



Building Organizational Capacity for Rigorous Evaluation: Lessons Learned from the Social Innovation Fund

"The SIF will support the use of evidence in several ways. First, the SIF will prioritize intermediaries that use rigorous evidence (see Section V) to select and invest in their subgrantees. Second, the SIF will encourage the use of data and evaluation tools by both intermediaries and subgrantees to monitor the replication and expansion of their programs. Third, the SIF will evaluate the efforts of intermediaries and their subgrantees to achieve measurable outcomes. Finally, the SIF seeks to connect the efforts of government, grantmakers, other donor institutions, and individuals to use evidence and evaluation in systematic ways. These efforts aim to help both SIF-funded programs and the nonprofit and philanthropic communities as a whole."
— 2010 Social Innovation Fund Notice of Funds Availability

Citation

Dietz, N., & the Urban Institute. (2017). *Building Organizational Capacity for Rigorous Evaluation: Lessons Learned from the Social Innovation Fund*. Washington, DC: Corporation for National and Community Service, Social Innovation Fund.

Acknowledgements

The Social Innovation Fund and Urban Institute would like to thank the following individuals for their contributions, revisions, and insight into this document: Christina Garcia and Kristin Lardas, REDF; Teresa Power and Gabriel Rhoads, Edna McConnell Clark Foundation; Dean Elson and Michael Lombardo, Reading Partners; Bryan Hall, BELL (Building Educated Leaders for Life); Navjeet Singh, National Fund for Workforce Solutions.

Authors

Nathan Dietz and the Urban Institute prepared this report for the Social Innovation Fund through a contract with JBS International.

TABLE OF CONTENTS

Foreword	iii
About the Social Innovation Fund.....	1
Introduction	1
Overview and Lessons Learned.....	2
Profiles of the Grantees and Subgrantees	3
Jobs for the Future, Inc. – National Fund for Workforce Solutions.....	3
REDF	3
The Edna McConnell Clark Foundation	4
Reading Partners	4
BELL (Building Educated Leaders for Life)	4
How Does the SIF Encourage Organizations to Build Evaluation Capacity?	4
Use of Evaluation Capacity in Grantee Selection	5
Goals for the Strength of the Evidence Base.....	6
Planning the Evaluation: Initial Steps	7
Budgeting for Evaluation.....	7
Assessing Your Organization’s Evaluation Capacity	8
Choosing an Evaluator	9
Selecting an Appropriate Evaluation Design.....	10
Managing the Evaluation.....	11
Working with Subgrantees or Sites	11
Working with Other Organizations in the Network.....	12
Day-to-Day Management of the Evaluators.....	14
Learning from the Process: Changes as a Result of Conducting the Evaluation	15
Lessons Learned and Key Takeaways.....	17
Conclusion: The SIF and Evaluation Capacity	20

Foreword

The Social Innovation Fund (SIF) places a premium on evidence-based programs and rigorous outcome evaluations. Grantees and subgrantees of the SIF receive support to build their capacity for planning and managing evaluations. This support fills a critical capacity gap and signals the importance of financing evaluation skills in social sector organizations that are expected to impact their communities.

As this report shows, an organization planning a rigorous evaluation needs to have sufficient financial resources and internal evaluation capacity to undertake a successful evaluation study. Recent research documents the extensive need for assistance that organizations require to begin measuring the effectiveness of their programs.¹ By requiring its grantees to build their own evaluation capacity, as well as the capacity of subgrantees, and by disseminating the knowledge these organizations acquire, the SIF is helping the social sector to improve its effectiveness.

While government agencies, driven by the requirements of the Government Performance and Results Act of 1993 (GPRA), have increased their emphasis on performance measurement and evaluation of social sector grantees, many foundations have lagged behind. According to a 2007-2008 survey of foundation practices, nearly 40 percent of foundations spent less than 1 percent of their grant budgets on evaluation activities.² Fortunately this is changing. SIF grantees take seriously the need to strengthen the evaluation capacity of their subgrantees and help them use evaluation results for program improvement.

In this report, grantees and subgrantees from the first SIF Classic cohort relate the experiences of conducting their required evaluation projects, and how they built their evaluation capacity by doing so. The interviewees discussed the importance of being realistic about what it takes to produce a high-quality rigorous evaluation: financial resources, staff time and expertise, and cooperative relationships with the other organizations in their networks. They should take an active role in managing the project, changing the plan when needed, and striving to build evaluation expertise among their own staff members.

By following the examples described by these SIF-funded organizations, and making use of the SIF's evaluation resources, nonprofit organizations can learn how to design and manage evaluation studies that can demonstrate that they are making a difference in the communities they serve.

Dr. Elizabeth Boris

Founding Director, Center on Nonprofits and Philanthropy

The Urban Institute

¹ Winkler, M.K., Theodos, B., and Grosz, M. 2009. *Evaluation Matters: Lessons from Youth-Serving Organizations*. Washington, DC: The Urban Institute and The World Bank.

² Patrizi Associates. 2010. "Use of Evaluative Information in Foundations: Benchmarking data" [unpublished PowerPoint slides].

About the Social Innovation Fund

The Social Innovation Fund (SIF) is a program of the Corporation for National and Community Service, a federal agency that engages millions of Americans in service through its AmeriCorps, Senior Corps, and Volunteer Generation Fund programs, and leads the nation's volunteer and service efforts. The SIF positions the federal government to be a catalyst for impact—using public and private resources to find and grow community-based nonprofits with evidence of results. The SIF focuses on overcoming challenges confronting low-income Americans in three areas of priority need: economic opportunity, healthy futures, and youth development. To learn more, visit www.nationalservice.gov/sif.

Introduction

The Social Innovation Fund (SIF), established in 2009, is one of the U.S. government's primary "evidence-based initiatives." The SIF is committed to the use of rigorous evaluation evidence to both select recipients of federal funding and to validate the impact of its funded program models. The SIF approach can be described through three core principles:

"(1) many of the most compelling solutions to persistent social problems in low-income communities are being developed in those communities and not in federal offices in Washington, D.C.;

(2) significant impact can be generated for society by proactively identifying the best community-based solutions, strengthening their evidence base, and supporting the growth of their impact; and

(3) the federal government can help drive social innovation by stimulating, focusing and enhancing public-private partnerships and cross-sector collaborations to grow the impact of the best community solutions."³

Almost all federal grantmaking initiatives, following the example of the Government Results and Progress Act of 1993, require their grantees to measure their own performance and report on their progress toward their goals. Many also conduct rigorous agency-level evaluations of their own programs, as recommended by the U.S. Office of Management and Budget.⁴ The SIF makes a unique contribution by helping to grow the impact of innovative community-based solutions that have compelling evidence of effectiveness, and by strengthening the evidence base, capacity, and scale of the organizations it funds.

As noted in a report by the Corporation for National and Community Service's Office of Research and Evaluation⁵, the SIF specifically emphasizes the importance of building a particular type of capacity:

"By developing a strong evaluation process, providing resources to support quality evaluations, and encouraging the SIF's intermediary grantmaking organizations and subgrantees to do the same, the SIF enables stakeholders to improve their own evaluation capacity and processes. This

³ Social Innovation Fund Notice of Funding Availability FY 2015, page 1. Available at <http://www.nationalservice.gov/build-your-capacity/grants/funding-opportunities/2015/social-innovation-fund-notice-funding>.

⁴ Zients, Jeffrey D. May 4, 2012. "Memorandum to the Heads of Executive Departments and Agencies." Available at <https://ojp.gov/docs/sabappendixlombuseofevidencememo062012.pdf>.

⁵ Corporation for National and Community Service, Office of Research and Evaluation. (2014). *The Social Innovation Fund: Pioneering an Evidence-Based Investment Model*. (by Christopher Spera, Adrienne DiTommaso, Mary Morris Hyde, and Lily Zandniapour). Washington, DC: Corporation for National and Community Service, page 2. Available at http://www.nationalservice.gov/sites/default/files/documents/Pioneering%20an%20Evidence-Based%20Investment%20Model_0.pdf.

transparent approach helps achieve another mission-critical goal: sharing best practices and lessons learned from evidence-based grantmaking with the broader social sector.”

The term “evaluation capacity” does not refer to the organization’s ability to conduct sophisticated and rigorous evaluation studies themselves. Instead, as the Office of Research and Evaluation report stresses, the SIF encourages grantees and subgrantees to build their capacity to prepare, manage, and learn from evaluation studies carried out by independent third-party evaluators and to improve their ability to implement needed changes from the study results. The SIF is also committed to identifying and sharing best practices from the experiences of SIF grantees and subgrantees, so that social service organizations can build their own evaluation capacity by drawing on those lessons.

Overview and Lessons Learned

This report focuses on the experiences of five organizations that were part of the first SIF Classic cohort of grantees and subgrantees. These organizations received grants either directly from the SIF or from an intermediary organization, and are among the first SIF-funded organizations within their cohort to complete their evaluations and publish the results. The five organizations interviewed provide insight into the following practical questions about how to conduct a high-quality evaluation:

1. What steps should a grantee take to make sure the evaluation starts properly? What should grantees consider when choosing an evaluator, and how can grantee staff prepare the evaluator so that the evaluation gets off to a good start?
2. What should SIF grantees do to prepare subgrantees and other stakeholder organizations for the start of the evaluation? How should SIF intermediaries and subgrantees involve other organizations from their networks in the planning and administration of the evaluation, and what role(s) should each organization play?
3. How does the experience of conducting an evaluation for SIF change the organization’s culture and its evaluation capacity? How can SIF-affiliated organizations make sure that these changes are positive and lasting?

In interviews, representatives from each of the organizations described how they managed the entire process of conducting an evaluation from start to finish, and reflected on the lasting impact that the process had on the way they approach evaluation and program management. One of the most important lessons that emerged is that to conduct a successful evaluation, organizations need to coordinate the activities of many other organizations that play crucial roles throughout the process. These include the evaluation team, the sites where their program models are actually implemented, and other community organizations that serve as partners and support both the programs and the evaluation.

The interviewees described how their organizations designed and planned their evaluations, and also how they worked with their evaluators and other organizations in their networks to organize and coordinate their activities. The interviews generated a number of practical lessons about how to manage the evaluation process and how to learn from the results to build evaluation capacity. These lessons include:

1. At the outset, be realistic about how ready your organization is for a rigorous evaluation, and choose an evaluation design that is appropriate for your organization.
2. Before the evaluation, try to learn as much as possible about the potential impact of variations in the implementation of the intervention.
3. Recognize that evaluators must be managed, and be prepared to correct course if the process goes off track.
4. Develop an awareness and understanding of the methods used and the process of the evaluation – don’t wait until the end to have the evaluator explain it to you.

5. Make sure that the SIF evaluation is situated in your organization's ongoing evaluation plan, and that your entire network of partners and stakeholders learns from the experience.
6. Strive to generate outcome measures that are most meaningful and relevant, to maximize what the organization gains from the evaluation.

Profiles of the Grantees and Subgrantees

In 2010, the SIF awarded grants to eleven intermediary organizations that distributed grant funds to 160 subgrantees located throughout the country.⁶ This report relies on interview evidence from some of the first SIF-funded organizations to publish results from their evaluations: SIF Classic grantees Jobs for the Future, Inc. – National Fund for Workforce Solutions, REDF, and the Edna McConnell Clark Foundation; and Reading Partners and BELL, two subgrantees of the Edna McConnell Clark Foundation. The Edna McConnell Clark Foundation did not conduct an evaluation of its own under its SIF grant, but its subgrantee organizations have completed and published evaluation studies conducted for the SIF by independent evaluators, as have the other two grantees.

Jobs for the Future, Inc. – National Fund for Workforce Solutions

Since its establishment within Jobs for the Future, Inc. in 2007, the National Fund for Workforce Solutions (NFWS) has invested in Regional Funding Collaboratives which match NFWS funds with funds from other sources to support workforce partnerships among local industry, nonprofits, and community colleges (i.e., Industry Partnerships). These partnerships identify local skills gaps, usually within a specific industry, and design and administer training programs to help low-income workers obtain the skills needed to meet the demands of those industries. JFF/NFWS contracted with an independent evaluator, IMPAQ International LLC, to conduct an evaluation of three Industry Partnerships in the Greater Cincinnati area. The study found that the programs had statistically significant positive effects on general employment rates for participants across all three partnerships, compared with unemployed nonparticipants seeking state employment services. More information can be found on the [Jobs for the Future](#) and the [National Fund for Workforce Solutions](#) websites and [CNCS Evidence Exchange](#).

REDF

REDF creates job opportunities for thousands of Californians with multiple barriers to employment – including dislocated youth, individuals who have been homeless or incarcerated, and those with severe mental illness – in sustainable nonprofit social enterprises in low-income communities throughout the state. REDF has been funding nonprofit organizations running social enterprises since 1997. As part of its five-year strategy to transform how people with many employment barriers transition into the workforce, REDF partnered with the Social Innovation Fund to support and evaluate this strategy. An evaluation conducted by independent evaluators from Mathematica tracked life and work-related outcomes for 282 workers at seven social enterprises over the course of the study, which found that social enterprise workers moved toward economic self-sufficiency and life stability and that social enterprises produced \$2.23 in social benefits for every \$1 they spent (by reducing dependence on government support and increasing tax revenue). More information can be found on [REDF's website](#) and the [CNCS Evidence Exchange](#).

⁶ Corporation for National and Community Service, 2014, page 3.

The Edna McConnell Clark Foundation

Over the last decade, the Edna McConnell Clark Foundation (EMCF) has implemented a grantmaking approach that identifies highly-promising nonprofits serving disadvantaged youth from low-income communities and makes large, long-term investments to strengthen their evidence base and replicate and expand them so they can serve more young people. EMCF, in collaboration with the Bridgespan Group and independent evaluators MDRC, has provided financial, strategic planning and evaluation support to a portfolio of nine youth development and school support nonprofit organizations working in both urban and rural communities throughout the nation. These subgrantees help youth ages 9-24: 1) improve their educational skills and academic achievement; 2) prepare for the world of work and make the transition to employment and economic independence; and, 3) avoid high-risk behaviors such as criminal activity and teen pregnancy. More information can be found on the [EMCF website](#).

Reading Partners

Reading Partners, an EMCF subgrantee, focuses on elementary school students in 37 schools throughout California and in Washington, D.C. who are at least six months behind in reading. Reading Partners recruits and uses community volunteers to provide individualized instruction to help students master the reading fundamentals they need to reach grade level. Reading Partners engaged independent evaluator MDRC to evaluate its program model. The study found that the program more effectively raised students' reading proficiency than other similar reading programs in the same schools that were studied; had a positive and statistically significant impact on students' reading comprehension, reading fluency, and sight word efficiency; showed suggestive evidence that it was particularly effective with students reading at the lowest levels; and cost the host school/school district on average \$710 per student, compared with an average of \$1690 per student for other reading programs offered in the same school. More information can be found at the [Reading Partners website](#) and [CNCS Evidence Exchange](#).

BELL (Building Educated Leaders for Life)

EMCF subgrantee BELL (Building Educated Leaders for Life) operates a middle school program that serves rising 6th through 8th grade students who are performing one to two years below grade level. The goals of the program are to increase student's literacy and math skills and to enhance their social development. BELL provides students with 6.5 hours of daily programming for approximately five weeks, five days per week. BELL worked with MDRC to conduct an evaluation of the Middle School Summer Program. The evaluation found that the program showed suggestive evidence of having had a positive impact on students' math achievement in the fall of the next school year (although the effects were statistically inconclusive); showed no effects in terms of higher reading scores in the fall of the next school year for BELL students compared to non-BELL students; and was implemented as designed (i.e., with fidelity) in the study districts. More information can be found on the [Bell website](#) and [CNCS Evidence Exchange](#).

How Does the SIF Encourage Organizations to Build Evaluation Capacity?

From the start, the SIF has made evaluation capacity a requirement for successful grant applicants, and an ongoing objective for funded grantees and subgrantees. The Notice of Funding Availability (NOFA) spells out the requirements and expectations for all applicants. Over time, the SIF has also developed detailed guidance for funded organizations to

help them plan and conduct high-quality evaluations. The SIF's Evaluation Plan Guidance,⁷ subtitled "A Step-By-Step Guide to Conducting a Rigorous Evaluation," serves as a resource for all organizations trying to strengthen the evidence base for their program interventions.

The SIF's official guidance documents make it clear that the goal of evaluation is not just to submit a required evaluation report that describes the impact of the grantee's program intervention(s), but to strengthen the evaluation capacity of the organization and to build an evaluation culture that encourages the use of data for program improvement. The original SIF Notice of Funding Availability, published in 2010, describes these expectations in detail; other CNCS publications describe how the agency helps organizations meet them.

Use of Evaluation Capacity in Grantee Selection

From the beginning, the SIF has made it clear that it intends to fund program interventions that have demonstrated levels of evidence, and that it expects grantees to be capable of planning and managing rigorous program evaluations. The SIF has strengthened its evaluation requirements and expectations for grantees and subgrantees since the initial grant year of 2010, and has also created a technical assistance process that helps organizations build their evaluation capacity.

However, the SIF has used the same three-tiered system, which is similar to systems used by other federal tiered-evidence initiatives, to define the strength of evaluation evidence:

"Preliminary evidence means evidence that is based on a reasonable hypothesis supported by research findings. Thus, research that has yielded promising results for either the program, or a similar program, will constitute preliminary evidence and will meet the Corporation's criteria. Examples of research that meet the standards include: 1) outcome studies that track program participants through a service 'pipeline' and measure participants' responses at the end of the program; and 2) pre- and post-test research that determines whether participants have improved on an outcome of interest."

"Moderate evidence means evidence from previous studies whose designs can support causal conclusions (i.e., studies with high internal validity) but have limited generalizability (i.e., moderate external validity), or studies with high external validity but moderate internal validity."

*"Strong evidence means evidence from previous studies whose designs 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)."*⁸

The SIF uses these three evidence tiers to measure the strength of the evidence base of a program intervention. Typically, strong evidence is the product of an evaluation that uses *random assignment*, or an *experimental design*, to determine who or what receives or experiences the program intervention. Random assignment ensures that the treatment group (which receives the intervention) and the control group (which does not) are equivalent, so that there can be confidence that any observed post-intervention differences in program outcomes can be attributed to the intervention. Randomized controlled trials (RCTs) are widely regarded as the strongest possible evidence, especially when they are conducted across multiple sites and demonstrate that the program intervention is effective in a variety of contexts.

In a *quasi-experimental design*, data on program outcomes are collected from a comparison group that is not formed by random assignment. Well-executed quasi-experimental designs that do an extremely thorough job of controlling for pre-existing differences between groups are capable of generating strong evidence. However, many studies with quasi-

⁷ Corporation for National and Community Service Social Innovation Fund. 2013. *Social Innovation Fund Evaluation Plan Guidance: A Step-By-Step Guide to Designing a Rigorous Evaluation*. Available at <http://www.nationalservice.gov/sites/default/files/documents/SIF%20Evaluation%20guidance%208%205%202014.pdf>.

⁸ 2010 SIF NOFA, pages 21-22.

experimental designs result in moderate rather than strong evidence because there is less confidence that the differences in outcomes are caused by the intervention.

As CNCS summarized in a 2014 report, the three types of evaluation evidence lead us to different conclusions about the effectiveness of a program intervention:

“Preliminary evidence indicates a reasonable hypothesis with credible research that shows promising results. Moderate evidence demonstrates that the program achieves the desired outcome for a limited population. Strong evidence documents the success of the program, including for larger populations.”⁹

The SIF also considers the applicant’s evaluation proposal, a reflection of the organization’s evaluation capacity, when making funding decisions. Intermediaries choose an evaluation approach that is best suited to their subgrantees and the evaluation or interventions they implement.¹⁰ No matter which approach they choose, all SIF grantees must submit a detailed evaluation plan for each of its funded program interventions. The evaluation plans do not need to be submitted with the application, but all applicants need to demonstrate that they are capable of designing an evaluation plan that covers each of its interventions. In addition, grant applicants need to demonstrate their commitment to growing the evaluation capacity of their subgrantees.

Goals for the Strength of the Evidence Base

Most of the program interventions supported by the SIF can cite some evidence to support their effectiveness, but few of them have undergone rigorous evaluations prior to their SIF grants. As of July 2015, the Social Innovation Fund had approved 86 evaluation plans representing a total of \$66 million in evaluation investment. For most of these interventions (74 percent), the evidence base when they entered the SIF was classified as “preliminary”; only 8 percent of the interventions are classified as having “strong” evidence, and only 18 percent have “moderate” evidence.¹¹

To help these organizations build their evaluation capacity, CNCS begins providing evaluation guidance to new SIF grantees and subgrantees shortly after the grants are awarded. Staff members start by reviewing the grantee’s assessment of the evidence for its program interventions; they also examine the methodology and scope of earlier evaluation projects, including evaluations of other organizations that used similar program interventions. By performing a thorough review of the evidence base for each grantee’s interventions, and reviewing the organization’s own experience with external evaluators, CNCS staff can help grantees and subgrantees develop appropriate evaluation plans.

Once the evaluation plan is finalized and approved, however, CNCS continues to provide guidance to SIF grantees and subgrantees and their external evaluators. CNCS sees the relationship between evidence and program growth as an iterative process that does not end when the evaluation study begins. In fact, SIF specifically recognizes that some organizations may require multiple evaluation studies within the grant period to generate moderate or strong evidence for their intervention, and CNCS works with the organizations throughout the process to generate the appropriate level of evidence.

CNCS’ evaluation technical assistance process ensures that SIF grantees and subgrantees receive support throughout the grant period to conduct their evaluation studies and build their evaluation capacity. This assistance continues up to the

⁹ Corporation for National and Community Service, 2014, page 7.

¹⁰ Nine of the twenty intermediaries that received grants between 2010 and 2012 decided to formulate a single evaluation plan that covered the intervention that all of their subgrantees were implementing. As CNCS’ Office of Research and Evaluation noted in its 2014 report, “Evaluations of subgrantee clusters with similar programs make it possible to compare effective practices, helping to achieve similar outcomes” (page 4).

¹¹ Social Innovation Fund, Corporation for National and Community Service. 2015. *Using Evidence to Scale What Works: 2015 State of the SIF Report*, page 22.

point when the study results are published to the CNCS Evidence Exchange, a resource developed so that other organizations can build their evaluation capacity by learning from their peers.

Planning the Evaluation: Initial Steps

The “SIF Evaluation Plan (SEP) Guidance” document contains a great deal of practical advice for organizations that are planning or considering rigorous evaluation studies. As the document notes, the planning process is an “intensive undertaking,”¹² but grantees report that the step-by-step process has helped them create budgets for both staff time and project costs.

The SIF relied heavily on the experiences of the 2010 grantees and subgrantees to make the SEP Guidance a resource for future applicants and funded organizations. As Gabriel Rhoads, Director of Evaluation and Learning at the Edna McConnell Clark Foundation recalled, “Our first subgrantee evaluation planning process was about a four- to eight-page document; we just put some good things down on paper, and we were ready to go. By the time our second cohort of subgrantees rolled around, I think that evaluation SEP Guidance document was up to about 70 pages and was extremely detailed.” Today, the SEP Guidance contains a wide variety of resources, planning documents and templates, and practical guidance about how to plan, implement and manage every step of a rigorous evaluation study.

Budgeting for Evaluation

Every organization must realize that evaluations in general are expensive, but should also know how the costs tend to increase for evaluations that generate moderate or strong evidence.¹³ Evidence from the first SIF-funded evaluations shows that the old “rule of thumb” – organizations should allocate five to ten percent of the overall project budget for evaluation – is likely to “result in serious under-budgeting of evaluations seeking to address both impact and implementation.”

CNCS has published a [report](#)¹⁴ containing an analysis of the approved evaluation budgets of 2010-2011 SIF grantees and subgrantees. The results from this analysis illustrate how seriously an organization might under-budget for an evaluation if they use the 5-10 percent guideline. Organizations that are seeking moderate evidence from their evaluations planned to spend about 17 percent of their overall project budget, on average, from their evaluations; the average organization seeking strong evidence budgeted 28 percent.

In the report, SIF grantees and subgrantees also relate how they created, and in some cases modified, their evaluation budgets. As the report summarizes:

“Although some grantees had conducted similar evaluations in the past and were able to start with reasonable budget estimates, program officers noted that others had to revisit evaluation budgets after the award to increase funding. Based on these reports, it seems reasonable to assume that a substantial number of evaluation budgets are still likely on the lower bound of feasibility for the planned study designs. Further, some grantees and subgrantees that targeted higher levels of evidence could not develop plans within their budgets that achieved them.”¹⁵

Organizations should also consider several other factors when developing budgets for their evaluation plans, such as:

¹² SEP Guidance, page 2.

¹³ SEP Guidance, page 33.

¹⁴ Corporation for National and Community Service, Office of Research and Evaluation. (2013). *Budgeting for Rigorous Evaluation: Insights from the Social Innovation Fund*. (by Lily Zandniapour and Nicole Vicinanza). Washington, DC: Corporation for National and Community Service. Available at http://www.nationalservice.gov/sites/default/files/documents/Budgeting_for_Evaluation.pdf.

¹⁵ Corporation for National and Community Service, 2013, page 10.

- **“Program factors, including the number of and distance between sites, the program services offered, and the type of population targeted;**
- **Design factors, such as key research questions, data collection strategies (including development and use of surveys), data sources, and amount of time to conduct data analysis;**
- **Management factors, including the need to build staff capacity for the desired evaluation; and**
- **Dissemination and use factors, such as the amount of time to document findings, prepare reports/deliverables, reflect on evaluation findings, and develop formal communications and dissemination plans to share evaluation results.”¹⁶**

Assessing Your Organization's Evaluation Capacity

The first 2010 cohort of SIF Classic grantees began planning for their evaluations with the help of CNCS staff and evaluation technical assistance providers, but without much published guidance about how to meet the SIF's requirements. The work from these early days helped CNCS to create the SEP Guidance document, which was SIF's response to the challenges faced by the early grantees.

The SEP Guidance, which is now [publicly available](#), contains detailed discussions of the most prominent evaluation designs – including practical tips for how to implement them – that can help organizations choose the evaluation design they want to conduct. The Guidance also contains a detailed checklist that organizations can use to plan a wide variety of impact evaluations, as well as feasibility assessments and implementation evaluations.

In interviews, several of the 2010 SIF cohort members recalled that they planned their own evaluation projects without the help of the SEP Guidance. As Navjeet Singh, Deputy Director for the National Fund for Workforce Solutions at Jobs for the Future, Inc. (JFF), noted, “Apparently, [the guidance] was something of an evolving process [...] I'm not sure we formally went through our evaluation feasibility assessment, or anything like that.” However, according to Singh, JFF had already conducted a baseline assessment of the program intervention, and had begun planning for their SIF-funded evaluation. The organization had committed to a strategy of collecting aggregate data on training outcomes from the workforce partnerships, many of which would later become SIF subgrantees.

Christina Garcia, Managing Director of REDF, also said that her organization did not consult much formal guidance from external sources before planning its evaluation. Garcia explained that REDF's sites, which are profit-making social enterprises, felt comfortable with the collection and use of data for program management:

“REDF as an organization has always been deeply committed to data and data tracking and the ability to really evaluate and demonstrate social impact. It's been something that REDF is deeply committed to and that very much is rooted within the business community. Clearly any business, whether in the nonprofit or for-profit context, has to be very focused on its metrics and being responsive to those metrics.”

Because its subgrantees operate a number of different program interventions, the Edna McConnell Clark Foundation (EMCF) allowed its subgrantees to operate their own programs and hire their own evaluators. As Gabriel Rhoads, EMCF's Director of Evaluation and Learning, noted, EMCF specifically selected its subgrantees based on “an orientation towards learning, where they're really embarking upon this work as something they're doing to improve program quality, not just to check a box that they've got evidence.” When selecting its subgrantees, EMCF placed less emphasis on the strength of a subgrantee's evidence base than on its evaluation capacity, as Rhoads explained:

“And the idea is that even if you have great evaluation results, that's fantastic; we'll learn from it; we'll see what you can do to even further improve. If you have mixed or confusing evaluation results, disappointing evaluation results, you are able to learn enough to be able to make changes and move forward and improve.”

¹⁶ Corporation for National and Community Service, 2013, page 11.

Reading Partners and BELL had differing amounts of previous experience with rigorous evaluations, but EMCF awarded subgrants to both organizations in 2010. BELL's elementary school intervention had been the subject of a random-assignment evaluation (or RCT) several years earlier; although Director of Evaluation Bryan Hall was not with the organization at that time, he noted that "we definitely had some institutional knowledge about some of the risks and opportunities of doing studies like this."

Michael Lombardo, Chief Executive Officer of Reading Partners since 2006, described how his organization built its evaluation capacity step-by-step, commissioning a study by Stanford Business School students and then a study that was more rigorous, but limited by a small sample size and "some acknowledged shortcomings in the structure." According to both Hall and Lombardo, by the time both organizations became EMCF and SIF subgrantees, they felt ready for an RCT and eager for the opportunity to participate in one.

Choosing an Evaluator

The SEP Guidance offers valuable practical advice to organizations seeking to hire an external evaluator for their SIF-sponsored evaluations. The Guidance stresses *qualifications* and *independence* as key criteria for selecting the right evaluator. When assessing the qualifications of an evaluation team, an organization should consider the credentials of the team members, but should also make sure that the team has experience with the type of evaluation that is being planned: Do they have experience with the proposed design? Have they worked on projects of this size and scope, especially when the design calls for participation from a number of different sites? Independence is important because the evaluator cannot have even a perceived conflict of interest, or a bias for or against positive results from the evaluation.

Selecting the right external evaluator can be a difficult and time-consuming process. According to Singh, when JFF was trying to hire an evaluator for its SIF evaluation, they needed to hold two competitions before finally selecting IMPAQ to conduct the study. In other cases, organizations have an easy time choosing an evaluator that is a good fit for the study and for the organization. REDF's Garcia recalls that they "just talked to several of the premier evaluators around availability and price, and Mathematica really was the best fit for us."

Eventually, REDF and Mathematica developed a "really collaborative process" in which the evaluator participated in the early planning discussions and advised the grantee about which evaluation design to choose. Kristin Lardas, REDF's Director of Learning and Impact, noted that REDF and Mathematica "really cooperated on how we could be less invasive and more effective at understanding how this program operates."

EMCF also considers its evaluator, MDRC, to be its "chief partner in this work" rather than a contractor. The foundation offered all of its subgrantees the opportunity to work with MDRC on their evaluations; only four of the twelve subgrantees chose to work with other evaluation contractors. EMCF allowed each of the other eight subgrantees to formulate its own budget for MDRC's evaluation, but before that, according to Teresa Power, EMCF Senior Portfolio Associate, "we went through a roadmapping process with them to think about how the sites needed the dollars for growth, how they needed them for infrastructure, capacity building, and what dollars were needed for evaluation. So in that sense, we were involved helping them budget."

In their interviews, EMCF subgrantees Reading Partners and BELL both discussed their satisfaction with the working relationship they developed with MDRC. As Lombardo recalls, at Reading Partners the work started right away: "Literally within a week of us getting the amazingly good news about the EMCF grant, we were involved in deep conversations with them, with MDRC and with [EMCF partners] Bridgespan to develop the evaluation plan." BELL's Bryan Hall says, "I really loved working with MDRC and we continue to work with them to this day on things related to this study that are sort of follow up pieces."

Selecting an Appropriate Evaluation Design

While the SEP Guidance contains detailed descriptions of a variety of evaluation designs, the process of choosing the design that's right for a given organization at a given time can be extremely difficult. This is especially true for SIF Classic grantees and subgrantees, given that the SIF's tiered-evidence approach encourages them to strive for the strongest possible level of evidence. The interviewees from all five of the organizations participating in this report recalled that finding the right design and committing to it took a considerable amount of time and energy.

JFF's Singh notes that "in workforce development, there is sort of a reluctance to use randomized control trials, which is based fundamentally on the question of how can we give some people a service and not others." Singh felt that a quasi-experimental design would be much easier to defend and to implement: "There was a belief in general that administrative data would be available, and that one could get data on those individuals, and therefore there would be a control group available" whose outcomes could be compared to those of the program participants.

In contrast, REDF staff realized that they would face the same problems with implementing a randomized control trials, but, as Christina Garcia put it, "We wanted to meet the ambition of SIF and really jump in and shoot for the stars with the randomized control trial." To assess the feasibility of random assignment, REDF's evaluator, Mathematica, visited each of the organization's California subgrantees. As Garcia recalls, "it did become very clear through the feasibility analysis that the field and the maturity of social enterprise as an intervention was not ready for an RCT." The design that REDF finally chose was intended to yield evidence at the "top end of preliminary" according to the SIF's standards – a choice that, as Garcia noted, left the sites feeling much more comfortable about their capacity to support the study. REDF's evaluation ultimately yielded evidence of moderate strength, exceeding the organization's original intentions.

Like REDF, EMCF's subgrantees also had high ambitions for their evaluations, as Gabriel Rhoads recalls: "Each one of them had targets of either moderate or strong evidence. And for the most part, they're aiming for strong, using rigorous experimental designs." Rhoads described the experiences of two subgrantees that are not striving for strong evidence from their SIF evaluations. One is conducting a "mini-RCT" with a small sample that is unlikely to return results that are statistically significant. However, EMCF, the subgrantee's funder, expects the study to return "really high-quality information" that will help both the intermediary and the subgrantee understand how to administer the program more effectively. Another subgrantee that is experiencing problems with fidelity (the sites are not all implementing the intervention the same way, which makes a multi-site evaluation not feasible) is receiving feedback and technical assistance from EMCF about implementation.¹⁷

Even Reading Partners and BELL, which both conducted RCTs with the help of MDRC, faced concerns about whether their interventions were ready for random assignment. Reading Partners' Michael Lombardo remembers that MDRC "had a lot of questions" about whether the program was ready for a large-scale RCT, but were eventually convinced after a long vetting process. Lombardo himself recalls having doubts – what would statistically insignificant findings do to the program's reputation? – but eventually decided that "you learn by taking risks" and that "we felt like this was a smart risk for us to take where we had a decent chance of success."

BELL's Hall also expressed concerns about the risks of an RCT, even though BELL's elementary-school intervention had already undergone a random-assignment evaluation. As Hall recalls, he and his organization saw the risks, but focused more on the opportunities: "I think we saw the middle school sphere as one where there wasn't a lot of evidence, particularly around summertime programs, so we felt that we could play a big role in that. I think also we saw it as a way to just set ourselves apart from many other organizations, the vast majority of whom don't even do one RCT, let alone

¹⁷ The 2015 NOFA required all subgrantees to have evidence that can be classified as "preliminary" strength to be funded. This requirement strengthens language from the 2010 NOFA, which states: "Where [strong or moderate evidence is] not available, the intermediaries will be expected to identify the existing levels of evidence (as defined in Section V of this Notice) of subgrantees and to use SIF resources to help build the evidence-base of these programs." (2010 NOFA, page 16)

two.” An organization that wants to conduct a rigorous evaluation – especially an RCT – needs to weigh the risks and opportunities carefully to determine whether such a design is feasible and appropriate at a given time.

Managing the Evaluation

The [SEP guidance](#) discusses key aspects of the planning and management process, but once the evaluation design is finalized, and the evaluators are ready to start collecting data, the work for an organization undergoing an evaluation really begins. Part of the job for the sponsoring organization is to identify the organizations that will be actively participating in the evaluation, either as study participants (in a multi-site evaluation) or in supporting roles. The sponsoring organization needs to prepare its network of sites and partner organizations for the study, making sure they are ready and willing to cooperate with the study and contribute necessary resources.

Managing an evaluation is also an important opportunity for an organization to build its evaluation capacity, especially if it takes a hands-on role in supervising the work of the evaluator. Watching how an evaluator conducts a study is not only good business – any contractor needs to be monitored – but also provides the organization with an eyewitness view of how to conduct evaluations. The knowledge an organization takes away from its direct involvement in the evaluation will increase its understanding of the study results, and will provide firsthand experience and knowledge about how to manage a rigorous evaluation.

Working with Subgrantees or Sites

Jobs for the Future Inc. has formed longstanding relationships with many of the workforce partnerships in its network, and has devoted considerable time and energy to nurture an evaluation culture that emphasizes the importance of using data for program improvements. In the beginning stages of the evaluation, as Navjeet Singh recalls, JFF and the National Fund for Workforce Solutions worked closely with their partner organizations and evaluation contractor IMPAQ to select sites for participation. First, JFF and the National Fund held a convening for all existing National Fund sites – some of which were also SIF subgrantees – to discuss the SIF evaluation requirement. JFF, the National Fund and IMPAQ realized after the convening that the SIF evaluation should not include all the SIF subgrantees, since some were relatively new additions to the National Fund’s portfolio.

Then, JFF, together with the National Fund and IMPAQ, made what Singh called a “critical decision” to focus only on longtime sites with well-established workforce partnerships. They narrowed the list of possible participating sites to about 25 SIF subgrantees located in 10 regions, using a variety of criteria:

1. Fidelity of the implementation of the program model
2. Effectiveness in raising matching funds
3. Length of partnership with employers and other stakeholders
4. Good working relationship with training providers
5. Sufficient number of participants to do analysis
6. Representation of the broad range of National Fund sites, in terms of region and the industry focus

After using these criteria to assess the pool of potential study sites, JFF added a final consideration: the availability of individual-level data on both program participants and nonparticipants who would form the study’s comparison group. This final criterion narrowed the scope of the study to three states – Ohio, Pennsylvania and Wisconsin – and JFF began working with state government agencies on data-sharing agreements that would provide access to data for all the study participants. According to Singh, after Pennsylvania and Wisconsin balked at sharing participant data, “We said, let’s

keep trying to get data from other states. In the meanwhile, why don't you [IMPAQ] do the analysis on the three partnerships in Ohio.”

REDF seriously considered an RCT design for its evaluation, but changed plans after gathering feedback from sites. REDF's Garcia described the “tensions and ruffled feathers” that resulted when evaluation contractor Mathematica visited all the SIF subgrantees to see if they would be capable of supporting a random-assignment evaluation. The decision to move away from the original plans for an RCT helped to ease those tensions, but, according to Garcia, also left REDF's leadership with a broader awareness of how to build evaluation capacity among their sites: “And so, after hearing those stories, it became really clear to me that REDF needed to be spending its time providing technical assistance, building capacity, and being a different kind of champion and cheerleader.” Organizations that conduct feasibility studies before committing to a full-scale, rigorous evaluation¹⁸ often learn something similar about how to build the evaluation capacity of their sites.

Interviewees from EMCF discussed similar challenges related to working with sites on rigorous evaluation studies. Senior Portfolio Associate Teresa Power observed that organizations in “growth mode,” many of which are focusing on building basic staff competencies about evaluation, present a particular challenge: “The focus can be on balancing competing priorities, rather than making sure that the enrollment is happening into the evaluation.” As an example, Power mentioned an organization that was struggling to enroll enough study participants, and how EMCF “put together a triangle of support between us, the evaluation firm and the grantee to brainstorm how to maintain that high enrollment month to month. And it actually wound up being a success story, but you can understand the challenge that the management teams are facing.”

EMCF's Gabriel Rhoads described the work that EMCF subgrantees have done on implementation challenges while they are undergoing or planning evaluations. One subgrantee that manages several sites of its own is developing a series of checklists that it will use to ensure that its intervention is being implemented in a consistent manner everywhere. As Rhoads notes, the subgrantee's work will also benefit the larger social sector: “And so they're actually going to have two things; one, show their alignment with models that have already been rigorously tested and shown to be the strong level of evidence, and the second being that they are able to add to the field and allow other people to do that same type of work and hopefully in an easier way.” The subgrantee's work also delivers a third benefit: by developing a process for checking the fidelity of the intervention in various places, the organization is developing evaluation capacity that will prepare it for a future rigorous multi-site evaluation.

Working with Other Organizations in the Network

One of the primary responsibilities of SIF Classic intermediaries is to build capacity among their subgrantees – not just evaluation capacity, but the capacity to provide high-quality services that meet community needs. To conduct an evaluation study, SIF grantees and subgrantees often need to secure the active participation of other organizations that are not formally part of the SIF network. All of the interviewees discussed the roles these non-SIF organizations play in an evaluation project, and the strategies that they used to manage their involvement.

In the interview, Jobs for the Future Inc.'s Navjeet Singh discussed the obstacles faced in acquiring administrative data for the SIF evaluation. However, he mentioned that JFF has finalized a data-sharing agreement with Wisconsin, and that JFF is continuing its search for a third state that might be willing to share data for a follow-up study.

Singh also discussed the organization's ongoing relationship with Program and Policy Insights (PPI), whose role in the JFF-National Fund network has shifted from evaluator to technical assistance provider, with a focus on building evaluation capacity among the workforce-partnership sites. PPI produces an annual report on participant demographics from the aggregate data provided by the sites, and also works on annual projects that build awareness about key

¹⁸ SEP Guidance, 2013, pp. 13-14.

programmatic issues. In 2011, PPI published a report based on a survey of employers, “since the key element of our model is that employers or businesses work directly with [the] workforce,” to determine whether their needs are being met by the education and training provided to participants. The 2012 report focused on fidelity to the program model; this report provided valuable information for the site selection process for IMPAQ, and also showed how the program intervention was implemented differently in different sites.

REDF described how it added features to its SIF evaluation based on the interest the study generated within its network, and within its subgrantees’ networks. As REDF’s Christina Garcia recalled, “When we started, it was just a very standard outcome study. As folks were seeing the questions, understanding what kind of data was coming out of it, using that to talk with their partners, there just was a growing interest in further maximizing the evaluation. What everyone wanted was to get every bit of juiciness out of this evaluation effort.”

Acting on this interest from stakeholders, REDF added an interim report and a cost-benefit analysis to the evaluation plan, which focused on impact on participants. Garcia noted that this engagement from stakeholders gave REDF the opportunity to build understanding of the importance of evaluation for social-service organizations. She added that despite the fact that the sites in REDF’s portfolio were social enterprises, “businesses that use data all the time,” most of them “used data very much as a survival mechanism... just looking at growth and revenue to make sure we are meeting the right margins.” Meanwhile, these social enterprises were also “embedded in much larger nonprofits that are doing other things. They’re a workforce entity; they’re a training entity.”

Garcia believes that the evaluation allowed all of the stakeholders – their partners, their sites and their sites’ partners – “to really see that entire continuum rather than looking at it as fragmented pieces... I think that was something that we were able to bring to the table in the way that we engaged with the groups.” The study also made REDF’s stakeholders aware of the implications of the study for adding much-needed strength to the evidence base for social enterprises: “But it became more explicitly clear that in order to scale and drive more resources into this field [of social enterprise], we need to have this evidence. And I think going through the evaluation, it made people more aware around the lack of evidence that existed.”

Because EMCF believes strongly that external viewpoints can strengthen an organization’s evaluation strategy, it has established an Evaluation Advisory Committee and encouraged each of its subgrantees to form one of its own. According to Gabriel Rhoads, who heads the foundation’s own committee, EMCF “reviewed all the evaluation designs with the Evaluation Advisory Committee and made sure that each subgrantee’s Evaluation Advisory Committee, if they had one, had the opportunity to review the design. Everyone tried to poke holes in it and make sure that we had the best design possible.” When the study results were available, the Evaluation Advisory Committees also joined EMCF and the evaluator in reviewing the results with the subgrantees and providing feedback.

Rhoads noted that “sometimes having an Evaluation Advisory Committee is really helpful for the subgrantees -- having leadership that’s really focused on the execution of the evaluation.” Bryan Hall of BELL agreed, describing the committee as “a volunteer group of external experts – many of them are skilled in RCT design and experimental design and were really, really incredibly valuable at just diving into the weeds behind the statistics and the methods and looking at early briefs of the report, and then helping flush out limitations and flush out things that should have more attention brought to them or less attention to them.”

An RCT requires the organization to recruit enough study participants to populate both the control group as well as the program group, which may place a strain on the organization’s recruitment capacity. As EMCF’s Teresa Power pointed out, the success of a rigorous evaluation depends on the organization’s relationships with other organizations that can help with recruitment: “I’d also say programs that depend on referrals from the county or state are somewhat at the mercy of those referrals. And at least our grantees have very good relationships with their government counterparts. But even with the best relationship, there can be slow times when the referrals are not coming in, and that can really affect the evaluation.”

Because the interventions of EMCF’s subgrantees are typically school-based, relationships with school districts are critical to achieve recruitment targets for an RCT. As EMCF’s Rhoads pointed out, subgrantees like Reading Partners tend to work directly with school districts, “so they’re able to have a little bit of a referral network.” However, maintaining and especially building relationships with school districts takes work. Bryan Hall from BELL remembered, “I think we knew that getting district buy-in was going to be a process and was going to require a lot of work,” even though the organization had a network of districts that supported their programs for elementary school-age children.

For their SIF evaluation, Hall noted, “all the partnerships in the study at the time were brand new to BELL. We didn’t have any sort of luxury of a learning curve or having worked through glitches, or being able to pilot anything prior to implementation.” Although the process of managing the evaluation has certainly strengthened these relationships, an organization that already has strong working relationships with other organizations in its network may also find it easier to manage a rigorous evaluation study.

Day-to-Day Management of the Evaluators

The SEP Guidance contains valuable advice about how to develop a reporting strategy, timeline and budget for an evaluation project. Regular reports can help program staff manage the process, and can also keep all the organization’s stakeholders – the organization’s leadership, board of directors, and funders – informed about the study’s progress so they can provide feedback when needed. A reporting schedule could contain the following elements:

- *“Report summarizing the design;*
- *Summary of instrument development;*
- *Detailed description of how sampling was done, including numbers of controls and participants and any issue(s) encountered;*
- *Report on each data collection, including each instrument used, how many were collected, from whom (sites, controls, participants, number completed), and when; and*
- *Detailed report of data analysis conducted, including any issues that arose or deviations from the original analysis plan and summaries of results to date.”¹⁹*

The SEP Guidance provides templates that organizations can use to plan the study’s budget and staffing plan for each phase of the project. The Guidance provides examples of project timetables, in Gantt-chart and in table format, for organizations to use in planning and managing their evaluation studies.

The day-to-day management of a rigorous evaluation involves much more work than management of an evaluation with a less rigorous design. Interviewees from JFF and REDF, which managed quasi-experimental evaluations for the SIF, described their working partnerships with their evaluators, but EMCF’s Rhoads offered much more detail about the process they use to manage subgrantees’ evaluations: “Well, we get quarterly reports from both the grantees and the evaluators to understand if the evaluation is on track or off track. When it looks like something has to materially change about the evaluation plan, we submit an official addendum or update the subgrantee evaluation plan.”

Bryan Hall, from EMCF subgrantee BELL, described the working relationship that emerged among his organization; its funder, EMCF; and MDRC, the evaluators. Hall recalled that BELL’s leadership had much more direct contact with EMCF than he did, but remembers them serving as “an advocate, in a way; a facilitator, helping us to answer certain questions.” His own role was to be “the real point of contact between MDRC and all of our various partners here around a lot of the nitty-gritty details.” He likened the work to working on an “election campaign”:

¹⁹ SEP Guidance, p. 30.

“We came down very much to the last minute in some of these districts around participation in the study. So we were mailing things out, mailing out reminder postcards, making phone calls, following up on consent forms that didn’t have enough information. And then we would call upon the various high level folks when there were sort of much bigger, larger picture issues that needed to be discussed. For example, when we were at a point where we weren’t really sure if we were going to reach the sample size or not, that’s when the discussion began to involve the much larger group of people around some of these much more serious, strategic directions of the study.”

Despite the hectic nature of his work on the SIF evaluation, Hall believes that program staff must take an active role in managing evaluations to make sure the evaluators have access to their subject matter expertise: “You kind of go into this expecting [the evaluators] to be the experts. And to some degree they are, but they also don’t necessarily understand the day-to-day of how your programs work and how your communities work, and how your families apply for the program, and all these other elements that take years and years of experience to understand. So I think that you need to educate your third-party researchers about a lot of things, just like they need to educate you. So knowing that you can have that back and forth conversation, and then knowing when to speak up and push back when necessary is a good approach to take.”

Learning from the Process: Changes as a Result of Conducting the Evaluation

Monitoring an evaluation study will leave the sponsoring organization with a fuller understanding of the results of the study, which will help staff translate the study’s results into program improvements. However, the process of managing of an evaluation study can also help the organization build a “culture” of evaluation that can have positive impacts on its internal management.

All of the intermediaries and subgrantees interviewed for this report described how the SIF evaluation changed the organization’s evaluation capacity and contributed to a culture of evaluation. According to the [evaluation brief](#) produced by CNCS, Jobs for the Future Inc. and the National Fund for Workforce Solutions and their partners “will be examining these results as they undergo their annual assessment of progress against goals. JFF/NFWS is also pursuing learning goals to enable its network, as well as other workforce development practitioners, to learn about effective approaches and exemplary practices for closing needs gaps and improving employment.”

The study, which was released in March 2014, was perfectly timed to fit into “the broader context of the different things we are trying to do as part of our evaluation and learning agenda,” according to JFF’s Navjeet Singh. In 2014, Singh said, JFF and the National Fund “had our evaluators work with each site, and do a survey of what were the key things they had learned in the previous six years, what were their major learnings, and what kind of reports had they done locally? We gathered those. We had a couple of sessions at our peer learning convening, to share those findings, and also to obtain input on what they wanted to learn that we weren’t doing.”

This input went into the learning agenda, which also contains a focus on data integrity and standardization, which Singh considers to be one of the major takeaways from the SIF evaluation: “I think it’s really important to focus a lot on data collection at the ground level in the field; the quality of data collection and consistency across sites. This is one of the things that I think if I were to do differently, I would focus a lot more on that early on.” Because, as Singh noted, “you can’t just do it through evaluators. The program team needs to work to do that.” JFF has made data integrity a priority for its evaluation and learning agenda. This will give JFF an opportunity to stress the importance of standardized data to its sites, which will also make it easier for its evaluators to strengthen the evidence base by producing follow-up reports to the SIF evaluation.

REDF also used the evaluation process to drive its larger agenda: to highlight the value of evaluation for the growth of the social enterprise program model. REDF explained that the study raised awareness about the lack of evidence of the effectiveness of social enterprises, and that the study’s ability “to talk about social enterprise as both kind of an economic

development function and a workforce function got a lot of people really jazzed.” The organization has concentrated on steering this energy and enthusiasm into two overarching goals: building the evaluation capacity of its portfolio of sites, and developing new outcome measures that quantify the value of social enterprises.

REDF understands that they benefit from their work with programs that are attractive in part because they are novel and innovative, and wants to use that to build support for further evaluation. As Kristin Lardas observes: “So everyone is behind funding the mission or the intervention. But when it comes to paying for those operational expenses, how do we actually make this work? How do we build this system that’s going to help us grow and adapt and add some new components or change and tap into the data so we can look back and say, ‘We’ve seen a growth arc’ all the way to the full hilt of real formal, third-party evaluation.” However, Lardas also realizes that REDF needs to work with its sites “to foster more intellectual curiosity around data and less fear. [...] It’s just about leaning in closer to the information and using the results and fielding them in a way that might just better serve your program.”

The second of REDF’s objectives is to build awareness of, and support for, cost-benefit analysis for social enterprises and for social service programs more generally. As Christina Garcia recalled, the idea for the cost-benefit analysis contained in the report originated when the team members realized that they were already collecting the necessary outcome data, and had access to the business and cost data from their subgrantees. The cost-benefit analysis allows REDF to argue that the social enterprise model is addressing “public good and public need but also is doing that in a way that is more efficient than the current kind of approach. Together, that is an incredibly compelling argument” that REDF plans to use to advertise the benefits of working with social enterprises.

EMCF did not conduct a SIF evaluation of its own, but did discuss the active role staff play in their subgrantees’ evaluations. Gabriel Rhoads believes that the SEP is a “win for the grantees” that allows organizations to manage their evaluations effectively and to “really visualize what this is going to look like at the end of the day. EMCF also provides a great deal of support for their SIF subgrantees, even when the original evaluation plan is shown to be not feasible.

As an example, Rhoads cited Children’s Home Society as one of EMCF’s success stories, even though they discovered during their implementation evaluation that the program intervention was not being implemented as planned. Instead of criticizing Children’s Home Society because their original evaluation plan fell through, EMCF praised them for discovering the fidelity problem and for altering the intervention (which they borrowed from another organization) to make it more effective in North Carolina, where the program operates. The SIF encourages intermediary organizations to conduct high-powered impact evaluations that generate strong evidence. However, funders that are willing to sponsor implementation evaluations and feasibility studies are providing their grantees with an invaluable opportunity to build evaluation capacity.

The two EMCF subgrantees that participated in this study, BELL and Reading Partners, both decided to undergo RCTs for their required SIF evaluations. Both also started implementing changes to their program model and internal processes even before the final evaluation reports were published. BELL’s Bryan Hall reports that the process of evaluating the middle-school programs encouraged them to change how they implemented training for all of their staff. However, he observed that BELL had planned these changes before seeing the study’s results: “I wouldn’t say it spurred a lot of new ideas for us; I think it reaffirmed that we were already making the right decisions.” Still, Hall credited the SIF evaluation with encouraging them to be more creative in considering future program changes: “So I think that this study, along with our own institutional knowledge, has prompted us to try to take not a different look, but a more enhanced look at how we deliver these programs and what other value-adds we can provide in our own programming.”

Reading Partners also decided to implement substantial changes to its intervention before the evaluation was completed. According to the CNCS [evaluation brief](#), these changes included: “1) reframing the target population to serve greater numbers of students in earlier grades before reading gaps become very large, 2) enhancing curriculum, assessment, training, and quality assurance systems, and 3) serving more students per site with extended sessions.” Reading Partners will continue to expand into more schools and new districts across the country. In addition, Reading Partners has planned

pilots of additional interventions aimed at increasing reading proficiency that include summer programming, multi-year support of high need students, and engaging additional community volunteers.

Reading Partners' Michael Lombardo recalled that he made the case for expanding the scope of the intervention simply and directly: "I sort of said to the organization, 'I don't think we can tutor our way out of this problem.'" Lombardo and Chief Knowledge Officer Dean Elson noted that some of their community partners and stakeholders objected at first to such broad changes. However, Lombardo defended the changes by arguing that "the end equation for creating a literate nation is going to involve more than just the tutoring strategy. I think our willingness to question the orthodoxy of tutoring is a really important part of how we thought about organizational impact" and how crucial it is for an organization to change its model in response to the results of a well-conducted evaluation.

According to Lombardo and Elson, it took at least as much effort to build support for these changes within Reading Partners as it did to convince external stakeholders of the need for change. Winning over the Reading Partners staff was a "two-year process," according to Lombardo, who also observed: "I would say the one thing that we learned the hard way was just how big the change was, how fundamentally we were shaping people's idea of what we were, and how much time and thoughtfulness it would take to navigate through that."

However, Lombardo said that the urgency of the problem allowed him and Elson to advocate for the strengthening of the organization's evaluation culture: "All of the teams focused on that objective of really making a meaningful difference, and a huge problem nationally made it easier for us to create that internal case, asking 'why is it important to be an evidence based organization? Why is it important that we prioritize [this] kind of rigorous research?'" In an organization with advanced evaluation capacity and a strong evaluation culture, internal and external stakeholders feel comfortable asking and answering such fundamental questions.

Lessons Learned and Key Takeaways

The Social Innovation Fund requires its grantees to have the capacity to conduct high-quality evaluations of each of its major program interventions, and to build the evaluation capacity of their subgrantees. Several of the grantees interviewed for this article felt that this requirement really pushed their organization to build evaluation capacity and the tools provided helped them to do so. Now that several of the organizations in the first SIF cohort have completed and published their evaluation reports, they have many valuable insights to share with organizations that are considering rigorous evaluations.

At the outset, be realistic about how ready your organization is for a rigorous evaluation, and choose an evaluation design that is appropriate for your organization.

The SIF Classic program requires its grantees to conduct evaluations that strengthen the evidence base of their program interventions, and to build the evaluation capacity of their subgrantees. However, organizations need to be realistic about what kind of evaluation design is appropriate in the current circumstances and feasible within a three- to five-year timeframe. Grantees should plan to consult with their external evaluators and internal evaluation specialists, and should also seek input from external stakeholders, as early as possible.

Takeaways for organizations trying to develop an evaluation plan:

- During the planning process, consult resources like the SEP Guidance for step-by-step advice about how to design an evaluation project that's appropriate for your organization. The SEP Guidance was designed for SIF programs, but the tools have wide applicability.
- Select a qualified, independent evaluation team as soon as possible, and work with them closely to develop a plan for the project.

- Be especially realistic about the amount of money, time, and internal staff capacity required to conduct a high-quality rigorous evaluation.

Before the evaluation, try to learn as much as possible about the potential impact of variations in the implementation of the intervention.

Organizations that are planning evaluations should be concerned about the issue of fidelity: whether the study sites have implemented the program intervention in a consistent manner. However, an organization that operates multiple sites also needs to allow at least some room for local variation in the way its programs operate. In many cases, sites administer their programs differently because of the influence of local organizations in their network. For a multi-site program intervention, a fidelity assessment can identify variations that can dilute the power of the evaluation. It can also help the organization and the evaluators develop plans for working with local networks to conduct the study.

Takeaways for organizations planning a multi-site evaluation:

- Learn as much as possible beforehand about the way that each of the sites (or subgrantees) in the study is implementing the intervention that is being evaluated. The evaluation will be strongest when all participating sites are implementing the intervention the same way.
- If the study requires sites to collect and submit data, work with them to make sure that everyone is following a common data collection process and submitting the data in a consistent format.
- Develop a plan for coordinating activities with the other stakeholders in your network, and also the networks of participating sites. This is especially important when the study requires consent or active support from other organizations in order to collect data.

Recognize that evaluators must be managed, and be prepared to correct course if the process goes off track.

The SEP Guidance contains practical advice and other resources that an organization can use to manage a study conducted by external evaluators. A well-developed plan for managing an evaluation includes a strategy and a process for learning about the study's progress from the evaluators, and for gathering feedback from stakeholders at key points. However, organizations must also make sure their own staff is prepared to handle the day-to-day work of managing an evaluation, especially during the data-collection phase of the project.

Takeaways for organizations managing an ongoing evaluation study:

- Develop a plan for internal management of the evaluation process. The plan should specify how to involve stakeholders in major decisions, as well as how internal staff will handle all of the day-to-day logistics of managing the study.
- Ask your evaluators to provide regular written reports on their progress, so that the information is easy to share.
- Make sure your organization's stakeholders know when you plan to ask them for their feedback on key deliverables.

Develop an awareness and understanding of the methods used and the process of the evaluation – don't wait until the end to have the evaluator explain it to you.

During an evaluation study, staff members and leaders may need to explain in detail how the program works for the evaluators. This will help the evaluators develop an appropriate strategy for conducting the evaluation and will help them understand the study's results. The organization's internal study leaders should also try to learn as much as possible about the science of conducting an evaluation. This will increase the staff's ability to manage the study and work

with the evaluators, but will also build internal evaluation capacity that will enable the organization to conduct more rigorous and ambitious studies in the future.

Takeaways for organizations working with evaluators:

- Be prepared to explain the strengths and weaknesses, as well as the risks and opportunities, of the study design to interested outsiders.
- Work closely with the evaluators to develop non-technical explanations of the results that are accessible to stakeholders.
- Work with the evaluators to understand how the study's results contribute to the evidence base for the intervention, and what the results imply about the next steps in the ongoing evaluation agenda.

Make sure that the program evaluation is situated in your organization's ongoing evaluation plan, and that your entire network of partners and stakeholders learns from the experience.

SIF grantees are required to work with their subgrantees to build their internal evaluation capacity, and the evaluation provides many opportunities to strengthen the evaluation culture within the grantee's network of sites. The process of working with the evaluators to collect data for the study can teach sites valuable lessons about how to conduct their own evaluations. Grantees should build on this experience by encouraging their sites to continue to use data for continuous program improvement, to reinforce the importance of a strong evaluation culture.

Takeaways for organizations about building evaluation capacity:

- Consider forming an Evaluation Advisory Committee so that your organization can have ongoing conversations about its evaluation agenda with knowledgeable outsiders.
- Debrief with the evaluators shortly after the study is completed, so that they can help interpret the results and translate them into program improvements.
- While preparing to discuss the study's results with sites or subgrantees, think about ways that they too can use the results to promote continuous improvement.

Strive to generate outcome measures that are most meaningful and relevant, to maximize what the organization gains from the evaluation.

Finally, grantees can strengthen their evaluation capacity, and the evaluation culture within their network of sites and stakeholders, by building on the results of a completed evaluation. Not all successful evaluation studies contain cost analyses, but evaluation evidence that shows that the program's benefits outweigh its costs is extremely powerful for the study's audience, as well as for current and future stakeholders of the organization and its programs.

Takeaways for organizations seeking to maximize the impact of their evaluations:

- Listen to stakeholder suggestions about how to follow up on the results of the evaluation.
- Consider incorporating a cost analysis into the study design, or into the ongoing evaluation agenda, to illustrate how costs vary across sites.
- If the evaluation does not include cost-benefit outcome measures, or a cost analysis, be prepared to explain what the study's results say about the net benefit of the program intervention.

Conclusion: The SIF and Evaluation Capacity

According to REDF's Christina Garcia, the SIF's evaluation requirement "set a really high bar" for the first SIF Classic grantees, but meeting these requirements helped grantees improve their evaluation capacity. The SIF has instituted ambitious requirements for program evaluation that help organizations strengthen the evidence base for their programs and their capacity to conduct rigorous evaluations. More importantly, the SIF has provided tools and resources for organizations to use to design and manage their own evaluations, and is committed to sharing the results publicly to help other organizations build their own evaluation capacity and strengthen their evaluation culture. As more SIF grantees and subgrantees complete their evaluations and publish the results, the SIF's overall impact on the social sector will continue to grow.

SIF Guidance Documents

- **Social Innovation Fund Notice of Funding Availability FY 2016:** This document describes the SIF's current evaluation requirements and how they are used to strengthen evaluation capacity across the social sector. Available at <https://www.nationalservice.gov/build-your-capacity/grants/funding-opportunities/2016/sif-classic-NOFA-FY16>.
- **Social Innovation Fund Evaluation Plan Guidance: A Step-By-Step Guide to Designing a Rigorous Evaluation:** The SEP Guidance can help intermediaries and subgrantees develop robust evaluation plans that meet the expectations of the SIF. It can serve as a resource for others seeking to conduct similar rigorous studies. Available at <http://www.nationalservice.gov/sites/default/files/documents/SIF%20Evaluation%20guidance%208%205%2014.pdf>.
- **Budgeting for Rigorous Evaluation: Insights from the Social Innovation Fund:** The accepted "rules of thumb" for evaluation budgets often understate the amount of money that organizations need to spend to conduct high-quality evaluations. This report provides important evidence about how much SIF grantees and subgrantees budget for rigorous evaluation projects. Available at http://www.nationalservice.gov/sites/default/files/documents/Budgeting_for_Evaluation.pdf.
- **The Social Innovation Fund: Pioneering an Evidence-Based Investment Model:** This report discusses the evaluation requirements of the SIF and how they promote the objectives of the program. Available at http://www.nationalservice.gov/sites/default/files/documents/Pioneering%20an%20Evidence-Based%20Investment%20Model_0.pdf.

"While many organizations may struggle to provide financial support for evaluation, in the SIF Classic program significant time and resources are dedicated to evaluation so that grantees and subgrantees build evaluation capacity and are able to continue to measure their program model's impact into the future."

Social Innovation Fund, Corporation for National and Community Service. (2015). Using Evidence to Scale What Works: 2015 State of the SIF Report.

"The SIF invests in promising programs that address pressing social and community challenges and uses a rigorous process to evaluate each program, building evidence of its impact and effectiveness. The evaluations create a robust evidence base that demonstrates whether SIF-funded programs are achieving intended outcomes and impact on communities. The evidence base also promotes an understanding of how programs achieve success and ways to improve them."

Corporation for National and Community Service, Office of Research and Evaluation. (2014). The Social Innovation Fund: Pioneering an Evidence-Based Investment Model.