting the evaluation.

---

[5] If the evaluation budget is limited, the grantee may choose to reduce costs by having internal project staff complete some of this work.

- **Data ownership and sharing**—This section outlines who "owns" the data collected by the evaluation and specifies what can be done with those data. This section should identify the need and protocol for developing data sharing agreements[6]—including the specific individuals who are allowed to provide and receive data. Finally, if articles or presentations can be created from the evaluation data collected, this section can specify whether both parties will need to give approval for publication and who will be listed as authors.
- **Nature of the financial arrangement and payment schedule**—This section states the type of contractual arrangement that will be made regarding fees to be paid for third-party evaluator services. It is a good idea to establish a phased payment schedule tied to observable milestones or completion of specific tasks or deliverables, including any penalties for late delivery or non-completion.
- **Timeline**—This section should include major steps and milestones in the evaluation process, data collection schedules, reports or other work components, and billing deadlines.
- **Exit Clauses**—This section states the terms under which the contract legally can be dissolved. The contract also should specify which parties can initiate dissolution—typically, both parties are given the power to dissolve a contract with sufficient written notice.
- **Deliverables and Reports**—This section includes the minimum number and types of reports required as well as the expected content of reports such as a narrative, charts, literature review, comparison to national or other statistics, number of drafts, and who will edit them.

Both the grantee and the third-party evaluator will need to sign the contract. Some important considerations for preparing the contract are discussed in more detail below.

2.3.1 Determining the nature of the financial arrangement

The contract should state what type of contractual arrangement the grantee will use regarding fees to be paid for third-party evaluator services. Such contracts generally fall into one of three categories:
1. **Fixed-Price (also called firm fixed price or lump sum)**—With these contracts, the third-party evaluator provides a defined scope of services for a preset price.
2. **Time and Expense (also called time and materials or a cost reimbursable contract)**—Under these contracts the grantee organization reimburses a third-party evaluator for professional time and expenses based on a formula. These contracts include a fixed daily or hourly billing rate for personnel with separate costs for allowable materials and expenses that will be reimbursed.
3. **Cost Plus (or Cost Plus Fixed Fee)**—For these contracts, the third-party evaluator is reimbursed for staff time and other expenses and is paid a predetermined fee that is a percentage of all allowable costs.

Grantees should check for any existing agency or organizational rules or procedures before contracting for external services, as these may specify what types of contract arrangements are available.

---

[6] Data sharing agreements are common, if not mandatory, when agencies partner to conduct an evaluation. It is important to note that data sharing (or shared use of data) can create liabilities for partner agencies, if it violates the confidentiality of the individuals whose data are represented. Grantees may want to consult with their parent agencies (as applicable) or OSEP to determine if formal data sharing agreements are necessary.

The contract also should include a payment schedule. In general, one of three payment options is used:

1. **Periodic payment schedule**—Fixed amount payments or reimbursements for actual costs are made at regular intervals (e.g., monthly or quarterly), presuming satisfactory progress on the evaluation.
2. **Pay-as-you-go**—Payment for each task or deliverable as it is completed.
3. **Lump-sum**—This may involve a single large payment at the end of the project period or one payment up front as a retainer with the final payment being due at the completion of the project.

Regardless of which payment option is selected, to the extent possible, it is a good idea to tie the payments to major third-party evaluation milestones or deliverables (e.g., a periodic payment schedule can be tied to submission of a monthly progress report). The contract should ensure the Project Director (or his or her designated representative) has the opportunity to review and approve deliverables to ensure that they meet expected standards of quality and utility, prior to generating payment. Projects should expect third-party evaluators to generate invoices that are consistent with the negotiated and contracted budget. Invoices that exceed contracted amounts may be delayed while the project and third-party evaluator discuss the reason for cost overruns and make any necessary or required contract amendments.

2.3.2 Outlining expectations/requirements for human subjects protections and protecting data confidentiality

An Institutional Review Board (IRB) is charged with protecting the rights and welfare of people involved in research. Additionally, the IRB review process provides an assurance that the methodologies employed in the study are appropriate for the nature and goals of the study. There are different requirements for research studies and evaluation studies. In general, research requires a controlled environment—the participants in a research study are not allowed to alter the study conditions. In contrast, evaluation seeks to understand how a project works within an often changing context or environment, in pursuit of a desired goal or outcome. However, it is not always easy to distinguish between research and evaluation because some projects have elements of both. In fact, some evaluators embed a research study within the overall project evaluation.

> **TIP: IRB assistance is available, if needed.** Universities and large research organizations typically maintain their own IRB responsible for approving evaluation activities while smaller evaluation companies often do not maintain an IRB. In those cases, the IRB process can be contracted out to independent firms that maintain and offer these services, often for a fee.

Grantees can consult within their agency or with their OSEP Project Officer for advice on whether or not to seek IRB approval. Ultimately, the decision about whether IRB review is required should be made in concert with an IRB. Grantees may find value in working with a third-party evaluator to complete the IRB process, as evaluation activities may be of particular interest for the review. This type of activity can be incorporated into the third-party evaluator's scope of work.

Grantees may or may not be required to submit their evaluation plan and instruments to the IRB panel. However, a third-party evaluator can work with the project team to ensure the evaluation plan adheres to the three basic ethical principles of human subject protections outlined in the Belmont Report[7]:

1. **Respect for Persons**—Individuals have the right to make decisions about their participation in any part of the evaluation should they deem it harmful in any way. Individuals who do not have decision-making power (i.e., those who are immature or incapacitated) need to be protected from harm.

2. **Beneficence**—Information should be shared with participants, when appropriate, about possible harm that may incur because of participation in the evaluation, and every effort should be made by the evaluator to ensure that participants benefit from the evaluation activities. Benefits may be short- (program improvement) or long-term (impact).

3. **Justice**—Justice ensures that any benefit of treatment or participation is not denied to an individual without good reason or and that burdens are not unduly imposed. Both the burdens and benefits of an evaluation should be distributed equally to all participants or according to individual need, individual effort, societal contribution, and/or merit.

One of the most common forms of human subject protection is the use of informed written consent. For example, when the evaluation design requires collecting or using data from individuals such as students, their parents/guardians, or school personnel, at a minimum the study team will need to obtain permission from the local school district. Further, adult scholars or graduates may need to obtain permission from employers to participate in the evaluation. Project staff may be required to complete training in human subject protections and district protocols for handling and securing data—as may the third-party evaluator.

Protecting Confidentiality and Anonymity of Data. When an evaluation promises confidentiality of information collected from evaluation participants, only the evaluators can identify the responses of individual subjects. However, when reporting findings evaluators must take care that the way the results are presented does not allow for the identification of individuals. For example, reporting the performance evaluation scores of the 15 special education teachers and 1 speech and language pathologist (SLP) employed by a school district, with their specific roles included, effectively identifies the score of the SLP, even if her name or other specific identifying information is not included.

When an evaluation promises anonymity of information collected from evaluation participants, it means that no identifying information is collected from individual respondents (e.g., name, address, email address), or there is no way to link individual responses with participants' identities.

Projects that work with third-party evaluators should clearly establish who will collect, enter, and maintain data and whether the data will be considered confidential or anonymous. Following these decisions, the staff responsible for collecting, entering, and maintaining data will need to establish protocols that protect identities (i.e., maintain the security of personally identifiable information), especially when data will be handled by both internal project staff and the third-party evaluator. See Appendix E for common practices for protecting personally identifiable information.

---

[7] More information about the Belmont Report can be found at: http://www.hhs.gov/ohrp/policy/belmont.html.

2.3.3 Identifying which products the project team expects to receive from the Third-Party Evaluator

The contract should include the specific type, number, and deadline for evaluation deliverables. Evaluation deliverables may include more than a final report—it is common for third-party evaluators to produce the following:

- Updated evaluation plans
- Interim reports or reports on specific evaluation tasks, such as the findings from one data collection event
- Guidance for internal project staff on how to complete aspects of the evaluation not conducted by the third-party evaluator (e.g., guidance on data collections conducted by internal project staff)
- Progress reports on the implementation of the contracted evaluation
- Compilation records that include an executive summary, report for public dissemination, policy brief, and technical report
- PowerPoint presentations
- Final data files with relevant documentation

The grantee also may ask the third-party evaluator for assistance in completing required reports for funders, including Annual Performance Reports; these tasks should be incorporated into the third-party evaluator's scope of work and included in the contracted evaluation budget. If this type of support is requested, grantees should share with the evaluator as much information as possible about the required reports, including report templates or forms, deadlines, and context (e.g., how the report will be used, why the report is required, recent changes in report template or questions). Samples of past reports and communications about past reports also may be helpful, if available. In addition, it is helpful to include the third-party evaluator when funders offer webinars or supplemental information or guidance to projects on how to complete upcoming reports. The third-party evaluator will need sufficient time to complete a draft of the report for grantee review well in advance of the report's deadline. Opportunities for review and communication about these reports should be included in the overall evaluation timeline. Ultimately, the grantee bears the responsibility for reviewing and approving the content of the report, prior to its submission to OSEP.

It is important to recognize that the third-party evaluator may produce findings or interpretations of data that the grantee does not agree with or would frame differently. During contract negotiations it is helpful for the grantee and third-party evaluator to discuss the potential for disagreement about findings and to determine how to resolve disagreements about findings for reporting purposes. This may include requiring the third-party evaluator to provide very specific information on the evaluation's limitations (e.g., the presence of missing data or a low response rate that may limit the generalizability of findings to a large population) or including footnotes in reports that refer to any differences in interpretation of the findings. Finally, it is important to acknowledge who will be held responsible for the reported findings or sections of the report—and to allow that person the final review and confirmation of what is reported.

In the next section we discuss steps grantees can take to monitor and manage the third-party evaluator's work.

Part 3. Monitoring and Managing the Third-Party Evaluator's Work

Once the contract has been signed, the grantee will need to monitor and manage the third-party evaluator's work. This entails establishing a strong working relationship with the third-party evaluator, maintaining regular communication between the project and evaluation teams, and keeping track of evaluation progress—even if the vast majority of the project's evaluation is assigned to the third-party evaluator.

3.1   Establishing a strong working relationship with a third-party evaluator

While it is important for the third-party evaluator to maintain objectivity, the grantee and evaluator can develop a strong working relationship that facilitates completion of evaluation tasks. The working relationship may be especially important when or if the evaluation experiences challenges. Two strategies for establishing a strong working relationship with a third-party evaluator are

- setting reasonable goals and expectations, and
- defining decision-making roles and responsibilities.

In some ways these will have been dealt with during the contract negotiation, but maintaining a good working relationship is an ongoing process.

3.1.1   Setting reasonable goals and expectations

Designing, implementing, and completing an evaluation can be a labor-intensive and time-consuming process, so it is important that both the grantee and the third-party evaluator have realistic expectations for what can be accomplished as well as the amount of time staff can—or should—be available to work on the evaluation project. The processes of conducting an evaluation needs assessment, creating a scope of work for the third-party evaluator, developing an RFP, and entering into a contract all will help grantees to clearly identify the tasks included in the overall evaluation project as well as the specific tasks the third-party evaluator will perform. Additionally, the third-party evaluation budget included in the contract will identify the amount of time the third-party evaluator expects to spend working on the evaluation project and the other resources the evaluator will provide in support of the evaluation.

> **TIP: Don't expect unlimited access.** Most third-party evaluators will have other clients. Therefore, there may be times when the third-party evaluator is unavailable to the grantee. The grantee can expect timely communications but also should recognize that immediate or constant access to the third-party evaluator likely will not be possible. Developing and agreeing upon a communication structure and schedule may help alleviate concerns about access, while limiting the amount of "extra work" the third-party evaluator is asked to do.

In a similar way, the overall evaluation plan will provide guidance on the amount of time and resources the project team should provide to the evaluation. This includes the number of internal staff that will be working on the evaluation (if any), the amount of logistical or financial support the project team will give to the evaluation (e.g., support to data collection efforts), the expected frequency of communication with the third-party evaluator, and any limitations on project resources and staff time. Grantees should make every effort to provide data and important background information about the project to the third-

party evaluator in a timely fashion. Additionally, it may be helpful for grantees to offer guidance and advice to the third-party evaluator regarding the best strategies for gathering data given the specific project context and client population.

The grantee should expect the evaluator to always be respectful of project staff, clients, and stakeholders; to limit any burden placed on staff, clients, and stakeholders; and to show appreciation for their contribution to the evaluation. The grantee also should expect the evaluator to ask for input at various stages of the evaluation and to take the appropriate measures to ensure the security of any data collected. Finally, the grantee should expect timely responses to inquiries, especially inquires about evaluation roadblocks or challenges. In turn, grantees should give the same level of respect to the third-party evaluator, be responsive to the evaluator's requests for support or feedback (as appropriate), and maintain reasonable expectations for the frequency and timeliness of responses to grantee requests for feedback or information from the third-party evaluator.

3.1.2   Defining decision-making responsibility

The overall evaluation plan outlines all of the tasks that must be completed for the evaluation, and the third-party evaluator's scope of work and contract state the specific evaluation tasks that are the responsibility of the third-party evaluator. As mentioned previously, in some circumstances the third-party evaluator will be responsible for conducting virtually all of the evaluation tasks; in others, internal project staff will work with the third-party evaluator to carry out different parts of the evaluation, as appropriate. In both situations, a clear definition of who has decision-making responsibility is extremely important.

Grantees should build clear guidelines for who has decision-making responsibility into the third-party evaluator's contract. It is important to define a decision-making process as early as possible to prevent any misunderstandings or conflicts between the grantee and the third-party evaluator. The process for decision-making may be centralized in a single individual in the project or shared by a committee that includes representatives of the project as well as the evaluation team. When assigning decision-making responsibilities, it is important to consider whether the evaluation (and particularly the summative evaluation) will be considered "independent." The objectivity and credibility that are a benefit of third-party evaluations may be jeopardized if the evaluator is not also allowed decision-making authority for at least some independent tasks. Examples of independent tasks are presented in Exhibit 4.

Often the need for decision-making revolves around unexpected expenses. These may be related to unanticipated challenges in data collection, staffing changes, travel and logistical demands, or infrastructure needs such as computer and software systems. The third-party evaluator should try to anticipate these types of expenses and include them in the initial budget, but the grantee may need to work with the evaluator during the course of an evaluation to make decisions about additional expenses.

Exhibit 4. Keeping Third-Party Evaluations Independent

One of the benefits of working with a third-party evaluator is the credibility and objectivity the external professional brings to the project. Grantees can structure the third-party evaluator's scope of work and contract in a way that maximizes the independence of the evaluation by using the strategies outlined below.

- Hire the third-party evaluator as early as possible and give him or her decision-making responsibility for technical or methodological aspects of evaluation planning and implementation.
- When using randomized experimental or quasi-experimental designs, give the third-party evaluator responsibility for the randomization or grouping processes. This step prevents the project from intentionally or unintentionally skewing participation in either the treatment or control/comparison groups.
- Give the third-party evaluator opportunities to provide input on or carry out the following aspects of data collection:
  - o choosing data collection instruments,
  - o creating and pilot testing new instruments,
  - o training and managing certification of data collectors,
  - o providing independent data collectors, and
  - o monitoring and overseeing the data collection process (when data are collected by internal project staff).
- Allow the third-party evaluator to
  - o take on responsibility for data entry and management,
  - o train project staff on data entry and management, and
  - o conduct data quality reviews and random data entry error checks.
- Ask the third-party evaluator to conduct the data analysis. An independent data analysis by an external professional also is one of the most important ways to guarantee the credibility and objectivity of the evaluation.
- Allow the third-party evaluator to draw conclusions from the data analysis and write the sections of the evaluation report—or other required reports such as APRs—that describe the findings. If the report writing is not included in the third-party evaluator's contract, grantees should at a minimum ask the third-party evaluator to review the sections of any reports that describe the evaluation findings; this practice can support the overall credibility and objectivity of the evaluation.
- When there are both internal and third-party evaluation staff, consider dividing the evaluation work so that the internal staff conduct the formative evaluation and the third-party evaluator carries out the summative evaluation.

3.2   Maintaining regular communication

As mentioned previously, the third-party evaluator's scope of work and the evaluation contract will establish the roles of the third-party evaluator and project staff involved with the evaluation, the evaluation timelines and deliverables schedule, and possibly even expectations for communication between the grantee and the third-party evaluator. Once the evaluation project begins, it is helpful for the grantee and third-party evaluator to agree to a communication structure and schedule that details the logistics of how and when the grantee and third-party evaluator will communicate. In-person meetings; email or other written communications; and telephone-, web-, or video-conferences are all viable communication options. The frequency of communication will likely depend upon the stage of the evaluation; grantees can expect a significant amount of communication at the beginning of the third-party evaluator's contract—when the evaluator is learning about the project and the evaluation— and during reporting periods. In general, grantees should plan for at least monthly communications with the third-party evaluator.

We recommend that grantees formerly designate or identify a liaison on the project team who will be responsible for regular communication with and monitoring of the overall project evaluation, including the third-party evaluator's work. The grantee will want to identify a staff person with the time and ability to
- serve as the evaluation team's primary point of contact for the evaluation, ideally for the duration of the project;
- ensure the evaluation team receives or has access to all of the information required to conduct or complete the evaluation tasks;
- keep track of deadlines and deliverables; and
- review and authorize mid-course corrections if the evaluation seems off track or if problems arise.

Similarly, the grantee can ask that a principal member of the third-party evaluator's team be designated as the point of contact with the project team. This person should have the authority to schedule and time to participate in ongoing communications between the evaluation team and the grantee. He or she also should have the technical capacity and authority to provide advice, guidance, and information or resources to the grantee or internal evaluation team members as needed to support the evaluation.

Good communication between the grantee and the third-party evaluator requires that both parties take an active part in keeping each other informed. On the one hand, it is essential that the grantee liaison monitor the project's implementation and communicate any potential deviations to the third-party evaluator so that any necessary adjustments can be made to the evaluation.[8] Issues to be discussed may include implementation challenges such as low enrollment in project activities or project staff turnover or changes to the scope of the project activities. On the other hand, the third-party evaluator should report potential challenges to implementation of the evaluation to the project liaison, including difficulties recruiting participants for the evaluation, problems making arrangements for data collections, low survey response rates, or evaluation staff turnover. Discussions between project and evaluator staff will be required to remedy or minimize such issues as they arise. For this reason, we urge all grantees to take an active role in monitoring and managing the progress of the evaluation, as discussed below.

3.3   Keeping track of evaluation progress

Even in cases when the third-party evaluator is responsible for most or all of the evaluation, grantees will need to monitor and manage the work of the third-party evaluator. This requires that grantees
- know how the separate pieces of the evaluation fit into the "big picture,"
- have a sense of what should be happening and when (e.g., date, time, and location data collections should be occurring),

> **TIP: Expect the unexpected.** It is exceedingly rare for an evaluation project to unfold without any roadblocks, barriers, delays, or other challenges. Similarly, progress monitoring or formative evaluation activities may point to the need to make changes to the project implementation or to the evaluation itself. It is important for the grantee and third-party evaluator to remain in regular communication throughout the evaluation, to address these issues as they arise and ensure that appropriate and effective remediation is enacted, as needed.

---

[8] Of course, any significant changes to the scope of the third-party evaluation may require a contract modification.

- receive timely and relevant information from project staff and the third-party evaluator, and
- know how to discern higher quality from lower quality work.

Grantees should expect to track the progress of an evaluation and ensure necessary events and tasks are completed on schedule and to expectations. The project staff person designated as the evaluation liaison must be very familiar with the evaluation plan, timelines, and deliverables schedule and should monitor completion of the specific evaluation tasks and contract requirements. The liaison should communicate regularly with the project director (or his or her designated representative) about any issues with implementation of the evaluation so that any problems can be mitigated or resolved. Exhibit 5 provides a sample checklist grantees might use at any stage in the evaluation to track the evaluation's progress.

Exhibit 5. Evaluation Progress Checklist

Questions to Consider	Evaluation is On Target If… (one or more of the following may apply)	Evaluation is Experiencing a Challenge If… (one or more of the following may apply)
Timing and Deadlines		
(1) When did the third-party evaluator get hired?	☐ Third-party evaluator was hired prior to the project's first day of implementation ☐ Third-party evaluator was hired on schedule (according to the project management plan).	☐ The process for hiring the evaluator is more than one month delayed. ☐ The third-party evaluator is in the process of being hired or has not yet been hired.
(2) Is the evaluation meeting its deadlines?	☐ The evaluation has met all deadlines. ☐ The evaluation has met most deadlines, with 1 or 2 delays.	☐ The evaluation has not established any deadlines. ☐ There have been 3 or more delays in meeting deadlines.
Evaluation Planning and Design		
(3) Is the evaluation design complete?	☐ The plan contains most or all key elements.	☐ The evaluation does not have a plan or design. ☐ The evaluation plan is missing 3 or more elements.
(4) Was the third-party evaluator involved in the evaluation's design?	☐ The evaluator was involved with the evaluation design from the beginning. ☐ The evaluator reviewed a preexisting evaluation plan and is satisfied with it. ☐ The evaluator helped revise and refine a preexisting evaluation plan.	☐ The evaluator disagrees with or has expressed reservations or limitations for some or all elements of the plan such as evaluation questions, evaluation approach and methodology, data collection instruments, response rate, etc.
(5) Were evaluation questions finalized prior to data collection?	☐ The evaluation questions were finalized before data collection began	☐ The evaluation questions were being developed or revised after data collection began.
(6) Was the data analysis plan finalized prior to data collection?	☐ The data analysis plan was finalized before data collection began.	☐ The data analysis plan was being developed or revised after data collection began.
Communications		
(7) How frequently does the grantee communicate with the evaluator?	☐ Communication occurs according to the communication schedule. ☐ Communication occurs at least once a month.	☐ There is no communication schedule. ☐ Communication with the evaluator is sporadic. ☐ Communication is infrequent--it is difficult to get in touch with the evaluator.
(8) Is the grantee receiving key deliverables?	☐ Key deliverables are submitted on time and in the format expected. ☐ Key deliverables are generally submitted on time, with 1 or 2 delays or changes.	☐ The evaluation has not established any key deliverables. ☐ Products or deliverables are in arrears or do not meet expectations.
Data Collection		
(9) Does the evaluation have access to the desired participants and sites for data collection?	☐ Desired participants and sites are accessible (i.e., the evaluator can collect data on or from the participant or site).	☐ Only some of the desired participants or sites are accessible to the evaluator. ☐ None of the desired participants and sites are accessible to the evaluator.
(10) Are data collectors considered highly reliable?	☐ Data collectors were trained by qualified staff to high reliability. ☐ Data collectors all produced high inter-rater reliability (i.e. raters produced scores, observations, or assessments that are consistent with each other, indicating that data collectors all were collecting or scoring data in the same way). (*Note: inter-rater reliability may be assessed during training, during a pilot test, and/or during formal, non-pilot test, data collection.*) ☐ Data collectors were certified by instrument development staff.	☐ Data collectors were not trained. ☐ Data collectors failed to receive certification or high inter-rater reliability.

Questions to Consider	Evaluation is On Target If… (one or more of the following may apply)		Evaluation is Experiencing a Challenge If… (one or more of the following may apply)	
(11) Are desired populations responding to data collection (e.g., completing data collection) in sufficient numbers?	☐	The evaluation is on track to achieve a high (e.g., ≥70%) response rate on most data collection events.	☐ ☐	The evaluation is struggling to achieve a high response rate on one or more data collection events. The evaluation has failed to collect any data on one or more data collection events.
(12) Is the evaluation able to collect the data needed to answer the evaluation questions?	☐ ☐	Instruments are designed for or well-aligned to the evaluation questions. Most forms and instruments are completed or nearly completed.	☐ ☐ ☐	Instruments were chosen prior to finalization of evaluation questions. The evaluator did not review or approve final instruments. Many forms and instruments are returned with missing data.
Data Entry, Management, and Quality				
(13) Is data transfer and storage secure?	☐ ☐	Data collectors and evaluation staff follow a protocol for handling and transferring data securely. Data are kept in secure locations by qualified staff.	☐ ☐	There is no protocol for how to securely handle, transfer, or store data after collection. Data remain in an unsecure location with data collectors after a data collection event.
(14) Is there a system for organizing raw data? *Note: This may include specifications on how to compile data into a database, how to merge raw data files, etc.*	☐	Staff follow a protocol (e.g., instructions or guidelines) for organizing raw data after collection.	☐ ☐ ☐	There is no protocol for organizing raw data after collection. There is no means of checking which data have been collected or transferred to data entry staff. There is no means of tracking a data element used in analysis back to its "raw" form (e.g., data that have been re-coded or transformed for analysis cannot be tracked back to original format or value).
(15) Is there a system for cleaning, entering, and coding data? *Note: Cleaning data refers to the process of reviewing and resolving missing, incomplete, or inconsistent data*	☐ ☐ ☐ ☐	Trained staff check data for completeness, consistency, and legitimacy. The evaluation has a data definitions and coding manual. Trained staff perform data entry. Trained staff code all data as needed.	☐ ☐ ☐ ☐	The evaluation does not have a system for cleaning data. The evaluation does not have a data definitions and coding manual. There are no staff responsible for or trained in data entry. There are no staff responsible for or trained in data coding.
Data Analysis				
(16) Is analysis proceeding with high quality data?	☐ ☐	Staff follow a protocol for reviewing data quality prior to analysis. Staff create a file of high-quality data to be used for the analysis (which is separate from the original file or raw, uncleaned data).	☐ ☐	There is no protocol for data quality review. Analyses proceed before data collection is completed.
(17) Is there a system for making data available for analysis?	☐ ☐	Evaluation staff (either internal or external) follow a protocol for data export or transfer of data spreadsheets or files. Data always are transferred securely.	☐ ☐	There is no protocol for data export or transfer of files among evaluation staff (e.g., the staff who collect data, enter data, conduct data quality checks, and analyze data). Data are not transferred securely.
(18) Is there a comprehensive analysis framework that is guiding data analysis?	☐	Analyses are aligned with the evaluation questions.	☐	Staff "mine" the data for possible analyses.

Questions to Consider	Evaluation is On Target If… (one or more of the following may apply)	Evaluation is Experiencing a Challenge If… (one or more of the following may apply)
Reporting		
(19) Are report templates in place?	☐ The evaluator works with the project to develop report templates. ☐ The evaluator uses funder's report templates.	☐ The evaluator has not received OSEP requirements or provided report templates..
(20) Are reported findings reliable and credible?	☐ Reporting of findings is based on high-quality data and analyses conducted by the third-party evaluator. . ☐ The third-party evaluator has authority to review and approve findings based on high-quality analysis conducted by the internal evaluator.	☐ The evaluation findings are not based on high-quality data or analysis. ☐ The grantee wants to "approve" the findings from the third-party evaluation before they are reported. ☐ The grantee will not allow the third-party evaluator to review some or all of the report, including findings reported by the internal evaluator.

3.4   Addressing problems with the third-party evaluation

Grantees may find that their contracted evaluation is not going well (e.g., has many logistical challenges or is having difficulty collecting data) or that the third-party evaluator is not completing contracted tasks. The recommended first course of action is to create an action plan that specifies how the evaluation should be brought back on track. Grantees also should ensure there is frequent communication with the third-party evaluator regarding the action plan and adequate completion of evaluation tasks. The strongest strategy for avoiding problems with an evaluation—whether internal or external—is to comprehensively plan the evaluation prior to its implementation and to invest in ongoing and regular communication with all evaluation staff and especially the third-party evaluator.

3.4.1   Dissolving the third-party evaluation contract

If it is not possible to resolve on-going problems with a third-party evaluator, grantees may investigate whether or not they have the option to end the contract. There are several considerations attached to this decision:

- Has the grantee exhausted all opportunities to monitor and manage the contracted evaluation? For example, has the grantee requested or increased the frequency of communication, or asked for more frequent updates on work products? Has the grantee placed the contracted evaluator on a performance improvement plan?
- What are the contract terms? The contract should clearly state the terms under which it legally can be dissolved, such as 30-days written notice. Grantees must be aware of these terms if they decide to end the contract before its term expires. If possible, grantees may find it necessary to consult with their organization's grants and contracts staff or their Board of Directors before initiating a contract termination.
- What would be the impact on the project and its evaluation of ending the contract early? Is the third-party evaluator in charge of many or most aspects of the evaluation? Does the evaluator have possession of critical documents and data? If so, the grantee may want to receive all documents, data, and evaluation materials before dissolving the contract.

It is important for grantees who are contemplating terminating a contract to carefully weigh the costs and benefits of this decision. After terminating the contract, the grantee will likely need to replace the contracted evaluator with either a different third-party evaluator or internal staff who can complete the evaluation. These staff will need to be trained in the project and the evaluation plan, which now are in the process of being implemented. Further, new evaluators may bring different perspectives to the evaluation that do not align well with the original work. If the grantee shifts responsibility for the evaluation to existing project team members, this comes at the cost of additional burden to these team members or the need to hire or re-assign other staff to assist these team members in their other duties. At a minimum, the grantee will lose time and have to spend resources bringing a new evaluation team up to speed.

Part 4. Concluding the Evaluation Project

At the end of the third-party evaluation (which may or may not correspond with the end of the overall project), there are four general tasks that grantees should carry out:

- verify that all contracted tasks have been completed,
- ensure that all evaluation documents have been submitted,
- establish what do to if there are future requests for information related to the evaluation, and
- ensure that data transfers have been completed and that any confidential files have been destroyed (as appropriate).

Identifying and responding to outstanding contractual items. As discussed earlier, the third-party evaluator's contract should specify those tasks and products the evaluator must perform and deliver to completely fulfill the terms of the contract. It is helpful for the grantee to incorporate these items into the ongoing management and monitoring of the contracted evaluation

> **TIP: Don't remit final payment until all third-party evaluation tasks are complete.** As a general rule of thumb, a project and its evaluation can be considered closed after OSEP has received and approved the final report, generated a final payment, or otherwise communicated to the grantee that the project is ended. This being said, a grantee's contract with a third-party evaluator may end prior to the official close of the project. We recommend making sure all evaluation tasks are complete and deliverables submitted before sending the final payment to the third-party evaluator.

and to complete a final review when the third-party evaluator says that the contracted evaluation is complete. In the event that some of the contracted tasks have not been completed, the grantee may withhold the evaluator's final payment until all contractual items have been fulfilled or the grantee and evaluator may negotiate a contract amendment or modification.

Ensuring the receipt of all evaluation documents. The grantee can expect to receive hard and electronic copies of all deliverables created by the evaluator, under the terms of the contract. The grantee may ask for other documents or products—however, the third-party evaluator is liable only for those products that were included in the contract. A grantee that desires additional products or deliverables may need to negotiate an amendment to the contract to ensure that these products are made available.

Establishing a chain of communication in the event of future information requests. Grantees may receive questions or requests for information or evaluation products after the project and its evaluation have concluded. There may be questions or requests that the grantee cannot answer—especially questions that target technical or methodological aspects of the third-party evaluation. The grantee may contact the third-party evaluator for assistance in responding to these questions or requests. If the third-party evaluator cannot be reached or cannot assist in responding to questions or requests, the grantee may need to reach out to other evaluators or qualified staff or may not be able to respond to the question or request.

Completing the appropriate transfer or destruction of data and files. If the evaluation plan called for the anonymous or confidential collection of data, these conditions apply even after the evaluation and project are concluded. If the grantee and third-party evaluator agree that the grantee will receive all confidential data and primary documents after the evaluation project concludes, the third-party evaluator should remove all identifying information and arrange for secure transfer of the data files prior to data and document transfer—this is to say, the grantee and the third-party evaluator should

agree on the procedures to use for transferring data and documents, prior to the actual transfer. Depending on the terms of the data sharing agreement, data files may include both raw (i.e., data in its original form) and cleaned data (i.e., data that have been checked for completeness, consistency, and legitimacy). If the grantee and third-party evaluator agree that the evaluator will maintain all data and primary documents, the evaluator should expect to securely store the data and documents until the prearranged time for destruction of the data files (e.g., projects may be required to keep data and files for three or more years after the project end date).

Exhibit 6 contains a checklist of evaluation "close-out" tasks that may be implemented by both the project evaluation liaison and the third-party evaluator.

Exhibit 6. Evaluation Close-Out Tasks

Close-Out Task	The grantee can…	The third-party evaluator can…
Complete all contracted items, including payments	☐ Review the contract and identify any outstanding contracted items. ☐ Develop a list of outstanding items to respond to, the order in which they can be addressed, and a timeline/deadline for addressing the items. ☐ Process any and all final invoices, per contracted terms.	☐ Review the contract and identify any outstanding contracted items. ☐ Submit any outstanding products or deliverables to the grantee. ☐ Submit a final request for payment or invoice, with a deadline for payment.
Ensure the receipt of all evaluation documents	☐ Review the evaluation documents received to-date, compare to the list of contracted deliverables, and create a list of documents that are outstanding. ☐ Remind the third-party evaluator that final payment may be delayed until all contracted documents are received.	☐ Review the evaluation documents produced to-date, compare to the list of contracted deliverables, and create a list of documents that are outstanding. ☐ Confirm with the grantee the documents that have been delivered to-date. ☐ Create a zip file, CD, or "thumb drive" with all evaluation materials and products. Ensure the secure transfer of the zip file, CD, or "thumb drive" to the grantee.
Establish a chain of communication for future information requests	☐ Establish, in writing, who will be responsible for handling future information requests. ☐ Provide contact information, including back-up contacts or contact information, in case the grantee cannot be contacted at his or her primary address, telephone number, cell phone number, or email address. ☐ Provide for or specifically exclude "extraordinary circumstances", in the event a funder or interested party cannot establish contact through the primary or secondary sources. *For example, if there are extraordinary circumstances, does the third-party evaluator have permission to respond to requests or should the interested party be routed to someone else in the grantee's agency?*	☐ Provide the grantee with primary and secondary contact information. ☐ Establish the length of time the third-party evaluator will be available for follow-up or requests for information.
Complete the transfer of or destruction of data and files	☐ Compile a list of all data and files the project will maintain in-house. Similarly, compile a list of all data and files that will be maintained by the third-party evaluator and a list of all data and files that are to be destroyed. *Note—the project may be required to maintain evaluation records (including data) for three or more years—OSEP and the grantee's agency may have different requirements.* ☐ Assign a staff person to oversee secure transfer of data and files. ☐ Assign a staff person the responsibility of secure in-house storage or destruction.	☐ Ensure all data and files that are transferred to the project or maintained in-house are scrubbed of identifying information (as appropriate) and transferred securely and completely. Receive a written receipt that details all data and files received, the date received, and the person who received the data and files. ☐ At the appropriate time, contract with or utilize available data and file destruction services. Maintain records of how and when data and files were destroyed. Provide copies of these records to the grantee, as required.

APPENDICES

Appendix A. An Evaluation Primer

This section presents a brief overview of evaluation planning basics.[9] Evaluations typically feature three components targeted at three distinct lines of inquiry: progress monitoring, formative evaluation, and summative evaluation.

Progress monitoring. Progress monitoring is a tool that helps project staff and evaluators to gauge implementation and its effects on the targeted population (or populations). Periodic monitoring of benchmarks helps identify areas of strength (benchmarks are attained or surpassed) or weakness (performance falls short of benchmarks), and provides implementers with evidence to make decisions regarding project improvements. Benchmarks relate to both project implementation (e.g., the number and types of participants that the project should reach each year or the number and types of services delivered and received by the target audience) and outcomes (e.g., the percentage of visually impaired students who are expected to enroll in postsecondary education upon graduation from high school).

To make the most of progress monitoring, evaluators should review project documents and work with project staff and implementers to define benchmarks as early as possible in the evaluation planning process; more detailed questions that guide the progress monitoring will depend on the benchmarks defined. Data for progress monitoring (including data to track costs or expenditures) should be planned for and collected periodically throughout the evaluation (e.g., state assessments are conducted annually, curricula-related assessments may occur at the end of each unit or more frequently) and are primarily quantitative.

During the early stages of project implementation, progress monitoring may serve as the primary gauge of whether the project is moving towards achieving its objectives. This information not only contributes to the evaluator's overall understanding of the achievements of the project (a goal typically served by summative evaluation), but it also helps the evaluators inform the project's leadership about needed changes or adjustments along the way (a goal typically served by formative evaluation). Additionally, progress monitoring data can be used to assess efficiency and guide cost containment strategies, if needed. Depending on the nature of the metrics and benchmarks, progress monitoring also may be helpful for the completion of Annual Performance Reports, or other required reports.

Formative evaluation. The purpose of a formative evaluation is to document and provide feedback to project staff on how well a project's components and strategies are being implemented. Evaluators may function as "critical friends" of the project and incorporate information obtained through progress monitoring (such as percent of benchmarks achieved to date or the project's ability to achieve desired milestones in project implementation, service use, or participation) and stakeholder data collections to answer questions related to whether the project is complying with established rules or policies, proceeding as planned, and producing the expected outputs or short-term outcomes. Formative evaluations also may assess the fidelity of implementation of the project model and a project's short-term impact.

Summative evaluation. The overall purpose of summative evaluation is to evaluate the effectiveness and efficiency of the project in achieving its outcomes or goals. This is accomplished in part through the progress monitoring and formative evaluation components described above, but summative evaluation questions typically require an investigation of the extent to which a change has occurred, the factors

---

[9] For more information and detailed guidance on evaluation, see Lammert, Heinemeier & Fiore (2013).

associated with a change, or the measurement of change among different populations. Further, a summative evaluation can establish a project's impact on the populations served or affected by the project, including students. An important role for the summative evaluation can be the determination of the unique contribution of the project to the desired change. As such, summative questions are best informed when there are comparison data (e.g., for treatment and control groups) to give the evaluator an idea of the counterfactual—that is, what would have happened if the project had not been implemented.

The Evaluation Plan

The evaluation plan outlines the questions that the evaluation is designed to address, defines and enumerates the data collection and analysis tasks to be accomplished, identifies the specific duties to be performed by staff on the project and evaluation teams (including the third-party evaluator), and lays out the expected outcomes of the work. In this way the evaluation plan can be considered an overall **scope of work** for the evaluation, although different evaluation tasks and responsibilities may be assigned to staff internal to the project team or to a third-party evaluator. An evaluation plan is a critical part of the evaluation and should be created prior to or soon after a project begins.[10]

An evaluation plan typically contains the following sections:
- **Introduction**—a review of the project and its theory of change[11]—often represented as a logic model—as well as contextual factors to be considered during the evaluation.
- **Evaluation questions**—commonly, an evaluation seeks to answer questions about a project's overall effectiveness and efficiency, as well as respond to specific questions defined by a funder. An evaluation may contain both formative and summative evaluation questions. Formative questions focus on the extent and quality of project implementation while summative questions focus on the extent to which a project achieved its goals (i.e., outcomes).
- **Methodology**—specific details regarding the evaluation design, data collection, data entry and management, data analysis, and reporting.
- **Timeframes and Responsibilities**—a timeline for the overall evaluation project as well as timelines for specific evaluation events such as data collection, analysis and reporting. This section also can identify the project staff that will complete different evaluation tasks and include the roles and responsibilities of the third-party evaluator.
- **Deliverables**—the reports and other products to be generated from the evaluation.
- **Budget**—the costs of conducting the different evaluation activities, including staff hours, travel costs, materials, etc.

---

[10] Unfortunately, this does not always happen or is not always possible. Grantees should bear in mind that if the project has been in operation for a few years without an evaluation plan, options may be very limited in terms of the outcomes an evaluation is able to capture. Or, if the project evaluation plan did not contain a design with both intervention and control/comparison groups, for instance, the project evaluation may not be able to establish that the project had a unique impact, or that project activities can be linked to changes in its clients or participants that would not otherwise have occurred.

[11] A theory of change is a description or explanation of how a project plans to use its strategies, activities, and services to meet or address a specific need or needs. Often, a theory of change links needs and services/project strategies with project deliverables and outcomes or long-term goals.

In the sections that follow we briefly discuss several aspects of evaluation planning that may benefit from the contribution of a third-party evaluator. These include the following:

- Developing evaluation questions
- Identifying high quality outcomes
- Selecting an evaluation design
- Creating a data collection plan
- Constructing a data analysis plan
- Developing a comprehensive timeline
- Creating an evaluation budget

These sections also contain examples, created using the fictional Anywhere State Speech and Language Pathologist (SLP) Support project (see Exhibit A1 below), that may help illustrate possible applications of evaluation design and planning concepts.

Exhibit A1. The Anywhere State Speech and Language Pathologist (SLP) Support Project

> Schools and other service providers have a need for highly qualified speech and language pathologists (SLPs) who can work with bilingual secondary students with disabilities. These SLPs must be proficient in evidence-based practices. The need for SLPs from diverse racial/ethnic backgrounds and those with disabilities is especially strong. The Anywhere State SLP Support project responds to these needs by educating, training, and graduating SLPs who: (a) demonstrate high- quality skills and performance in appropriate work settings and (b) demonstrate success with children/ students.

Developing evaluation questions

The evaluation questions reflect the goals of the evaluation and often reference the primary outcomes associated with the project. It is helpful for grantees to develop questions in collaboration with key project staff and stakeholders, including OSEP, as appropriate. Grantees also may work with third-party evaluators to develop or refine evaluation questions, which should be based on a thorough understanding of the project's overarching objectives and theory of change.

There generally are three types of evaluation questions: progress monitoring, formative, and summative. Progress monitoring questions address the rate or pace of project implementation and the extent to which the target population is served. Progress monitoring questions tend to rely on quantitative or numeric data and often incorporate the use of benchmarks or milestones such as deadlines for implementation, enrollment to capacity, and so on. Examples of progress monitoring questions include the following:

- Did the project start services on time or as expected?
- Is the project fully enrolled? How many individuals or participants are receiving services?
- Is the project moving towards achieving its objectives?
- Are any changes or adjustments needed?
- Is the project using resources efficiently?

In comparison, formative evaluation questions focus on the project's processes and address the extent to which (and how well) the project is being implemented according to design. This includes questions related to whether the project is complying with established rules or policies, proceeding as planned, and producing the expected outputs—sometimes called monitoring questions—as well as questions

addressing the fidelity of implementation of the project model. Formative types of questions include the following:

- Are key inputs being utilized as expected (e.g., as designed)?
- To what extent are the project's key components being implemented with fidelity?
- To what extent is the project reaching its target population?
- Is the project achieving key outputs at desired or targeted levels?
- Are the project participants receiving an effective level of services?
- How are activities being received by participants? (What do participants and stakeholders like about the current program?)
- What changes do participants and stakeholders suggest should be made?

Finally, summative evaluation questions target the extent to which a project achieves its expected outcomes. These questions typically require an investigation into what, if any, change has occurred, the factors associated with a change, or the measurement of change among different populations. Summative questions are best informed when there are comparison data (e.g., for treatment and control or comparison groups) to give the evaluator an idea of the counterfactual—that is, what would have happened if the project had not been implemented. Examples of summative types of questions include the following:

- What outcomes (expected and unexpected) have occurred?
- What expected outcomes have not occurred?
- Where is change the greatest?
- To what degree have outcomes occurred?
- What is the unique contribution of the program to the observed change?
- What is the cost/benefit of these outcomes?
- To what extent do the same outcomes occur in treatment and control or comparison groups?

Identifying high quality outcomes

The quality of an evaluation depends in large part on the quality of the outcomes used to demonstrate project effects as well as the rigor of the data collection and analysis procedures used to evaluate the outcomes. The best outcomes are rigorous, have a high degree of utility, and are informed by high-quality data.

One approach to creating high quality outcomes is to determine if they are SMART: "*Specific and Clearly Stated, Measurable and Based on Data, Attainable and Realistic, Relevant to Student Achievement and Performance, and Time-Bound*" (Doran, 1981). More specifically, SMART outcomes are:

- **Specific (and Clearly Stated)**—Outcome statements should reflect the need or problem the program is responding to, in sufficient detail, such that the reader can determine if the originating need or problem has been addressed.
- **Measureable (and Based on Data)**—All outcomes must be assessed using standardized procedures for the collection, aggregation, and analysis of data that are relevant to the outcome. It is important to determine *a priori* if the data are, or can be, available for use in outcome assessment. If high-quality, outcome-specific data are not available through existing data collection and management systems, the evaluator must determine if and how the data can be collected. Ultimately, if high-quality, outcome-specific data are not available and cannot be collected for a specific outcome (within the limits of the available time and

resources), it is better to select another outcome for which such data *are* available than to report results for an outcome that are based on poor-quality data.

- **Attainable (and Realistic)**—Good outcomes reflect changes that are achievable within a given timeframe.
- **Relevant (to Student Achievement and Performance)**—Relevant outcomes address the degree to which the underlying need or problem has been alleviated, reflect needs and problems of consequence to communities and schools, and generate information for future decision-making. The concept of "relevant" outcomes captures the desired changes that can occur after services are deployed.
- **Time-Bound**—Good outcomes are achievable within a defined period of time. Direct, or short-term, outcomes generally can be observed within one program year, whereas longer-term outcomes can take the entire grant period or more to assess and achieve. Intermediate-outcomes generally represent changes that occur between the direct and long-term outcomes and, because they are likely to fall within the time boundaries of the grant period (e.g., 3-5 years), may be the most distal outcomes on which the evaluation can realistically focus.

Another important part of using high-quality outcomes is identifying high-quality outcome measures that can be used in the evaluation. The U.S. Department of Education's Institute of Education Sciences' What Works Clearinghouse (WWC, http://ies.ed.gov/ncee/wwc/) has stated that high-quality outcome measures share four characteristics:

- **Face validity**—The measure must appear to be a valid measure of the outcome (e.g., a reading fluency test should not be used to measure mathematics outcomes).
- **Adequate reliability**—Reliability is informed by the type of outcome measure (e.g., test score, scale, observation measure) and whether or not the measure is based on a standardized test or state- required achievement test.[12]
- **Lack of over-alignment with the study intervention**—The measure must not be designed or administered in ways that are specifically aligned to an intervention (e.g., a student should not be tested for reading fluency using the 50 words that she practiced reading aloud repeatedly during an intervention).
- **Consistency of data collection across groups**—The outcome data must be collected using the same rules or procedures across groups of participants in the study (e.g., student outcome data should not be collected by special education teachers as part of their regular classroom activities in one school and by graduate research assistants in a pull-out activity in another school).

Finally, high quality measures are **consistently defined across groups**. For example, if an evaluation is looking at the performance of transition specialists and wants to use rates of college enrollment among students who are deaf/hard of hearing (DHH) as an outcome, it is important to know whether the college enrollment rate for DHH students is calculated in similar ways across schools and school districts.

High quality questions incorporate high quality outcomes and require rigorous data collection and analysis. High quality questions are structured so as to produce testable hypotheses. Exhibit A2 presents the five questions created for the Anywhere State SLP project. At least two hypotheses can be created for each question, which are testable through the rigorous collection of service statistics and other evaluation data. Note, for each question, the evaluator will have to carefully define the

---

[12] See What Works Clearinghouse for specific guidance on reliability metrics: http://ies.ed.gov/ncee/wwc/

conditions, context, and findings that allow the project to reject or fail to reject a hypothesis.

Exhibit A2. Anywhere State SLP Support Project Evaluation Questions and Hypotheses

	Evaluation Question	Hypothesis 1	Hypothesis 2
A.	**Are we collecting valid data on the performance of our graduates?**	Valid data are being collected.	Valid data are not being collected.
B.	**Do the expected number and percentage of graduates work in appropriate settings for 3 years?**	The expected number and percent of graduates work in appropriate settings for 3 years.	The expected number and percent of graduates do not work in appropriate settings for 3 years.
C.	**To what extent do graduates exit the program with the skills and knowledge necessary to perform at a high level?**	Graduates exit the program with the skills and knowledge necessary to perform at a high level.	Graduates do not exit the program with the skills and knowledge necessary to perform at a high level.
D.	**To what extent do graduates demonstrate in the workplace the high quality skills and knowledge needed to improve outcomes for children with disabilities?**	Graduates demonstrate high quality skills and knowledge.	Graduates do not demonstrate high quality skills and knowledge.
E.	**To what extent do graduates demonstrate success with children/students with disabilities?**	Graduates demonstrate success with children/students with disabilities.	Graduates do not demonstrate success with children/students with disabilities.

Selecting an evaluation design

Evaluation designs fall into one of three categories: experimental, quasi-experimental and non-experimental. Evaluators can use one design for their evaluation or can combine two or more types of evaluation designs into an overall "mixed-method" evaluation plan; the term "mixed-method" also may refer to the collection of both quantitative and qualitative data. The selection of the evaluation design depends on (a) the questions that the evaluation is trying to answer; (b) the resources available for data collection, management, and analysis; (c) the availability and feasibility of control or comparison groups; and (d) the availability of data to measure outcomes. When considering which evaluation design to choose, it is important to think about issues related to the validity of the study—or issues such as confounding variables, alternate explanations for observed results, and generalizability of findings. While experimental designs are considered to be the most rigorous, OSEP does not typically require this evaluation design.

Experimental Design: In randomized experimental designs, or randomized controlled trials (RCT), the researcher controls all aspects of the intervention[13] such as
- Manipulation of the independent variable;
- Fully randomized assignment of participants to treatment groups; and
- Controlling for possible confounding variables.

---

[13] See Shadish, Cook & Campbell, 2002 and Dimitrov, 2010.

Randomization can take two forms: **random selection** of individuals to participate in the study and **random assignment** of participants to treatment and control groups. A fully randomized study includes both random selection *and* random assignment. **Confounding variables** are those that are correlated (either positively or negatively) with both the dependent and independent variable, thereby affecting the study's ability to clearly associate an intervention or project with an observed outcome.

Quasi-Experimental Design: Generally speaking, the main difference between randomized experiments and quasi-experiments, or quasi-experimental designs (QEDs), is that QEDs do not feature random assignment of study participants to treatment groups.[14] Researchers still may control other aspects of the intervention such as

- selecting and scheduling measures,
- execution of non-random assignment,
- selection of the comparison group, and
- treatment schedule.

> **TIP: Be sure the project has staff trained in carrying out the specific type of evaluation design required to answer all evaluation questions.** Experimental and quasi-experimental designs often require statistical support, or the input and participation of trained and qualified statisticians—the third-party evaluator may or may not have sufficient statistical training for the project design. Similarly, projects that rely heavily on qualitative data such as case studies and interviews may want to work with a staff member or third-party evaluator who specializes in or has training and experience in qualitative methods.

QEDs generally are easier to conduct than experiments, while still providing a measure of methodological rigor. However, QEDs provide less support for counterfactual inferences—that is, making inferences about what would have happened if the intervention or project had not been implemented—than RCTs since the lack of random assignment to groups means that the treatment groups may differ in systematic ways that may affect the outcomes.[15] Consequently, researchers conducting QEDs should outline as many *plausible* alternative explanations for the study results as possible "*and then use logic, design, and measurement to assess whether each one is operating in a way that might explain any observed effect.*"[16] Of course, this has an impact on the complexity of the study design and, by extension, the difficulty of study implementation, so the study team will need to decide whether ruling out a plausible alternative explanation is worth the time, money, and effort required.

Non-Experimental Designs: This category of evaluation design may include case studies, descriptive studies or surveys, correlational studies, and ex post facto studies (i.e., studies that take place after the fact using secondary data). It is possible to investigate a presumed cause and effect in a non-experimental study, but the structural features of experiments that help to rule out possible alternative explanations and identify the counterfactual are often missing.

Non-experimental designs generally are considered to be more appropriate for formative evaluations or monitoring the progress of (or fidelity to) project implementation than for summative evaluations. However, sometimes non-experimental studies are the only viable option for evaluators—especially if the evaluation was not planned prior to beginning implementation of the project or if the evaluator has little control over events during a study. Exhibit A3, below, discusses the use of different designs in the Anywhere State SLP project evaluation.

---

[14] Shadish et al., 2002.
[15] Dimitrov, 2010; Shadish et al., 2002.
[16] Shadish et al., 2002, p. 14.

Mixed-Method Designs: At the most basic level, mixed-method designs are those that combine quantitative and qualitative data collection and analysis approaches. An entire field of literature has developed related to the nature of mixed methods versus mixed methodology (that is, in general, mixed methods studies combine different data collection and analysis methods while mixed methodology studies combine different theoretical approaches as well as data collection and analysis methods).[17]

Grantees also should be aware of the type of data used in the evaluation. Data are either qualitative or quantitative in nature, where qualitative data include narrative data and data collected during interviews and open-ended questions, for example. Qualitative data typically are used to describe phenomena, experiences, impact and so on, without standardized metrics such as a scale or scoring rubric. In comparison, quantitative data include numeric data, such as the score on an achievement test, and data collected with close-ended or multiple choice questions, such as the Likert-scale. Evaluation designs often incorporate both types of data—these designs are often referred to as mixed-methods designs.

Exhibit A3. Evaluation Designs in the Anywhere State SLP Project Evaluation

> The project's plan incorporates both quasi-experimental and non-experimental designs. A quasi-experimental approach is planned for evaluation questions B and E, as follows:
>
> - Question B: Do the expected number and percentage of graduates work in appropriate settings for 3 years? The project plans to determine the retention rates for project graduates in comparison to graduates of other programs. To do this, the project has made an arrangement with Districts X and Y, where 60 percent of its graduates are employed, to use administrative records of employment of speech and language therapists. To protect personnel confidentiality, the data will be provided to the project in masked form in two datasets. The first dataset will be the employment status of project graduates. The second dataset will be the corresponding statistics for all other recently employed speech and language therapists. The project plans to use the second dataset as the comparison data.
>
> - Question E: To what extent do graduates demonstrate success with child/students with disabilities? The project also made an arrangement with Districts X and Y to use administrative records of progress on a relevant speech/language assessment. To protect child confidentiality, the data will be provided to the project in aggregated or masked form in two datasets. The first dataset will be the goal attainment scores for the students served by project graduates. The second dataset will be the corresponding statistics for all other special education students receiving speech and language services. The project will use the second dataset as the comparison data. When possible without risking the exposure of individual students' identities, the data will include individual student descriptors, such as sex, SES, and disability category.
>
> Evaluation questions A, C, and D will be treated non-experimentally. That is, the project only plans to capture project, and not comparison, data to respond to the evaluation questions.

Exhibit A4 presents the strategies, activities, outputs, outcomes, and evaluation questions for the project. In the following exhibit, consider the outcomes and evaluation questions. Are the outcomes SMART? Do the questions include progress monitoring, formative, and summative opportunities? Are the outcomes and questions well-aligned with the need for the projects and its goals?

---

[17] For more information on the design and conduct of a mixed method study, see Creswell, 2003; Brewer & Hunter, 2006; and Tashakkori, & Teddlie, 1998.

Exhibit A4. Evaluation questions aligned with Anywhere State Speech Therapist SLP Support project goals, strategies/activities, outputs, and outcomes

Goals	Strategies/Activities	Outputs	Outcomes	Evaluation Questions
1. Graduates demonstrate high-quality skills and performance in appropriate work settings.	■ Develop a tracking system to maintain contact with graduates. ▫ Develop a reliable and efficient system that will allow graduates to be followed for a minimum of 3 years post-completion. ▫ Obtain commitments from candidates to participate in follow-up activities post-graduation. ■ Develop a plan for systematically obtaining data on the performance of graduates and their students. ▫ Assemble a set of measures for determining skills and knowledge of candidates at time of program completion ▫ Work in conjunction with districts and agencies where graduates are employed (or likely to be employed) to develop valid and practical measures for determining graduate performance. ▫ Work in conjunction with districts and agencies to determine the most valid and efficient means for collecting data on the performance of children/students that graduates serve. ■ Implement the tracking system and integrate with other data collection plans	- Set of measures that capture candidate exit skills and knowledge assembled. - Work plan for tracking and follow-up of graduates. - Work plan for the collection of valid data on the performance of graduates and of their students.	Direct ● Graduates exit the program having demonstrated the skills and knowledge to perform at a high level. ● Graduates are tracked for 3 years. ● Data on the performance of graduates are collected for 3 years. Intermediate ● Graduates are working in appropriate settings for a minimum of 3 years. ● Graduates are providing high quality services within three years of exiting the program.	A. Are we collecting valid data on the performance of our graduates? B. Do the expected number and percentage of graduates work in appropriate settings for 3 years? C. To what extent do graduates exit the program with the skills and knowledge necessary to perform at a high level? D. To what extent do graduates demonstrate in the workplace the high quality skills and knowledge needed to improve outcomes for children with disabilities?

Goals	Strategies/Activities	Outputs	Outcomes	Evaluation Questions
2. Graduates demonstrate success with children/ students.	▪ Select/develop valid measures of student outcomes ▫ Examine extant data availability. ▫ Explore testing options. ▫ Develop graduate reporting protocol. ▪ Establish data collection system ▫ Prepare plan ▫ Develop timeline ▪ Develop analysis plan	- Instruments selected or developed and tested. - Data collection plan and timeline. - Analysis plan.	Direct ● A measurement system is implemented. ● Valid data are collected. Intermediate ● Evidence of success with children/students served by graduates. Long-term ● Students of graduates demonstrate improved outcomes.	E. To what extent do graduates demonstrate success with children/students with disabilities?

Westat®

Creating a data analysis plan

A data analysis plan identifies the specific steps the grantee or evaluator will use to analyze data, in response to the evaluation questions. (Ideally, a third-party evaluator will be responsible for analyzing summative data so as to maximize the objectivity and credibility of the evaluation findings.) The data analysis plan requires identification of the unit of analysis for each evaluation question (e.g., number of individuals, number of classes, number of trainings), as well as the possible comparison points such as treatment versus control or comparison group, special needs versus non-special needs status, or comparisons based on gender. Including a data analysis plan in the evaluation plan helps ensure that (a) the instrumentation chosen or developed for the evaluation will gather the needed   data in the correct format or scale, and (b) sufficient numbers and types of respondents or data sources   will be included in data collection.

Exhibit A5 presents a sample data analysis plan for the SLP Support project. As can be seen in the exhibit, the data analysis plan includes information **relevant to each evaluation question**, including:
- Study design (experimental, quasi-experimental, or non-experimental)
- Types of data collected (qualitative and quantitative)
- Presence of treatment and control (or comparison) groups
- Type of data analysis (using the most rigorous approach possible for answering each question)
- Variables to be used for quantitative analyses
- Instruments and data collection techniques
- Information on whether data is available for the entire evaluation population (i.e., a census) or for a sample of the evaluation's population—keeping in mind that the evaluation's population may include both treatment and control/comparison groups
- Minimum number of responses and/or response rate

The data analysis plan allows for an evaluation to use different designs for each question and allows for both quantitative and qualitative analyses to be applied. As shown in Exhibit A5, an evaluation may contain some questions that are treated experimentally or quasi-experimentally and others that are treated non-experimentally. Similarly, the evaluation may contain some questions addressed with qualitative data and analyses and others addressed with quantitative data and analyses. In the example presented in Exhibit A5, the project has decided to use descriptive quantitative analysis for questions A, C, and D. These analyses may include disaggregation or cross-tabulation of data to investigate patterns among graduates based on race, ethnicity, or gender. The project also decided to use statistical analyses (also quantitative) for questions B and E—comparison populations for these questions along with high response rate (i.e., a sufficiently high number of data points for both project and comparison groups) allow for two-group comparisons such as independent samples t-tests.

Exhibit A5. Sample Data Analysis Plan Template for the Anywhere State Speech Therapist Support project

Evaluation Question	Design		Data Analysis	Necessary Variables for Quantitative Analyses		Variable Sources (instruments or data collection techniques)	Population from which Data will be Collected	Minimum number of responses and/or response rate
	Design Type	If experimental or quasi-experimental, who constitutes the…		Statistical Tests	Descriptive statistics			
A. Are we collecting valid data on the performance of our graduates?	☐ Experimental ☐ Quasi-experimental ☒ Non-experimental	Treatment group: **N/A** Control or comparison group: **N/A**	☐ Statistical tests ☒ Descriptive statistics ☐ Qualitative analysis	Dependent: **N/A** Independent: **N/A** Covariates: **N/A**	☒ Frequency: **Percent of graduates with complete performance data** ☐ Mean	**Performance data as provided by partner districts (teacher evaluations and student achievement measures)**	☒ Census ☐ Sample (indicate sampling framework)	**A minimum 70% response rate from project graduates**
B. Do the expected number and percentage of graduates work in appropriate settings for 3 years?	☐ Experimental ☒ Quasi-experimental ☐ Non-experimental	Treatment group: **Project graduates** Comparison group: **Non-project graduates working in partner districts**	☒ Statistical tests ☒ Descriptive statistics ☐ Qualitative analysis	Dependent: **Length of time working in appropriate settings** Independent: Covariates: **Graduate project or comparison group affiliation, race, ethnicity, or gender**	☒ Frequency: **Percent of project and comparison group graduates who work in appropriate settings for at least 3 years.** ☒ Mean: **Average number of years project and comparison group graduates work in appropriate settings.**	**Employment data provided by partner districts**	☒ Census ☐ Sample (indicate sampling framework)	**A minimum 70% response rate from project graduates and comparison group graduates**

| Evaluation Question | Design | | Data Analysis | Necessary Variables for Quantitative Analyses | | Variable Sources (instruments or data collection techniques) | Population from which Data will be Collected | Minimum number of responses and/or response rate |
	Design Type	If experimental or quasi-experimental, who constitutes the…		Statistical Tests	Descriptive statistics			
C. To what extent do graduates exit the program with the skills and knowledge necessary to perform at a high level?	☐ Experimental ☐ Quasi-experimental ☒ Non-experimental	Treatment group: **N/A** Control or comparison group: **N/A**	☐ Statistical tests ☒ Descriptive statistics ☐ Qualitative analysis	Dependent: **N/A** Independent: **N/A** Covariates: **N/A**	☒ Frequency **Percent of graduates who exit the project with skills and knowledge necessary to perform at a high level** ☐ Mean	**Proficiency tests Short and extended form observation protocol**	☒ Census ☒ Sample: **a randomly selected sample of graduates will receive the extended observation**	**Data necessary for at least 70% of project graduates**
D. To what extent do graduates demonstrate in the workplace the high quality skills and knowledge needed to improve outcomes for children with disabilities?	☐ Experimental ☐ Quasi-experimental ☒ Non-experimental	Treatment group: **N/A** Control or comparison group: **N/A**	☐ Statistical tests ☒ Descriptive statistics ☐ Qualitative analysis	Dependent: **N/A** Independent: **N/A** Covariates: **N/A**	☒ Frequency: **Percent of graduates who demonstrate high quality skills and knowledge** ☐ Mean	**Performance data as provided by partner districts (teacher evaluations and student achievement measures)**	☒ Census ☐ Sample (indicate sampling framework)	**Data necessary for at least 70% of project graduates**
E. To what extent do graduates demonstrate success with child/students with disabilities?	☐ Experimental ☒ Quasi-experimental ☐ Non-experimental	Treatment group: **Project graduates** Comparison group: **Graduates from other projects**	☒ Statistical tests ☒ Descriptive statistics ☐ Qualitative analysis	Dependent: **Student achievement measures** Independent: Covariates: **Project or comparison group affiliation, race, ethnicity, or gender**	☐ Frequency: ☒ Mean: **Mean academic achievement of students served by project and comparison group graduates**	**Achievement data provided by partner districts**	☒ Census ☒ Sample: **a randomly selected sample of students will receive the Language Use Inventory**	**Data necessary for at least 70% of students served by project graduates**

> **TIP: If the evaluation design features sampling, obtain assistance from qualified staff.** Grantees serving large numbers of participants may elect to use a sample for their evaluation. For example, for a project that serves 200 or more individuals, collecting data on all of these individuals may be time-consuming, costly, and likely to result in a low response rate. Selecting a sample from this population enables the study team to save time and money, and, if done correctly, can help to improve the quality and accuracy of the data collected. This becomes even more relevant if the evaluation plan includes use of a comparison group, or when the evaluation seeks to collect data on the performance of individuals where the numbers may reach 1,000 or more. Some considerations when making decisions about how to select a sample include:
> - Available resources—How much time and money can be spent? How many people are available to work on the study (e.g., to recruit, collect data, analyze data)? Is a census cost-prohibitive?
> - Desired precision of estimates – What is the minimum sample size needed in order to reach conclusions with a pre-specified level of confidence? Will results be used to make comparisons between groups?
>
> Sampling is one aspect of an evaluation where a team member with specific training, expertise, or experience is needed (a third-party evaluator may be of particular assistance, if qualified for this task), as there are multiple technical details that must be accommodated when the evaluation uses a sample, including how to design a sampling framework and how to analyze the data that is collected using sample weights.

Constructing a data collection plan

The data collection plan identifies the specific tasks to be completed to ensure the required evaluation data are collected during the planned and appropriate timeframe(s). The following questions may help guide the creation of this plan:
- What instruments or data collection techniques will supply the variables that are needed—standardized assessment, survey, interview, etc.?
- Can some data be accessed through existing data sources? If so, what is the process for obtaining these data?
- What, if any, types of instrumentation or forms need to be identified or developed? Who will create and pilot test the new instruments?
- How will data collectors and data entry staff be trained? What materials, if any, need to be developed? Have all data collectors and data entry staff received training in the protection of human subjects[18]?
- When will data be collected? How frequently will data be collected?
- How will data be entered into a database and verified for accuracy? Where will data be stored?
- What security protocols will be developed to maintain data confidentiality? Who will have access to the data after it has been entered?

Exhibits A6 and A7 present sample data from the Anywhere State SLP Support project. Exhibit A6 presents a sample data collection schedule that identifies, for each data collection activity, the timing of the first and any additional data collections. The timing and use of multiple data collections is informed

---

[18] Cf: Federal Policy for the Protection of Human Research Subjects http://www.hhs.gov/ohrp/humansubjects/guidance/45cfr46.html or http://www2.ed.gov/about/offices/list/ocfo/humansub.html); Health Insurance Portability and Accountability Act of 1996
(HIPAA; http://www.hhs.gov/ocr/privacy/); Family Educational Rights and Privacy Act (FERPA; http://www.ed.gov/policy/gen/guid/fpco/ferpa/index.html).

by the nature and means of answering the evaluation question. For example, a question that focuses on whether or not students have achieved gains in knowledge or skills typically needs at least two data collections—a baseline or "pre" assessment and at least one "post" assessment. Exhibit A7 presents a data collection summary table, which links each evaluation question to data sources, possible instrument(s) or data collection protocol, the status of the instruments/ protocols, required training(s), timing of first and any subsequent data collections, and persons responsible for data collection, data entry and coding, and data quality and accuracy reviews.

As shown in these exhibits, at the start of the project, many data collection tasks have been identified, along with the timing of the task. However, in some cases, appropriate personnel have not yet been assigned to specific tasks.

Exhibit A6. Sample evaluation data collection schedule for the Anywhere State Speech Therapist Support Project

Data Collection Activity	First Data Collection	Additional Data Collections
SLP performance evaluation data submitted by partner districts	June 20XX	Annually for 3 years
SLP employment data submitted by partner districts	June 20XX	Annually for 3 years
Proficiency tests for project graduates collected from project staff	Final semester of training	N/A
Short and extended form observations conducted by third-party evaluator	Final semester of training	18 months after placement
Student achievement data provided by partner districts	June 20XX	Annually for 3 years

Exhibit A7. Sample Data Collection Summary Table for the Anywhere State Speech Therapist Support Project

Evaluation Question	Data Source(s)	Instrument/ data collection protocol	What is the status of the instrument/ protocol? (E=exists; UD=under development; TBD= to be developed)	Required Training	Timing of First Data Collection	Timing of Additional Data Collections, as necessary	Persons responsible for data collection	Persons responsible for data entry and coding	Persons responsible for data quality and accuracy checks
A. Are we collecting valid data on the performance of our graduates?	Partner districts	SLP evaluation data export	UD—the protocol for evaluation data export is under development	None	June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined
		Student achievement data export	TBD—the protocol for student achievement data export is awaiting development	None	June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined
B. Do the expected number and percentage of graduates work in appropriate settings for 3 years?	Partner districts	Employment data	E	None	June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined
C. To what extent do graduates exit the program with the skills and knowledge necessary to perform at a high level?	Project graduates	Proficiency tests	E	None	Final semester of training	N/A	Project staff	To be determined	To be determined
	Project graduates	Short and extended form observation protocol	UD	Observer training required to administer the observation protocol	Final semester of training	18 months after placement	To be determined	To be determined	To be determined

Westat®

Evaluation Question	Data Source(s)	Instrument/ data collection protocol	What is the status of the instrument/ protocol? (E=exists; UD=under development; TBD= to be developed)	Required Training	Timing of First Data Collection	Timing of Additional Data Collections, as necessary	Persons responsible for data collection	Persons responsible for data entry and coding	Persons responsible for data quality and accuracy checks
D. To what extent do graduates demonstrate in the workplace the high quality skills and knowledge needed to improve outcomes for children with disabilities?	Partner districts	Teacher evaluation data export	TBD—the protocol for teacher evaluation data export is awaiting development	None	June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined
	Partner districts	Student achievement data export	TBD—the protocol for student achievement data export is awaiting development	None	June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined
E. To what extent do graduates demonstrate success with child/students with disabilities?	Partner districts	Student achievement data export	TBD—the protocol for student achievement data export is awaiting development		June 20XX	Annually for 3 years	Project staff will receive data from district	N/A	To be determined

Developing a comprehensive timeline

The data collection and data analysis plans will inform development of a comprehensive evaluation timeline. The evaluation timeline combines data collection, analysis, and reporting tasks and can be included as part of the third-party evaluator's scope of work (provided that the roles and responsibilities for the different activities are clearly outlined). Specific deliverable dates can be included as well, or listed as a separate schedule of deliverables. Different types of project management tools can be used to create a timeline; one example is the Gantt Chart, as shown in Exhibit A8.

Exhibit A8. Anywhere State SLP Support Project's Evaluation Timeline: Year 1

Evaluation Task	Year 1: Project Implementation Months											
	1	2	3	4	5	6	7	8	9	10	11	12
Develop logic model	█											
Develop evaluation plan	█											
Develop analysis plan	█											
Prepare data collection instruments		█	█									
Complete IRB process		█										
Secure district participation		█	█									
Prepare training materials			█									
Conduct data collector training				█								
Conduct data collection				█	█	█	█					
Enter and clean data					█	█	█	█				
Conduct coder training									█			
Analysis and reporting										█	█	█	█

Westat®

Appendix B. Evaluation Needs Assessment Template

Question	Check the best option…	Possible Third-Party Evaluator Tasks

(1) Does the project already have an evaluation plan (a description of the evaluation questions, data collection tools and methods, analysis approach, and reporting requirements)?
 	
Note: Very often projects have some or all of an evaluation plan in place but the plan requires review or revision after a project is funded.
	☐ Yes, there is a complete evaluation plan in place, which responds in full to the evaluation requirements—**proceed to question 2. If you want to double check your answer, complete the checklist at right to identify possible third-party evaluator tasks**
	
	☐ There is a plan, but I'm not sure if it is complete or if it responds to requirements in full— **complete the checklist at right to identify possible third-party evaluator tasks**
	
	☐ No— **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Create or review the comprehensive evaluation plan; **OR**
		☐ Review, develop, or refine formative evaluation questions
		☐ Review, develop, or refine summative evaluation questions
		☐ Identify or review data collection sources
		☐ Identify or review data collection instruments
		☐ Create/pilot test data collection instrument(s)
		☐ Design data collection procedures
		   ☐ Implementation progress monitoring
		   ☐ Service statistics (e.g., numbers served; numbers of services provided)
		   ☐ Fidelity of implementation
		   ☐ Outcomes/impact data
		☐ Design data entry/ management procedures
		☐ Create data analysis plan
		☐ Design or review evaluation budget
		☐ Design or review report template(s)

(2) Are there internal staff with skills necessary to conduct the evaluation?
 	
Note: Very often projects will ensure statisticians and qualitative specialists (team members who specialize in qualitative research) are available to work on or support the evaluation.
	☐ Yes, internal staff are qualified for the types of evaluation required--**check off the applicable and needed skills below and proceed to question 3**
	
	  ☐ Formative evaluation—the evaluation will collect data on implementation progress and provide periodic feedback to project implementers to support project improvement
	
	  ☐ Measuring Fidelity of Implementation—the evaluation will collect data on implementation of the core components of the project, measure fidelity to the proposed theory of change, create and assign fidelity scores, and determine the level of component-level and overall fidelity of implementation
	
	  ☐ Experimental design—the evaluation will collect data on individuals randomly assigned into treatment and control groups; the evaluation will rigorously monitor treatment and control group conditions over the duration of the project
	
	  ☐ Quasi-experimental design—the evaluation will collect data on individuals placed into treatment and comparison groups by the evaluator; the evaluation will rigorously monitor treatment and comparison group conditions over the duration of the project
	
	  ☐ Non-experimental—the evaluation will collect data on the treatment group; a comparison group may be created post hoc (the evaluation will not track comparison group conditions over the duration of the project)
	
	  ☐ Design and implementation of a sampling plan—the evaluation will design a sample that is sufficient for the evaluation's approach, methodology, and analysis framework. The evaluation will identify how to treat sampled data (e.g., establish sample weights and any limitations on interpretation of data)
	
	☐ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Conduct formative evaluation activities
		☐ Conduct study of fidelity of implementation
		☐ Implement experimental or quasi-experimental design study (evaluator should have advanced background and expertise or training in sampling, research methodology)
		☐ Implement non-experimental study (evaluator should have basic background and expertise or training in research methodology)
		☐ Design and implement a sampling plan

Question	Check the best option…	Possible Third-Party Evaluator Tasks

(3) Can internal staff be sufficiently allocated to perform all evaluation tasks and responsibilities?	☐ Yes—**proceed to question 4**
	
	☐ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Create/pilot test data collection instruments
		☐ Collect data on
		    ☐ Implementation progress
		    ☐ Service Statistics (e.g., numbers served; numbers of services provided)
		    ☐ Fidelity of implementation
		    ☐ Outcomes/impact
		☐ Perform data entry/management
		☐ Conduct data analysis
		☐ Provide performance feedback to project team
		☐ Write reports
		☐ Other: _____

(4) Can internal staff perform all evaluation tasks and responsibilities objectively and without jeopardizing the credibility of evaluation findings?	☐ Yes—**proceed to item 5**
	
	☐ Unsure or No — **complete the checklist at right to identify possible third-party evaluator tasks**
		☐ Collect data on
		    ☐ Implementation progress
		    ☐ Service Statistics (e.g., numbers served; numbers of services provided)
		    ☐ Fidelity of implementation
		    ☐ Outcomes/impact
		☐ Perform data entry/management
		☐ Conduct data analysis
		☐ Provide performance feedback to project team
		☐ Write reports
		☐ Other: _____

(5) NEEDS ASSESSMENT COMPLETED
- If the answer to all questions is "yes", the project may not need a third-party evaluator.
- If the answer to one or more questions is "unsure or no," the project may benefit from hiring a third-party evaluator to perform specific tasks, as identified in this assessment.

Appendix C. Budgeting Guidance

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
Timing and Deadlines			

How much of the project evaluation will be assigned to a third-party evaluator?	• Deciding to use a third-party evaluator may add costs to the budget.
	
	Note: The third-party evaluator will likely provide a time and budget estimate for his or her work.
		• Working with a third-party evaluator requires additional time in the form of communications, oversight, and contract management.	• Projects frequently budget for the following evaluation staff:
			   o Project Director and/or Principal Investigator
			   o Administrative Assistant and/ or Project Manager
			   o Data collectors
			   o Data entry staff/ data technicians
			   o Third-party evaluator

If the project already has started, is the evaluation behind schedule on specific tasks?	• If the project is behind schedule when budgeting or budget revisions are calculated, grantees may consider adding staff to ensure evaluation tasks are completed in a timely fashion.	• If the project is behind schedule, grantees may need to allocate additional (often unbudgeted) hours so that evaluation tasks can be completed on schedule. This is especially important when data have to be collected within a specific time frame (e.g., if observational data must be completed before state testing starts in schools).	• Projects can easily fall behind schedule if the evaluation planners underestimate the amount of time necessary to complete specific tasks. For example, if the evaluation budget allocates 1 hour for collecting interview data but spends more than 1 hour to complete data collection, the evaluation may fall behind schedule and over-budget on that data collection task.
Evaluation Planning and Design			
Is the evaluation design complete when the project is funded?	• If the design is not complete, allocate time for staff (either internal or third-party) to review and complete the design.	• The amount of time needed will be related to the complexity of the evaluation approach. For example, a complex, long-term, experimental design may require more time to complete than a relatively simple, short-term, non-experimental design.	• Grantees may submit evaluation ideas or a draft evaluation plan with their applications. However, grantees may want to allocate staff and time to a review of the evaluation design after funding is approved. This may be necessary if OSEP has provided comments specific to the evaluation.
How many evaluation questions are there? *Note: The number of questions may be the first indication of the complexity and intensity of an evaluation.*	• The evaluation should have sufficient staff to fully implement all the evaluation activities in a timely fashion—ensuring all data are collected, analyzed, and reported within any required timeframes.	• The number of evaluation questions can impact the total time allocation since each evaluation question may require independent training of staff, data collection, data entry, etc.	• Create a complete evaluation approach for each question: Is the question experimental, quasi-experimental, or non-experimental in design? What is the means of data collection? Who will collect data, enter and manage data, and complete analyses and reporting? Consider working with a third-party evaluator to review and complete the evaluation design. • Evaluation reporting often takes longer than expected. Plan to create draft evaluation reports and allow time for review and revisions. • Plan evaluation reports to align with OSEP-required reporting periods so that the evaluation feeds into OSEP reports.

Westat®

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
For each evaluation question, what is the evaluation approach or design?	• Some designs (e.g., those that incorporate sampling or require substantial qualitative data collection) will need staff with training and experience in specific research methodologies.	• Evaluations that specify "changes", such as achievement gains, or change over time, will require at least two data collections. • Evaluations with many variables or possible explanations for observed changes may require more extensive or time-consuming data collections.	• Create a logic model or set of hypotheses for how the project will achieve its desired outcomes. Identify the possible explanatory or confounding variables and make sure there is a reliable, high-quality data source for each variable.
Communications and Management			
How frequently are communications planned?	• Grantees should identify a project liaison responsible for on-going communication with the evaluator and clearly identify how much other project staff will be involved with communications.	• More frequent communications will require more staff time. • Evaluations experiencing challenges may require more frequent communications.	• At a minimum, projects should plan for monthly, 1-2 hour communications. • Expect total communication frequency and duration to vary over the course of the project, with more communication usually taking place during the beginning and ending phases of the evaluation and whenever data collection occurs. • If an evaluation is experiencing challenges, grantees may need to add communication opportunities to the budget.
How frequently will the project monitor the evaluation?	• Grantees should identify a project liaison responsible for monitoring evaluation tasks and clearly identify how much other project staff will be involved with monitoring.	• Grantees may need to allocate time for developing a monitoring protocol or template, if one does not exist. • More frequent monitoring will require more staff time. • Evaluations experiencing challenges may require more frequent monitoring.	• The timing and duration of monitoring may vary across different phases of the evaluation. • Monitor data collection events during or directly after they occur to ensure the evaluation is achieving a sufficient quantity of valid and reliable data. • If an evaluation is experiencing challenges, grantees may need to add more monitoring opportunities to the budget.
Data Collection			
What types of data collection are expected (e.g., survey, interview, focus group, observational, and achievement data)?	• Evaluations that incorporate more than one type of data collection (e.g., collection of standardized assessment data, survey data, or interview data) may require staff with experience and training in specific techniques.	Different types of data collections require different amounts of time to complete. For example, assessments, surveys, and interviews will vary in length. Obtaining standardized data from a partner such as the local education agency may require a significant investment of time over one or more months.	• Consider utilizing at least two methods of data collection to answer evaluation questions—for example, interviews and standardized assessment data. This provides a richer range and scope of data.
How much data will be collected?	• If large amounts of data are to be collected, it may be necessary to budget for additional staff to conduct data collection activities in various locations simultaneously or to collect data multiple times over an extended period of time. • The amount of data to be collected can be influenced by whether or not the project uses sampling. Projects that use sampling may require staff with training and experience in sampling techniques.	• Large-scale data collections require larger time allocations.	• Consider including design and implementation of a sampling framework in the third-party evaluator scope of work and qualifications. • Determine the sample size as early as possible to adequately budget data collection time.

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
How frequently will data be collected?	• Evaluations with frequent data collections may require a larger data collection team to complete all of the data collection activities on time.	• Each data collection should require the same time "per unit" (e.g., per survey, per data export), although initial data collections sometimes require more time.	• Track and calculate data collection metrics. Calculate total and per unit data collection time and compare to budgeted estimates. Use findings to revise time allocations, as necessary and possible, for future collections. This may require budget changes.
Are established instruments in place or do instruments need to be developed?	• Staff should be qualified to identify whether existing instruments can be appropriately used to respond to the evaluation questions. • If the project does not plan to use an existing instrument, it may be necessary to identify staff qualified in creating data collection instruments.	• Choosing or creating data collection instruments can be time consuming since evaluations often require multiple instruments to answer different evaluation questions. • Instrument creation generally requires substantially more time than choosing an existing, valid, and reliable instrument.	• Allocate time for reviewing existing instruments for validity and reliability for the project evaluation. • Evaluators commonly underestimate the amount of time needed to develop and refine a new data collection instrument; allocate plenty of time for this (if needed). • If the project plans to use a published instrument(s), include the costs of the instrument's data collection forms, scoring guidelines, technical manual, and other necessary tools in the evaluation budget.
Will there be a pilot test? *Note: Pilot tests are "dry runs" of the instrument in its desired context.*	• If data collection instruments will be created for the evaluation, it may be necessary to identify qualified staff to conduct a pilot test for validity and reliability.	• If data collection instruments will be created for the evaluation, it may be necessary to allocate staff time and compensation for pilot testing, (i.e., recruiting pilot testers, conducting the test, analyzing the results, and making changes to the instruments).	• Allocate as much if not more time for the pilot test as for non-pilot data collections.
Will data collectors be trained to high levels of inter-rater reliability?	• Evaluations should use data collection staff that produce highly reliable results. It may be necessary to identify staff who are experienced in training data collection staff to high reliability. • Data collection staff should be trained to reliably use the instruments. Thus, if existing staff are not trained, they should receive training to ensure a high degree of reliability of the data collected.	• It may be necessary to allocate time for staff training and reliability testing on each of the evaluation's instruments. • Evaluations that incorporate multiple instruments may need to provide multiple trainings for data collection staff. • It may be necessary to allow time to create a training protocol, if one does not already exist (e.g., if the project is creating a data collection instrument).	• Publishers of existing instruments often require and provide training to ensure high reliability of the data collection. Check with the instrument's publisher to determine the training requirements. • Consider allocating time to conducting reliability checks on data.
Data Entry, Management, and Quality			
Who will enter data?	• Evaluations typically employ data technicians for data entry. • It is helpful to identify whether staff require training or specific qualifications to enter data.	• It may be necessary to allow time to create a data entry protocol, including a variable dictionary and data codes. • It may be necessary to allow time for training in data entry.	• Whenever possible, do not use "higher level" staff such as the project director or principal investigator for data entry. • Data technicians should be competent in word processing and spreadsheet programs.
How much data will be entered?	• If large amounts of data need to be entered in a relatively short amount of time, a larger data entry team may be required.	• Large-scale data collections (and longer data collection instruments) require larger time allocations. *Note—this assumes the initial data collection estimate is correct.*	• Track and calculate data entry metrics. Calculate total and per unit data entry time and compare to budgeted estimates. Use findings to revise time allocations, as necessary and possible, for future data collections. This may require changes to the budget.

Westat

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
Is a data management system in place or will one need to be constructed?	• Staff with training and experience in creating data management systems may be needed to create the spreadsheet or database that will house the data.	• Allow time for identifying or creating a data management system that will capture all data necessary to complete the evaluation's analyses. • Relatively complex evaluations may require relatively complex data systems (e.g., that allow multiple collectors to enter data simultaneously or remotely).	• Plan to enter data following each data collection event. • Plan for a means of linking data across data collection events (e.g., use of unique identifiers).
Are protocols for data entry, coding, management and quality checks in place or do they need to be developed?	• Qualified staff should be responsible for managing data entry and coding and reviewing data for quality.	• Allow time for data quality checks and management of the data entry and coding process.	• Plan to generate descriptive statistics to review data for data quality after the data have been entered. • Plan to review spreadsheets or databases for missing data and outliers.
Will data entry be checked for errors?	• Qualified staff should conduct data checks to ensure data entry contains minimal or no typographic or clerical errors.	• Allow time to check for data entry errors.	• Plan to select a random sample of data that has been entered to review for data entry errors—compare data that is entered to the original, raw data.
Data Analysis			
Is a data analysis plan or protocol in place or does one need to be developed?	• Qualified staff should be responsible for creating a data analysis framework that appropriately addresses all evaluation questions.	• Allow sufficient time to create and revise the analysis framework as the evaluation proceeds.	• Plan to revisit and revise the analysis framework (the expected process and sequence of data analysis) at least once during an evaluation.
Are there qualified staff to conduct data analyses?	• Evaluations with experimental and quasi-experimental designs, as well as those that incorporate sampling, require staff with statistical training and experience. • Evaluations with qualitative elements (e.g., case studies) require staff with specific training and experience in qualitative methods.	• Allow sufficient time for a complete data analysis for each evaluation question. Quantitative and qualitative analyses both can be very time consuming—the project may find it helpful to consult with a third-party evaluator, statistician, or qualitative specialist to determine the amount of time necessary for analyses.	• Complete an analysis framework for each question that identifies the steps and estimated time necessary to complete the data analysis. • Include the costs of consulting with a third-party evaluator, statistician, or qualitative specialist in the evaluation's budget.
Will analyses be reviewed within the team?	• Staff who will be responsible for reviewing the analyses should be able to assess whether they are appropriate for the type of data and the evaluation question, and determine if the findings are accurate.	• It may be necessary to allow time for different types of analyses to be confirmed by different staff members.	• Plan to have a second staff member review the more complex analyses—or analyses that incorporate advanced statistical techniques.

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
Reporting			
How many reports/products are expected (e.g., technical report, policy brief, PowerPoint presentation)?	• Specific products may require staff with specific skills in preparation or editing (e.g., in making documents Section 508 compliant).	• Allow sufficient time for writing, review, and revision of each product. It may be helpful to consult with a third-party evaluator, statistician, or qualitative specialist to determine the amount of time necessary for thorough reporting.	• Include extra time for reporting in the project timeline and budget—it usually takes longer than expected. • Include the costs of consulting with a third-party evaluator, statistician, or qualitative specialist in the evaluation's budget.
Are report templates in place or do they need to be developed?	• It may be helpful to make a qualified and knowledgeable staff person responsible for developing report templates that meet specific requirements.	• It may be necessary to allow time for developing and receiving approval from project managers on draft report templates.	• OSEP provides an Annual Performance Report template. • The project may want to produce additional reports, for varied audiences.
Who will write and who will review the report(s) and product(s)?	• It may be helpful to identify o the staff who are qualified and who will be responsible for <u>writing</u> the report(s). o the staff who are qualified and who will be responsible for <u>reviewing</u> the report(s).	• Allow sufficient time for report writing and report review. • Allow time for report revisions. • More complex evaluations, with a larger number of evaluation questions, likely will require a larger time allocation for report writing and review.	• Allocate at least 2 days (each) for report writing, review, and revision for each evaluation question. • Allocate additional time for the report's introduction and conclusion. • Allocate additional time for submission of various drafts of reports for feedback, as needed. • Time the evaluation's report writing to align with required OSEP reporting so that the evaluation findings can feed into any required reports.

Westat®

Appendix D. Time Frame Estimates for Common Data Collection Activities

Exhibit D1. Focus Group/Interview Timeline (approximately 12 weeks)

Task	Timeframe
Write the focus group/interview purpose statement	6-8 weeks prior to session
Identify the participants	6-8 weeks prior
Gather participant contact information	6-8 weeks prior
Assign group facilitator/interviewer	4-5 weeks prior
Develop the questions	4-5 weeks prior
Develop script/protocol	4-5 weeks prior
Arrange and reserve the session site for focus group	4 weeks prior
Write and send the invitation	3-4 weeks prior
Follow up the invitation with a phone call	2 weeks prior
Make room/meeting arrangements (seating, equipment for focus group)	1 week prior
Place a reminder call to participants	2 days prior
Conduct Focus Group/Interview	1 – 1.5 hours each
Transcribe Focus Group/Interview recording	1 week following session
Analyze, Interpret, and Report Results	3-4 weeks following session

Adapted from *The Fieldstone Alliance Nonprofit Field Guide Series Conducting Successful Focus Groups* by Judith Sharken Simon, 1999, Fieldstone Alliance

Exhibit D2. Web-based Survey Timeline (approximately 12 weeks)

Task	Timeframe
Identify research objectives	9 weeks prior to administering survey
Create sampling plan	8 weeks prior
Select the sample; collect contact information	6-7 weeks prior
Write, review, and revise the survey	4-5 weeks prior
Write, review, and revise the invitation letter/email	3.5 weeks prior
Conduct pilot test	3 weeks prior
Revise survey (if necessary)	2 weeks prior
Enter survey and database of potential respondents online	1 week prior
Begin survey	--
Monitor and attempt to increase response rate	2-3 weeks after survey begins
End survey	3-4 weeks after survey begins
Code open-ended responses	1-2 weeks after survey ends
Analyze the data	3 weeks after survey ends
Write, review, and revise report	4-6 weeks after survey ends

Adapted from *Conducting Online Surveys* by Valerie M. Sue & Lois A. Ritter, 2007, Sage Publications, Inc.

Appendix E. Common Practices for Protecting Personally Identifiable Information

1. Use of identification (ID) codes on data collection instruments in place of identifying information. The use of codes instead of names may prevent anyone who may see the data from determining the participants' identities.

2. Separation of files that contain participant identifiers (e.g., names and addresses, if these are necessary aspects of data collection) from evaluation instruments containing data. A unique, randomized, identification number can be used to link the files—what is key is that personally identifiable information is not found in the same file as other data fields.

3. Encryption of personally identifiable information.

4. Use of security codes and password protection to restrict access to computerized records.

5. Use of secure data transfer protocols (e.g., FTPS, SFTP).

6. Use of locked and limited access data filing for hard copies of data collection instruments, documents, and forms.

7. Proper disposal, destruction, or deletion of study data/documents.

Appendix F. Recommended Readings on Research/Evaluation Methodology

Brewer, J. & Hunter, A. (2006). *Foundations of multimethod research: Synthesizing styles.* Thousand Oaks, CA: Sage.

Campbell, D. T., & Stanley, J. C. (1963). Experimental and quasi-experimental designs for research. In N. L. Gage (Ed.), *Handbook of research on teaching* (pp. 1-76). Chicago: Rand-McNally.

Creswell, J. W. (2011). *Educational research: Planning, conducting, and evaluating quantitative and qualitative research (4th ed.).* Boston: Pearson.

Creswell, J. W. (2003). *Research design: Qualitative, quantitative and mixed methods approaches (2nd edition).* Thousand Oaks, CA: Sage.

Dillman, D., Smyth, J. & Christian, L. M. (2009). *Internet, mail, and mixed-mode surveys (3rd edition).* Hoboken, NJ: John Wiley & Sons.

Dimitrov, D. M. (2010). *Quantitative research in education: Intermediate and advanced methods (2nd edition).* New York: Whittier Publications.

Dong, N., & Maynard, R. A. (2013). PowerUp!: A tool for calculating minimum detectable effect sizes and minimum required sample sizes for experimental and quasi-experimental design studies. *Journal of Research on Educational Effectiveness, 6*(1), 24-67. DOI: 10.1080/19345747.2012.673143

Frechtling, J., with M. Melvin, D. Rog, & E. Johnson. (2010). *The 2010 user-friendly handbook for project evaluation.* Rockville, MD: Westat & Directorate for Education and Human Resources, Division of Research and Learning in Formal and Informal Settings, National Science Foundation.

Groves, R., Flower, F., Couper, M., Lepkowski, J., Singer, E., & Tourangeau, R. (2004). *Survey methodology.* Hoboken, NJ: John Wiley & Sons.

Harkness, J., Braun, M., Edwards, B., & Johnson, T. P. (2010). *Survey methods in multinational, multiregional, and multicultural contexts.* Hoboken, NJ: John Wiley & Sons.

Hedges, L. V., & Rhoads, C. (2010). *Statistical power analysis in education research* (NCSER 2010-3006). Washington, DC: U.S. Department of Education Institute for Education Sciences. Retrieved from: http://ies.ed.gov/ncser/pubs/20103006/pdf/20103006.pdf

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (2003). *Applied statistics for the behavioral sciences* (Fifth edition). Boston: Houghton Mifflin.

Hox, J. J. (2010). *Multilevel analysis: Techniques and applications* (2nd edition), (Quantitative Methodology Series). New York: Routledge.

Kennedy, C. H. (2005). *Single-case designs for educational research.* Boston: Pearson Education.

Lammert, J. D., Heinemeier, S., & Fiore, T.A. (2013). *Evaluating special education preservice programs using graduate and student outcomes: Resource Toolkit*. Rockville, MD: Westat.

Maxwell, J. A. (2005). *Qualitative research design: An interactive approach (2nd ed.).* Applied Social Research Methods Series, vol. 41. Thousand Oaks, CA: Sage.

Maxwell, J. A., & Miller, B. A. (2010). Categorizing and connecting strategies in qualitative data analysis. In S. N. Hesse-Biber & P. Leavy. (Eds.), *Handbook of emergent methods*. New York: The Guilford Press.

Patton, M. Q. (2002). *Qualitative research and evaluation methods (3rd ed.)* Thousand Oaks, CA: Sage.

Raudenbush, S. W., et al. (2011). Optimal Design Software for Multi-level and Longitudinal Research (Version 3.01) [Software]. Retrieved from: http://hlmsoft.net/od/

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Belmont, CA: Wadsworth.

Tashakkori, A., & Teddlie, C. (1998). *Mixed methodology: Combining qualitative and quantitative approaches*. Thousand Oaks, CA: Sage.

Todman, J. B., & Dugard, P. (2001). *Single-case and small-n experimental designs: A practical guide to randomization tests.* Mahwah, NJ: Lawrence Erlbaum Associates

U.S. General Accounting Office, GAO. (2012). *Designing evaluations: 2012 revision* (Report No. GAO-12-208G). Washington, DC: Author. Retrieved from http://www.gao.gov/assets/590/588146.pdf

U.S. General Accounting Office, GAO. (1992). *Quantitative data analysis: An introduction* (Report to Program Evaluation and Methodology Division, Report No. GAO/PEMD-10.1.11). Washington, DC: Author. Retrieved from http://www.gao.gov/special.pubs/pe10111.pdf

Willis, G. (2005). *Cognitive interviewing: a tool for improving questionnaire design*. Thousand Oaks: Sage.

W. K. Kellogg Foundation. (2006). *Logic Model Development Guide*. Battle Creek, MI: Author. Retrieved from: http://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook

W. K. Kellogg Foundation. (2004). *Evaluation Handbook*. Battle Creek, MI: Author. Retrieved from: http://www.wkkf.org/resource-directory/resource/2010/w-k-kellogg-foundation-evaluation-handbook