Selected Evaluation Project Management Tools

Presented by Jill D. Lammert, Ph.D.
Westat

May 15, 2015

Evaluation Needs Assessment Template

Question	Check the best option…	Evaluation Team Tasks

(1) Does the State SSIP team already have an evaluation plan (a description of the evaluation questions, data collection tools and methods, analysis approach, and reporting requirements)?
 	
Note: Even if the State team has some or all of an evaluation plan in place, it maybe helpful to have an external (third-party) evaluator review the plan prior to submission.
	☐ Yes, there is a complete evaluation plan in place, which responds in full to the SSIP Phase II Evaluation Plan requirements—**proceed to question 2. If you want to double check your answer, complete the checklist at right to identify possible evaluation team 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 evaluation team tasks**
	☐ No— **complete the checklist at right to identify possible evaluation team 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) Do State SSIP team members have the skills necessary to conduct the evaluation?	☐ Yes, State SSIP team 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 evaluation team; 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 evaluation team task**
		☐ Conduct formative evaluation activities
		☐ Conduct study of fidelity of implementation
		☐ Implement experimental or quasi-experimental design study (staff should have advanced background and expertise or training in sampling, research methodology)
		☐ Implement non-experimental study (staff should have basic background and expertise or training in research methodology)
		☐ Design and implement a sampling plan

Question	Check the best option…	Evaluation Team Tasks

(3) Can State SSIP team 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 evaluation team 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 State SSIP team 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 evaluation team 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 additional evaluaton team members. • If the answer to one or more questions is "unsure or no," the project may benefit from identifying specific staff to serve on the evaluation team or hiring a third-party evaluator to perform specific evaluation team tasks, as identified in this assessment.		

This tool was adapted from the *Guidelines for Working with Third-Party Evaluators*, which were developed by the Center to Improve Project Performance (CIPP) operated by Westat for the U.S. Department of Education, Office of Special Education Programs (OSEP). Reference: Heinemeier, S., D'Agostino, A., Lammert, J.D., & Fiore, T.A. (2014). *Guidelines for Working with Third-Party Evaluators*. Rockville, MD: Westat.

Budgeting for Evaluation: Key Factors to onsider

Knowing how much to budget for an evaluation requires an understanding of the evaluation process and of the various factors that might influence costs. In simple terms, the amount of money that you will need depends on the scope and complexity of both the program to be evaluated and the evaluation itself.

SCOPE refers to factors such as the size and reach of the program or project to be evaluated, the evaluation focus, the number of evaluation questions, the evaluation time period, whether and how stakeholders will be involved, and the number and type of reports that will be required.

COMPLEXITY refers to factors such as the nature of the evaluation questions, the type of evaluation design needed to answer each evaluation question, the number of participants and sites to be included in the evaluation, data collection methods, frequency and duration of data collection, and data management and analysis requirements.

No matter the scope or complexity, all evaluations require certain costs, the most significant of which is staffing. Other likely costs include materials and supplies, computer equipment and software, and local or out-of-town travel.

If necessary, the cost of different types of evaluations can be roughly estimated in terms of a percentage of the program or project budget, as follows:

- Low cost = <10%
- Moderate cost = 10-20%
- High cost = >20%

Ideally, you should estimate evaluation costs in absolute dollar amounts, after carefully considering scope and complexity in the context of the specific evaluation needs. The *Evaluation Cost Considerations Worksheet* included in this brief can help you think about the different factors affecting the scope and complexity of an evaluation. In the worksheet, evaluation elements are assigned relative costs based on how they may affect staffing, supplies, travel, etc. The list of evaluation elements presented in the worksheet is not exhaustive and the different categories are not mutually exclusive; the goal is to help you think more deliberately about the factors affecting costs in an evaluation.

Instructions for completing the Cost Considerations Worksheet: As you think about your evaluation, check off each item you will need, calculate your score for each row by summing the items across all three columns (e.g., 1 point for items in column one, 2 points for column two, and 3 points for column three), and then calculate your total score. The scores ranges give you an idea of the relative cost of your evaluation, keeping in mind that the actual cost will depend on factors such as labor rates and travel costs. Note: In some rows (e.g., Interview Mode) you might pick no items or only one item across all three columns, while in other rows you might pick multiple items in multiple columns.

Bibliography

Corporation for National and Community Service. (2014). *A quick guide for social innovation fund applicants: budgeting for rigorous evaluations*. Washington, DC: Author. Retrieved from: http://www.nationalservice.gov/build-your-capacity/grants/funding-opportunities/2014/social-innovation-fund-grants-fy-2014/financial-management-system-requirements

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

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

Horn, J. (2001). A checklist for developing and evaluating evaluation budgets. Western Michigan University, Evaluations Checklist Project. Available at: http://scholar.google.com/scholar?hl=en&q=A+checklist+for+developing+and+evaluating+evaluation+budgets&btnG=&as_sdt=1%2C21&as_sdtp

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

The Pell Institute for the Study of Opportunity in Higher Education, the Institute for Higher Education Policy, & Pathways to College Network (2014). *Evaluaton toolkit: Create a budget.* Washington, DC: The Pell Institute. Retrieved from: http://toolkit.pellinstitute.org/evaluation-guide/plan-budget/develop-a-budget/

Twersky, F., & Arbreton, A. (2014). *Benchmarks for spending on evaluation*. Menlo Park, CA: The William and Flora Hewlett Foundation

U.S. Department of Health and Human Services. (2010). *The program manager's guide to evaluation (2nd Ed.)*. Washington, DC: U.S. Department of Health and Human Services, Office of Planning, Research and Evaluation, Administration for Children and Families. Retrieved from: http://www.acf.hhs.gov/programs/opre/resource/the-program-managers-guide-to-evaluation-second-edition

This brief was prepared 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) under contract number ED-OSE-13-C-0049. The project officer is Dr. Patricia Gonzalez. The content of this publication does not necessarily reflect the views or policies of the U.S. Department of Education nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. government. Suggested citation: Lammert, J. D., & Fiore, T. A. (2015). *Budgeting for evaluation: Key factors to consider*. Rockville, MD: Westat.

Evaluation Cost Considerations Worksheet

Evaluation Element	Low Cost (1 point/item)	Moderate Cost (2 points/item)	High Cost (3 points/item)	Score
Evaluation Design Elements				
Focus of Formative Study	— Participant satisfaction — Project implementation	— Outputs (e.g., satisfaction, quality, relevance) — Implementation fidelity (key components, activities, outputs, possibly some direct outcomes)	— Intervention fidelity that includes mediators, intermediate outcomes	_____
Focus of Summative Study	— Changes in participant satisfaction — Changes in existing data (e.g., student scores on state tests)	— Short-term outcomes	— Intermediate/long-term outcomes — Comparative outcomes (e.g., treatment vs. control groups) — Causal attribution	_____
Evaluation Study Design	— Non-experimental (descriptive study, basic qualitative methods)	— Non-experimental (case studies, advanced qualitative methods) — Simple quasi-experiment (QED, e.g., basic comparison study) — Single-case design (SCD; reversal design)	— Complex QED (e.g., with matching, multiple comparison groups) — Randomized controlled trial (RCT) — Multi-site or cluster RCT — SCD (multiple baseline, alternating treatment design)	_____
# of Participants/ Sites, Sampling	— Small target population — 1-2 sites — Simple sampling plan (e.g., purposive, simple random)	— Moderate-size target population — 3-5 sites — Somewhat complex sampling plan (e.g., stratified)	— Large target population — >5 sites — Highly complex sampling plan (e.g. stratified, clustered, weighted)	_____
Data Collection Elements				
Document review	— Limited document search — Basic document summaries	— Extensive document search — Detailed document summaries — Limited document synthesis — Limited qualitative analysis of documents	— Extensive document synthesis — Extensive qualitative analysis of documents	_____
Survey (existing or new)	— Existing (available, no changes needed)	— Existing (fee to use) — Existing (some changes needed) — New survey (with limited pilot testing & no validation study)	— New survey (with extensive pilot testing) — New survey (with validation)	_____
Survey mode of administration	— Simple online (e.g., basic Survey Monkey survey)	— Customized online survey (e.g. customized Survey Monkey with skip patterns) — Self-adminstered paper-and-pencil suvey (few respondents) — Telephone survey (few respondents, brief responses)	— Customized online survey with integrated data management (e.g., survey with highly complex skip patterns & linkages to data management system) — Self-adminstered paper-and-pencil suvey (many respondents) — Telephone survey (many respondents, lengthy responses)	_____
Survey non-response follow-up	— Limited email follow-up	— Extensive email follow-up — Limited telephone follow-up	— Extensive telephone follow-up — Mail follow-up with reminder cards	_____
Interview mode	— Online	— Telephone	— Face-to-face[a]	_____
Interview type	— Structured (i.e., asking specific, close-ended questions)	— Semi-structured (i.e., asking some close-ended & some open-ended questions)	— Unstructured (i.e., asking open-ended questions, with focus potentially varying by respondent)	_____
Interview data capture	— Interviewer takes notes during interview	— Interview recorded & transcribed	— Note-taker present at interview	_____
Observation location	— Local[a]	— Driving distance (overnight)[a]	— Long-distance (air travel required)[a]	_____

Observation protocol	— Checklist (i.e., specific activities or behaviors to observe; limited training required)	— Guided/structured protocol (i.e., general categories of activities or behaviors to observe; some content knowledge and training required)	— Unstructured protocol (i.e., open-ended with focus varying by site; deep content knowledge and/or extensive training required)	_____
Assessments	— Existing (conducted at no cost to evaluation)	— Administer small scale pre-post assessments (with no specialized credential or training required)[b]	— Administer large scale pre-post assessments[b] — Administer repeated assessments[b] — Administer assessments (with specialized credential or training required)	_____
Data Collection Frequency & Duration	— Limited frequency (1 time per year or less)	— Moderate frequency (2 or 3 times per year) — Multi-year, but not annual	— Frequent data collection (4 or more times per year) — Annual, or longitudinal data collection	_____
Data Management Elements				
Data management software & hardware	— New software required (low cost) — New hardware required (low cost)	— New software required (moderate cost) — New hardware required (moderate cost)	— New software required (high cost) — New software required (high cost)	_____
Data control & cleaning	— Limited need for data quality control (i.e., multiple choice items; data collected electronically, etc.) — Limited need for data cleaning (i.e., few duplicate records & outliers, little need for coding, etc.)	— Moderate need for data quality control (i.e., multiple choice with some write-in; some field scoring of assessments; some missing data, etc.) — Moderate need for data cleaning (i.e., moderate number of duplicate records & outliers, need to recode some data, etc.)	— Extensive need for data quality control (i.e., open-ended questions; field scoring of assessments; need to merge & reconcile diverse databases; extensive missing data, etc.) — Extensive need for data cleaning (i.e., extensive duplicates & outliers, extensive recoding needed, etc.)	_____
Data Entry	— Automated data entry (e.g., online survey)	— Data entry mostly automated, with some need for hand entry	— Data entry entirely by hand	_____
Database	— Existing database	— Create new database with limited functionality/data sharing	— Create new database with multi-user functionality/data sharing	_____
Data Analysis Elements				
Type of analysis	— Basic descriptive quantitative analysis (e.g., frequencies, t-tests, chi-square tests, ANOVA) — Very limited qualitative analysis	— Intermediate quantitative analysis (e.g., regression, ANCOVA) — Somewhat limited qualitative analysis	— Advanced quantitative analysis (e.g., HLM, SEM) — Extensive qualitative analysis	_____
Data analysis software	— New software required (low cost)	— New software required (moderate cost)	— New software required (high cost)	_____
Technical expertise for analysis	— Existing staff have all needed expertise	— External consultant needed to conduct some analysis or train existing staff	— External consultant needed to conduct most or all analysis or to provide extensive staff training & support	_____
Reporting Elements				
Reporting frequency	— Annual report only	— Interim & annual reports	— Monthly, interim, & annual reports	_____
Types of presentations/ audiences	— Oral evaluation updates — Presentations to project staff only	— Policy briefs (limited audiences) — Presentations to project staff & few stakeholders	— Policy briefs (multiple audiences) — Presentations to project staff & multiple stakeholders	_____
Low Cost (<36 points); Moderate Cost (36-79 points); High Cost (≥80 points)			**Total Score**	_____

Notes: a. Cost = frequency x travel cost; b. Cost = frequency x number of assessment instruments

IDC IDEA DATA CENTER

Budgeting Guidance for Evaluations

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
Timing and Deadlines			
How much of the SSIP 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.	• It is common to budget for the following evaluation staff: ○ Project Director and/or Principal Investigator ○ Administrative Assistant and/ or Project Manager ○ Data collectors ○ Data entry staff/ data technicians ○ Data analysts ○ Report writer ○ Third-party evaluator
If the SSIP implementation already has started, is the evaluation behind schedule on specific tasks?	• If the SSIP implementation is behind schedule when budgeting or budget revisions are calculated, State teams may consider adding staff to ensure evaluation tasks are completed in a timely fashion.	• If the SSIP implementation is behind schedule, State teams 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).	• Evaluations 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			
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 amount of time needed to develop the evaluation plan 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. • 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.
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 SSIP 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.

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
Communications and Management			
How frequently are communications planned?	• State teams should identify a project liaison responsible for on-going communication with the evaluation team and clearly identify how much other 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, State teams should plan for monthly, 1-2 hour communications. • Expect total communication frequency and duration to vary over the course of the implementaton period, 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, State teams may need to add communication opportunities to the budget.
How frequently will the State SSIP team monitor the evaluation?	• State teams should identify a liaison responsible for monitoring evaluation tasks and clearly identify how much other staff will be involved with monitoring.	• State teams 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, State teams 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 evaluation uses sampling. Evaluations 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 evaluation team scope of work and qualifications. • Determine the sample size as early as possible to adequately budget data collection time.
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.

Questions to Guide Budgeting	Staffing Considerations	Time Considerations	Tips
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 evaluation team 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 SSIP evaluation. • Evaluation teams 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 evaluation team 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 evaluation team 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.

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 State team 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 evaluation 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 administrators or managers on draft report templates.	• The State team 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 ○ the staff who are qualified and who will be responsible for <u>writing</u> the report(s). ○ 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.

This tool was adapted from the *Guidelines for Working with Third-Party Evaluators*, which were developed by the Center to Improve Project Performance (CIPP) operated by Westat for the U.S. Department of Education, Office of Special Education Programs (OSEP). Reference: Heinemeier, S., D'Agostino, A., Lammert, J.D., & Fiore, T.A. (2014). *Guidelines for Working with Third-Party Evaluators*. Rockville, MD: Westat.

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 was the evaluation team assembled?	☐ The evaluation team was assembled prior to submission of the SSIP Evaluation Plan. ☐ The evaluation team was assembled on schedule (according to the SSIP management plan).	☐ There is no evaluation team. ☐ The evaluation team was assembled after submission of the SSIP Evaluation Plan.
(2) Have stakeholders been identified and included in the evaluation planning?	☐ Stakeholders have been identified and will be included in the evaluation planning (when it begins). ☐ Stakeholders have been identified and included in the evaluation planning.	☐ Stakeholders have not been identified. ☐ Stakeholders are not expected to be included in the evaluation planning (when it begins). ☐ Stakeholders have not been included in the on-going evaluation planning.
(3) 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		
(4) 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.
(5) Was the evaluation team involved in the evaluation's design?	☐ The evaluation team was involved with the evaluation design from the beginning. ☐ The evaluation team reviewed a preexisting evaluation plan and is satisfied with it. ☐ The evaluation team helped revise and refine a preexisting evaluation plan.	☐ The evaluation team 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.
(6) 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.
(7) 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		
(8) How frequently does the State team communicate with the evaluation team and stakeholders?	☐ Communication occurs according to the communication schedule. ☐ Communication occurs at least once a month.	☐ There is no communication schedule. ☐ Communication with the evaluation team is sporadic. ☐ Communicaton with stakeholders is sporadic. ☐ Communication is infrequent--it is difficult to get in touch with the evaluation team (or stakeholders).
(9) Is the State team 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		
(10) Does the evaluation have access to the desired participants and sites for data collection?	☐ Desired participants and sites are accessible (i.e., the evaluation team can collect data on or from the participant or site).	☐ Only some of the desired participants or sites are accessible to the evaluation team. ☐ None of the desired participants and sites are accessible to the evaluation team.

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 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.
(12) 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.
(13) 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 evaluation team did not review or approve final instruments. ☐ Many forms and instruments are returned with missing data.
	Data Entry, Management, and Quality	
(14) 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.
(15) 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).
(16) 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	
(17) 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.
(18) 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.

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)
(19) 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.
	Reporting	
(20) Are report templates in place?	☐ The evaluation team works with the State team to develop report templates. ☐ The evaluation team uses existing report templates.	☐ The evaluation team has not received SSIP reporting requirements or provided report templates.
(21) Are reported findings reliable and credible?	☐ Reporting of findings is based on high-quality data and analyses. ☐ There is no evidence of bias in the reported findings. ☐ The evaluation team has authority to report findings without prior “approval” from the State team. [Note: this does not mean that the evaluation team has primary responsibility for generating all reports on the SSIP implementation and outcomes – only that the findings do not have to be “approved.”]	☐ The evaluation findings are not based on high-quality data or analysis. ☐ There is evidence of bias in the reported findings. ☐ The State team wants to “approve” the findings generated by the evaluation team before they are reported.

This tool was adapted from the *Guidelines for Working with Third-Party Evaluators*, which were developed by the Center to Improve Project Performance (CIPP) operated by Westat for the U.S. Department of Education, Office of Special Education Programs (OSEP). Reference: Heinemeier, S., D’Agostino, A., Lammert, J.D., & Fiore, T.A. (2014). *Guidelines for Working with Third-Party Evaluators*. Rockville, MD: Westat.

Evaluation Close-Out Tasks

Close-Out Task	The State team can…	The evaluation team 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 (if applicable).	☐ Review the contract and identify any outstanding contracted items (if applicable). ☐ Submit any outstanding products or deliverables to the State team. ☐ Submit a final request for payment or invoice, with a deadline for payment (if applicable).
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 evaluation team that final payment may be delayed until all contracted documents are received (if applicable).	☐ 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 State team 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 State team.
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 primary State team contact cannot be reached 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 evaluation team have permission to respond to requests or should the interested party be routed to someone else in the State team's agency?*	☐ Provide the State team with primary and secondary contact information (if external to the State team). ☐ Establish the length of time the evaluation team will be available for follow-up or requests for information (if external to the State team).
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 evaluation team and a list of all data and files that are to be destroyed. *Note—it may be necessary to maintain evaluation records (including data) for a pre-specified amount of time, depending on OSEP or State 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 State team, as required.

This tool was adapted from the *Guidelines for Working with Third-Party Evaluators*, which were developed by the Center to Improve Project Performance (CIPP) operated by Westat for the U.S. Department of Education, Office of Special Education Programs (OSEP). Reference: Heinemeier, S., D'Agostino, A., Lammert, J.D., & Fiore, T.A. (2014). *Guidelines for Working with Third-Party Evaluators*. Rockville, MD: Westat.