

SUPPLEMENTAL EVALUATION RESOURCES

RUBRIC FOR IDENTIFYING INCOMING LEVELS OF EVIDENCE

GUIDANCE

For grantmaking institutions applying to the Social Innovation Fund (SIF) with a pre-identified intervention, the evidence level of that intervention will be assessed during the application review. This document explains how that assessment will be conducted and includes the rubric that will be used.

The process of determining the level of evidence for an intervention can be complicated and often requires technical knowledge of research and evaluation design and methods. The rubric introduced in this document provides a framework for assessing the existing body of evidence based on past research and evaluation studies.

I. ATTAINING A LEVEL OF EVIDENCE

As described in the SIF NOFA, all SIF-funded interventions require at least a Preliminary level of evidence upon entering the SIF. This level of evidence will be demonstrated by studies conducted prior to applying for SIF funding.

- To attain the **Preliminary** level of evidence required for SIF funding, an intervention must, at a minimum, have a study that has “yielded promising results for either the program or a similar program.” Specifically, the intervention must have at least some outcome information such as pre- and post-tests without a comparison group, or post-test comparison between program and comparison groups.
- To attain a **Moderate** level of evidence, an intervention needs to have evidence “from studies whose designs can support causal conclusions (i.e., studies with high internal validity¹), but have limited generalizability (i.e., moderate external validity²), or studies with high external validity, but moderate internal validity.”³ Studies with high internal validity will likely use Quasi-experimental Designs (QED) (such as a matched comparison group or a comparative interrupted time series design) or Randomized Controlled Trials (RCT) also known as Experimental Designs. At least one study with high internal or external validity is typically needed to attain a **Moderate** level of evidence.

¹Internal validity for a study is the extent to which the observed difference in the average group outcomes (usually program participants versus control or comparison group members) can be causally attributed to the intervention or program.

²External validity for a study is the extent to which evaluation results are applicable to groups other than those in the research.

³Moderate internal validity could come from a study having a comparison group formed without statistical matching techniques, statistical matching techniques that resulted in lower than desirable pre-test group equivalence, or an interrupted time series design without a comparison group.

- To attain a **Strong** level of evidence, an intervention should have designs that “can support causal conclusions (i.e., studies with high internal validity), and studies that in total include enough of the range of participants and settings to support scaling up to the state, regional, or national level (i.e., studies with high external validity).” Interventions that enter the SIF with a **Strong** level of evidence would have conducted either one large, multisite RCT or QED study or several smaller RCT or QED studies either in different locations or with different populations.

II. ASSESSING INCOMING LEVEL OF EVIDENCE

Although there are several factors to consider when assessing an intervention’s level of evidence, this rubric focuses on two important sets of factors: 1) the similarity of the intervention under consideration to the previously studied intervention(s) in terms of where and how they were implemented; and 2) the type of study or studies previously conducted.

- 1) Similarity of the intervention under consideration to previously studied interventions: Identify how the previously studied interventions relate to the intervention you are considering in the following ways:
 - a) Was the intervention implemented by your organization or a different one?
 - b) How closely matched is the previously studied intervention to the proposed intervention? In other words, was the studied intervention identical or very similar to the proposed intervention in terms of content, delivery or target population, or was it substantially modified, adapted, or combined with other interventions?
- 2) Type of study conducted: Identify which types of research or evaluation designs were used in prior studies of the proposed intervention. Only consider studies which yielded positive results.
 - What types of studies showed positive results - rather than null or negative results - for the outcomes targeted by the applicant program?
 - For example, are there positive results from studies that have used designs such as pre- and post-tests with a single group? Are there studies that used a matched comparison group? Was there a randomized controlled trial?

III. ISSUES TO CONSIDER

Additional issues to consider when assessing incoming evidence:

- 1) Adapting an intervention or combining multiple interventions may lower the assessed evidence level.
- 2) Even with the same study design (e.g., a single site RCT), an intervention using evidence from studies of a similar intervention may have a lower assessed level of evidence than an intervention using studies from the identical intervention.

- 3) Unless an intervention is being intentionally replicated with fidelity, studies for the same intervention conducted by a different program or organization may also offer lower levels of evidence than studies conducted by the proposing subgrantee.
- 4) A study or studies conducted in a different organizational context than the one being proposed does not likely have sufficient evidence to be considered preliminary under SIF standards. This is due to the fact that the preliminary evaluations ((i.e., single group pre-post-tests) do not have sufficient internal validity to show that the program “causes” the outcome. For studies that only offer pre-post testing, it is possible that something in the program context other than the intervention (e.g., how participants are selected) may be causing the changes seen by that program.

IV. USING THE RUBRIC

To use the Evidence Level Review Rubric on the next page and find an intervention’s incoming level of evidence:

- 1) Review each previously conducted study. Identify those that generally show positive, rather than null or negative, results for the outcomes targeted by the applicant’s program.
- 2) Determine the connection of the proposed intervention to the studied intervention and use the labels in the top row, “Similarity to Proposed Intervention,” to select the column that best represents how that study relates to the proposed intervention.
- 3) Put a check in the box(es) of the column selected in step 2 above, where it intersects with the row for the design type used in the study from the choices in the left hand column “Study Design.”
- 4) Review each study and check each relevant box.
- 5) After following this procedure for each study, the highest-ranked checked box (e.g., Preliminary, Moderate, Strong) is the level of evidence for the proposed intervention.

INCOMING EVIDENCE REVIEW RUBRIC

Intervention: [Click here to enter text.](#)

Highest Ranked Checked Box: [Click here to enter text.](#)

		How matched – by organization and similarity - is the previously studied intervention to the proposed Intervention? Was it done by:				
		A different organization doing a similar, but not identical intervention?	A different organization doing an identical intervention? (<i>The proposed intervention will be replicated with fidelity</i>)	The same organization doing a combination of interventions that include the one studied?	The same organization doing an intervention that is similar, but not identical to the studied intervention?	The same organization, doing exactly the same intervention?
What type of design was used for studies where the results on relevant outcomes were positive?	None or none known	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()
	Implementation only	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()
	Pre-post testing	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary ()	Not yet preliminary/ Preliminary (Depending on the extent of similarity) ()	Preliminary ()
	Pre-post or post only with non-matched comparison group, or interrupted time series with no comparison group	Not yet preliminary ()	Not yet preliminary ()	Preliminary ()	Preliminary ()	Preliminary ()
	Single site, well designed and implemented QED or RCT	Preliminary ()	Preliminary ()	Preliminary ()	Preliminary ()	Moderate ()
	Two or three well designed and well implemented single site RCTs or QEDs	Preliminary ()	Moderate ()	Preliminary ()	Preliminary/Moderate (Depending on extent of modification) ()	Moderate ()
	National/large scale multi-site well designed and well implemented QED or RCT, or multiple (three or more) well designed and well implemented QEDs or RCTs in different locations	Preliminary ()	Strong* ()	Preliminary ()	Preliminary/Moderate (Depending on extent of modification) ()	Strong* ()

*(To be designated Strong+ and to be exempted from the requirement to attain moderate evidence with its SIF evaluation, the program would need an extensive, multi-site history of RCT's/QED's with the population in question.)

EVALUATION PLANNING PROCESS

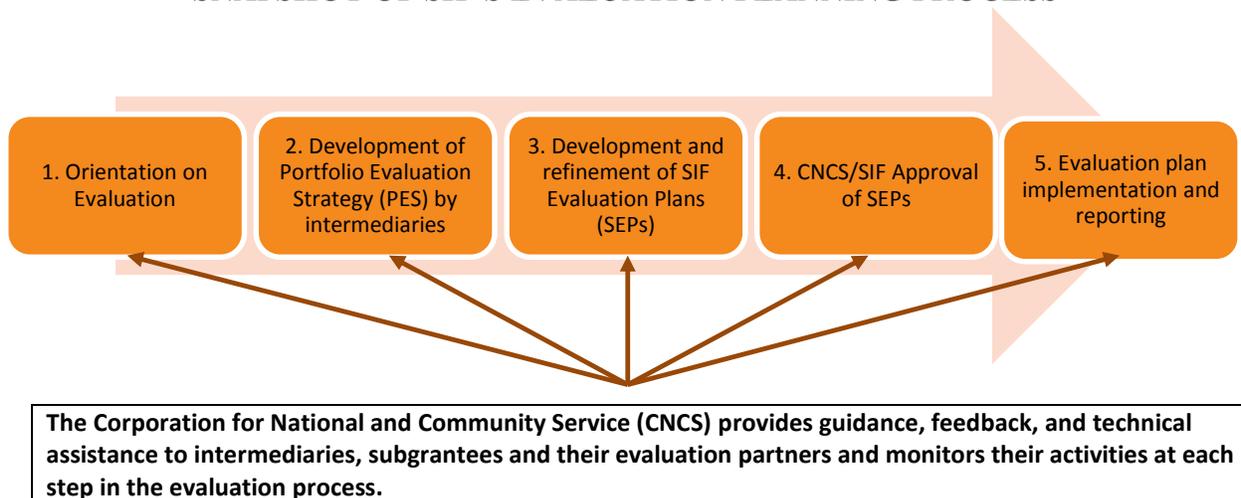
OVERVIEW

Interventions funded through the SIF must enter the program with at least a preliminary level of evidence. Then through rigorous, external, and unbiased evaluation, each intermediary is expected to ensure that the evidence base behind its funded intervention(s) is advanced and provides moderate or strong levels of evidence of impact by the end of the three- to five-year subgrant period. SIF’s definitions of preliminary, moderate, and strong levels of evidence can be found in the NOFA and in “Rubric for Identifying Incoming Levels of Evidence.”

I. THE PLANNING PROCESS

The diagram and description below outline the evaluation planning process you will undergo if you are selected as a SIF intermediary.

SNAPSHOT OF SIF’S EVALUATION PLANNING PROCESS



- 1) After being selected for a SIF grant award, intermediaries receive an orientation by CNCS staff on SIF’s evaluation expectations.
- 2) Intermediaries select subgrantees to implement interventions with at least a preliminary level of evidence.
- 3) Once subgrantees are selected, intermediaries submit a Portfolio Evaluation Strategy (PES) plan to CNCS for discussion and feedback. The PES provides a framework for CNCS and the SIF intermediaries to discuss how the intermediary will approach the evaluation of its portfolio of subgrantees.
 - CNCS and its evaluation technical assistance provider discuss the PES with intermediaries to clarify any questions, develop a shared understanding of the approach, and offer advice as intermediaries begin to develop their SIF Evaluation Plans (SEPs) using the SEP Guidance document provided by CNCS.

- 4) The next step in the evaluation planning process involves the drafting and submission of the SEP, following the guidelines provided in the SEP Guidance, which lays out in detail what should be included. The SEPs submitted to CNCS typically go through two rounds of review, feedback and revision before they are finalized, a process which typically takes six to eight weeks. The plans are approved by CNCS only when all outstanding questions, issues, or concerns about the proposed evaluations are considered and addressed.

Approved plans provide the basis for the rigorous evaluations conducted by third-party evaluators as part of the SIF program. It takes about nine months to one year following the grant awards for intermediaries to have approved evaluation plans in place.

II. ROLE AND IMPORTANCE OF PARTNERSHIPS IN SEP DEVELOPMENT AND IMPLEMENTATION

SIF intermediaries typically work with an evaluation partner that could be internal or external to the organization. Internal evaluators are typically staff members who hold positions as evaluation officers at the intermediary organization. More frequently, evaluators are external to the organization and may be individual consultants, evaluation consulting firms or university or university-affiliated partners. Intermediaries and their evaluation partners usually work together to oversee the evaluation work of subgrantees and provide evaluation training and technical assistance as needed. In limited cases, the intermediary evaluator conducts evaluation of all supported interventions. In others, the external evaluator works across the portfolio to support the intermediary and the subgrantees in their evaluation activities. The subgrantees also work with other external evaluation partners who develop and implement the SEPs at the local level. Partnerships are often defined in contractual agreements at the intermediary and subgrantee level. CNCS supports intermediaries and their evaluation partners through its own set of activities. Although CNCS is in direct communication with SIF intermediaries only, as needed and at the discretion of the intermediaries CNCS communicates with subgrantees and evaluation partners.

BUDGETING FOR RIGOROUS EVALUATION

OVERVIEW

The Social Innovation Fund requires that intermediaries select subgrantees that have experience conducting outcome evaluations, or are implementing programs supported by some existing evaluation evidence. Grantees are required to conduct third-party evaluations that advance the evidence base for the funded programs and increase the number of interventions with moderate and strong levels of evidence of effectiveness. SIF evaluations can utilize a range of study designs including experimental, quasi-experimental and non-experimental designs; however, experimental or quasi-experimental designs are required for an evaluation plan to be approved as providing causal evidence. Most evaluations include a combination of implementation studies and impact evaluations that provide causal evidence regarding program effectiveness within the timeframe of SIF funding (usually three to five years).

Because evaluation is a key component of the SIF, intermediaries and subgrantees are encouraged to allocate sufficient resources to ensure that commissioned studies produce scientifically valid and rigorous evidence. The challenge faced by many SIF intermediaries and subgrantees, however, is how to identify what is "sufficient" before external evaluators are hired, and evaluation plans and detailed evaluation budgets are developed.

When estimated evaluation budgets are realistic and adequate, the evaluation experience can be very positive and informative. In general, evaluation budgets should be:

- Commensurate with stakeholder expectations and involvement;
- Appropriate for the research design used and key questions to be answered;
- Adequate for ensuring quality and rigor, and;
- In line with the level of program and organizational resources available.

The following evaluation budgeting guidelines are based on a review of evaluation and program budgets for 2010 and 2011 SIF intermediaries and on their experiences and reflections.

I. KEY FINDINGS

- The rule of thumb ratios in use to date (i.e., between 5% and 10% of the total budget allocated for evaluation) result in serious under-budgeting of evaluations seeking to address both impact and implementation. Available data indicate that between 15% and 20% is more realistic for single site quasi-experimental designs (QEDs) and randomized controlled trials (RCTs), with some designs (e.g., multisite RCTs, designs with intensive implementation studies) requiring 25% or more.
- In general, using a percentage of program budget is not an ideal method for allocating evaluation funds. Evaluation and program costs should be considered in absolute dollar amounts as well as in relative terms. For example, you likely cannot conduct an evaluation that targets a moderate level of evidence as defined by the SIF for less than \$75,000 per year, unless your study is subsidized (e.g., you receive pro-bono services from the evaluator).

- Evaluation costs and evaluation-to-program budget ratios vary based on the study design chosen and increase with designs that seek to establish causal impact.
- The price of evaluation goes up as the level of evidence desired goes up. Strong evidence is disproportionately more expensive. One driving factor is whether or not the study is conducted across multiple sites.
- All design types have the potential to be expensive.
- There are types of studies yielding preliminary evidence that can be very costly (e.g., multisite implementation studies).
- It is not possible to conduct a rigorous evaluation on a shoestring budget, and in order to conduct a robust evaluation that targets a high level of evidence you have to budget accordingly.

II. FACTORS THAT INFLUENCE BUDGET ESTIMATES

Consider the following when developing an evaluation budget estimate:

PROGRAM FACTORS

- The number of sites studied. The greater the number of sites and distance between sites, the higher the costs for any face-to-face or observational data collection. Further, programs with multiple sites will need to include either all or a sufficient number of sites (depending on sampling approach), if targeting moderate or strong levels of evidence.
- The type of population targeted and program services delivered as well as anticipated challenges in collecting information on that population for those services.
- Overall program and organizational budget and resources available.

EVALUATION DESIGN FACTORS

- The level of technical assistance and capacity building provided to grantees. Providing evaluation capacity building to grantees requires dedicated resources and will increase overall evaluation costs.
- The level of stakeholder engagement. Higher levels of engagement and stakeholder requirements will mean more time dedicated to evaluation, making it more costly. On the other hand, there are significant benefits of involving stakeholders. Engagement of stakeholders can facilitate development of a shared understanding about the program and evaluation, increase buy-in, and result in greater use of evaluation findings in decision making.

- The type of study design and the key questions the study intends to answer. More sophisticated evaluation designs are typically more expensive, and some research questions can be more difficult and costly to answer than others.
- The types of data collection strategies and data sources used and the level of effort associated with implementing data collection approaches. Surveys, for example, may be more costly to develop, pilot test and implement, compared to accessing existing data sources, particularly when the follow-up period is long.
- The amount of time and level of technical expertise required to conduct data analysis and interpretation. Analysis of data using sophisticated statistical methods will require technical expertise which can increase the cost of evaluation.

DISSEMINATION AND USE FACTORS

- Amount of time and resources needed to document evaluation findings and prepare reports, briefs, presentations, and other evaluation deliverables.
- Time and resources needed by internal audiences and stakeholders to discuss and reflect on evaluation findings, and ensure findings are used for informing decision making within the organization as needed (internal utilization).
- Time, effort, and resources required to implement (external) communications and dissemination plans around the evaluation study and its results and sharing of lessons learned in the process.

CATEGORIES OF COSTS

There are two key areas of costs to consider when planning for evaluation. The first is the cost of the evaluation itself; that is the time, materials and other direct costs expended by the evaluation team on evaluation activities. These costs are typically included as line items in a detailed evaluation budget. The second category includes program costs for supporting the evaluation. These are not typically included in an evaluation budget, but represent real costs nonetheless. These costs can include program staff and volunteer time spent in evaluation planning, oversight and supervision; data collection, entry and review; report development; program staff travel to support the evaluation; and other supplies and support.

When planning for evaluation, consider and set aside resources to support both of these cost categories. Failure to allocate sufficient resources for contracted evaluators can negatively impact the quality of the evaluation and the level of evidence attained. In addition, failure to establish clear expectations for program staff participation in the evaluation can lead to a perception on the part of program staff that evaluation activities are an additional or unreasonable burden.

COSTS TO INCLUDE IN A DETAILED BUDGET

	Evaluation Team	Program Staff
<p>Staff time to conduct, or support evaluation activities Include evaluation staff and subcontractor salary and benefits and consultant time to conduct activities. Some contractors may provide separate line items for salary and benefits, while others may present a single, loaded rate, which includes salary, benefits indirect rates and fees.</p>		
Evaluation planning (e.g., development of written evaluation, sampling, analysis and reporting plans, if needed)	x	x
Instrument selection, development, and any needed validation	x	x
Development of Institutional Review Board (IRB) packages	x	x
Data collection, entry, cleaning, and coding	x	x
Data analysis	x	
Reporting (e.g., funder-required & evaluation-specific reports)	x	
Review and acceptance of reports		x
Travel required by the evaluation (e.g., to and from data collection and reporting activities)	x	
Interfacing for project and contract management	x	x
Development of evaluation capacity building/training activities	x	
Participation in training/capacity building	x	x
<p>Travel Travel expenses for staff and/or evaluators should be included as a line item in the budget. Travel costs vary from project to project. Projects across multiple sites around the country will likely need larger travel budgets compared to those located in one site. Proximity of the evaluator can also affect travel costs. There may be travel costs associated with data collection, capacity building activities, communication and dissemination plans. Ideally travel should be estimated in association with specific tasks such as data collection or reporting. Detailed travel estimates should include separate line items or a breakout for airfare, ground transportation, lodging, and per diem/meals/incidentals.</p>		
Airfare	x	x
Ground transportation	x	x
Lodging	x	x
Per diem/meals/incidental travel costs	x	x
<p>Other direct costs Other costs associated with the evaluation should be detailed in the budget.</p>		
Communications—postage, telephone calls, etc.	x	x
Printing and copying—including both task-specific and general duplication	x	x
Supplies and equipment that must be purchased or rented for the evaluation	x	x

CONSIDERATIONS FOR DEVELOPING EVALUATION BUDGETS

- Start by thinking about the expectations of the evaluation. This can include requirements set forth by funding agencies (e.g., evaluation requirements articulated by the Notice of Funding Opportunity) as well as expectations of program staff and other internal or external stakeholders.
- Consider the expected level of engagement from stakeholder groups, including your board of directors, senior leadership, program and evaluation personnel, and others. How much time will different groups need to devote to planning, implementation and dissemination of evaluation findings, and how much time and effort will be needed to ensure buy-in and uptake?
- Consider whether the evaluation approach will require capacity building (i.e., training, technical assistance, and coaching). If so, estimate the internal and external costs associated with evaluation capacity-building activities with attention to the modes of delivery of activities and their intensity.
- Think about who the key partners (evaluation firm[s], consultants, program partners, capacity-builders, etc.) should be for this work to move forward.
- Consider existing evidence regarding the program from past evaluations and how that base of evidence can be advanced. What approach makes most sense for evaluating your program in terms of overall design and targeted level of evidence?
- Develop or request from your evaluation contractor a detailed budget worksheet and estimated time and costs on a weekly, monthly or annual basis (as appropriate) for all items and cost categories relevant to your evaluation needs. The timeframe reflected in the budget worksheet may need to be extended to cover relevant preparation and follow-up periods as well as the timeframe of the evaluation itself.
- Get multiple eyes on the estimated budget to make sure you have covered all the bases.
- Plan for contingencies that may arise over the course of the evaluation. Be prepared to revise and adjust the budget as you move forward from planning to implementation. Often budgets need to be revisited due to realities on the ground over the course of project timeline.
- Plan to ensure that evaluation processes and findings are used to inform and improve your work.