Issue Brief #5: Designing and Implementing Rigorous Evaluation for the Social Innovation Fund

A product of the National Assessment of the Social Innovation Fund (SIF)

September 2016
Issue Brief #5:

Designing and Implementing Rigorous Evaluation for the Social Innovation Fund

September 2016

Authors
Xiaodong Zhang, Ph.D.
Johanna Barrero, M.I.A., M.A.

Submitted To
Lily Zandniapour, Ph.D.
Office of Research and Evaluation
Corporation for National and Community Service
For inquiries, please contact evaluation@cns.gov

Acknowledgements
Janet Griffith, Ph.D. and Janet Pershing, Ph.D., of ICF International, provided review and insights that informed the development of this product.

About This Document
This document was produced pursuant to a contract between the Corporation for National and Community Service and ICF International (Contract Number: GS-23F-8182H; CNSHQ13F0091).

ICF International | 9300 Lee Highway | Fairfax, VA 22031-1207 | www.icfi.com

This document was commissioned as an independent study by the Corporation for National and Community Service. Views expressed in the document do not necessarily reflect the official viewpoints of the agency or its staff. This document is in the public domain. Authorization to reproduce it in whole or in part is granted. This document has been formatted to be readable by assistive technologies, in accordance with Section 508 regulations.

Recommended Citation

About CNCS
The Corporation for National and Community Service (CNCS) is a federal agency that improves lives, strengthens communities, and fosters civic engagement through service and volunteering. Each year, CNCS engages five million Americans of all ages and backgrounds through AmeriCorps, Senior Corps, the Social Innovation Fund, the Volunteer Generation Fund, and other programs, and leads the President's national call to service initiative, United We Serve. For more information, visit NationalService.gov.
# Table of Contents

1. Introduction .................................................................................................................. 1
2. Evaluation Context ......................................................................................................... 1
3. Organizational Readiness and Capacity ....................................................................... 2
   3.1 Financial Capacity .................................................................................................... 3
   3.2 Technical Capacity .................................................................................................. 3
   3.3 CNCS Support ......................................................................................................... 4
4. Designing and Implementing the Evaluation ............................................................... 5
   4.1 Choosing a Rigorous Evaluation Design ................................................................ 6
   4.2 Importance of an Implementation Study ............................................................... 6
   4.3 Challenges with Data Collection ........................................................................... 7
   4.4 Dissemination of Results ....................................................................................... 7
5. Organizational and Field Impact ................................................................................. 8
   5.1 Effects on Organizational Practice ....................................................................... 8
   5.2 Effects on the Evidence Base ............................................................................... 9
   5.3 Effects on the Philanthropic Community ............................................................... 9
About This Issue Brief .................................................................................................... 11
About The Social Innovation Fund ................................................................................. 11
1. Introduction

The Social Innovation Fund (SIF) leverages public and private resources to grow community solutions based on evidence of results. The National Assessment of the SIF seeks to document and capture the impact that the SIF has on key program stakeholders. Findings from this multi-component, independent assessment, sponsored by the Corporation for National and Community Service (CNCS) Office of Research and Evaluation and conducted by ICF International, will tell the story of the SIF and identify lessons learned.

The SIF makes grants to experienced grant-making organizations, which in turn identify promising programs within communities through an open and competitive process and distribute funds to high-performing nonprofit organizations. These nonprofits—SIF subgrantees—implement the programs. The federal funds are matched dollar-for-dollar at both grantee and subgrantee levels. All SIF-funded interventions are required to undergo rigorous, independent evaluations to advance the base of evidence for the funded intervention.

This issue brief shares best practices and lessons about how SIF grantees and subgrantees strive to conduct rigorous third-party evaluations. It provides action-oriented recommendations for current and prospective SIF grantees, nonprofit organizations, and policy makers. This brief is informed by interviews with four SIF grantees (also known as intermediaries), two subgrantees, and four evaluators, as well as insights gleaned from other SIF-related sources, including document reviews and interviews conducted for other briefs. With this small purposive sample, the issue brief is intended to illustrate rather than generalize from the range of experiences of SIF grantees, subgrantees, and their evaluators. The brief begins with an overview of the evaluation context before presenting findings with regard to organizational readiness and capacity, designing and implementing rigorous evaluation, and organizational and field impact.

2. Evaluation Context

The SIF supports programs that demonstrate evidence of effectiveness and the potential for scale. Funded programs must show at least preliminary evidence of effectiveness and undergo rigorous, independent evaluations intended to substantially advance the sector’s knowledge base. Each evaluation is based on an approved SIF Evaluation Plan (SEP) — jointly developed by the grantee, the subgrantee(s), and their respective evaluators, with guidance from CNCS — and implemented during the five year performance period.

To date, the SIF has received a total of 108 evaluation plans covering 87 programs and approved 77 plans. These approved SEPs range from randomized controlled trials (RCTs or experimental studies) to quasi-experimental studies using appropriate comparison groups, to non-experimental studies such as process and outcome evaluations, case studies, and cost-benefit analyses (CNCS, 2015). ¹ The SIF evaluations have been conducted in a wide range of contexts. Each SIF grantee selects a number of subgrantees to fund. The subgrantees may implement different interventions or multiple replications of a single model. Each unique intervention is evaluated separately. Table 1 lists the SIF grantees, their subgrantees, and evaluation contractors interviewed for this issue brief. Two grantees conducted evaluations of single

interventions at the grant-level; two grantees carried out evaluations of multiple interventions at the subgrant-level.

<table>
<thead>
<tr>
<th>Organization</th>
<th>SIF role</th>
<th>Cohort</th>
<th>Type of Evaluation</th>
<th>Intervention</th>
<th>Evaluator</th>
</tr>
</thead>
<tbody>
<tr>
<td>U.S. Soccer Foundation</td>
<td>Grantee</td>
<td>2011</td>
<td>Multi-site evaluation of a single intervention; quasi-experimental design impact study and implementation study</td>
<td>After-school sports-based youth development program that aims to reduce obesity through physical activity and nutrition education</td>
<td>Healthy Networks Design and Research — impact study only</td>
</tr>
<tr>
<td>Methodist Healthcare Ministries of South Texas (MHM)</td>
<td>Grantee</td>
<td>2014</td>
<td>Multiple interventions, with nine subgrantees evaluated by one evaluator</td>
<td>Integrated behavioral health projects</td>
<td>Health Resources in Action (HRIA)</td>
</tr>
<tr>
<td>University of Texas Rio Grande Valley</td>
<td>Subgrantee of MHM</td>
<td>NA</td>
<td>Multi-site evaluation of a single intervention; Quasi-experimental design and implementation study</td>
<td>Behavioral health model implemented at two family medicine residency clinics</td>
<td>NA</td>
</tr>
<tr>
<td>GreenLight Fund</td>
<td>Grantee</td>
<td>2012</td>
<td>Multiple interventions with six subgrantees running their own evaluations with six separate evaluators</td>
<td>Programs that seek to close the achievement and opportunity gap in children and youth</td>
<td>NA</td>
</tr>
<tr>
<td>Year Up²</td>
<td>Subgrantee of GreenLight Fund</td>
<td>NA</td>
<td>Multi-site evaluation of a single intervention; Randomized controlled trial impact study and implementation study</td>
<td>Professional Training Corps in Philadelphia, PA, that combines skills training, college credit accumulation, and career trajectory work experience for low-income young adults</td>
<td>Abt Associates</td>
</tr>
</tbody>
</table>

3. **Organizational Readiness and Capacity**

Although SIF grantees and subgrantees are high-performing grantmakers³, several interviewed for this brief reported little experience developing, implementing, and overseeing rigorous third-party evaluations as required by the SIF. The SIF grant was their first experience managing large-scale

---

² Year Up is also a subgrantee of another SIF grant awarded to New Profit and Venture Philanthropy Partners, which was not included in this brief.

³ Grantees are existing grantmaking institutions or a partnership of an existing grantmaking institution and a second grantmaking institution, a state commission, or a chief executive officer of a unit of general local government. Subrecipients (subgrantees) are nonprofit organizations.
evaluations aiming to achieve a moderate to strong level of evidence. The funded organizations, particularly those in the earlier SIF cohorts, were surprised to discover the cost of high-quality evaluation and the time commitment required.

3.1 Financial Capacity

The costs of the required evaluations were sometimes higher than grantees expected. When National Fund for Workforce Solutions/Jobs for the Future (NFWS/JFF) issued the Request for Proposal (RFP) for an external evaluator, the proposals far exceeded the budgeted amount. The grantee had to reissue the RFP and go through a second round of competition to find an evaluator whose proposal fit the budget. The U.S. Soccer Foundation had a similar experience when staff realized that their original budget would not cover the cost of the initially envisioned evaluation design. The organization had to request additional funds from its board of directors and contract a different evaluator to conduct the impact study.

In response to the grantees’ desire for more guidance on the costs associated with rigorous evaluation, CNCS published a report titled Budgeting for Rigorous Evaluation: Insights from the Social Innovation Fund. Based on information from 70 SIF-supported interventions from the earlier cohorts, the report pointed out that many SIF grantees and subgrantees faced the challenge of budgeting accurately for evaluation. Later cohorts appear to have benefited from the experience of the earlier ones and the guidance provided by CNCS. According to the Health Resources in Action (HRIA) representative, the SIF evaluation budget was appropriate: “Too often, there’s not enough money set aside for evaluation to implement a rigorous design. In this case, SIF and Methodist Healthcare Ministries of South Texas (MHM) ensured evaluation was a top priority by allocating the appropriate level of funding for an intensive level of study. That’s what was so exciting about this particular opportunity.”

3.2 Technical Capacity

Financial resources alone did not guarantee grantees and subgrantees would have the organizational capacity to carry out a successful evaluation. Managing rigorous evaluations requires specialized technical capacity from these organizations.

---


When U.S. Soccer Foundation was awarded the SIF grant, staff quickly realized that the process would require a greater time commitment and a higher level of involvement than initially anticipated. They decided to assign a staff member to engage with partners in every step of the process and play a significant role in the evaluation. The staff person and the evaluator traveled to every school site participating in the evaluation and met with school leaders to make sure the sites were well prepared for data collection.

Other grantees used external consultants to increase their capacity and help navigate the evaluation process. For example, NFWS/JFF engaged a contractor to help aggregate the data from its subgrantees, and grantee GreenLight Fund hired an external consultant, the Center for Youth and Communities at Brandeis University, to provide technical assistance (TA) to its six subgrantees and their evaluators.

MHM created peer learning opportunities to share evaluation knowledge in order to build the capacity of its subgrantees, most of which are small organizations operating in a low resource area of the United States–Mexico border. Their learning collaboratives meet in-person on a quarterly basis to exchange experiences and ideas related to their evaluations and share presentations on evaluation-related topics. Subgrantees not only benefit from the technical support received from their individual evaluators, but also from each other’s contributions to their collective knowledge.

### 3.3 CNCS Support

CNCS provides not only grant funding but substantial assistance through its Technical Assistance (TA) contractor to support rigorous evaluations of the programs it funds. The TA was offered to grantees, subgrantees, and evaluators engaged in planning, implementation, and reporting for SIF-sponsored evaluations.

Providing sufficient funding to allow for rigorous evaluations was recognized by all interviewees as one of CNCS’s most important contributions. They appreciated the financial commitment from CNCS to building evidence and emphasizing evaluation. Year Up staff noted that “More than anything the funding created the opportunity. None of this would’ve been possible if we did not have the funding that was then matched and leveraged and then matched and leveraged again.”

CNCS’s evaluation guidance, such as the Step-by-Step Guidance on Rigorous Evaluations and Budgeting for Rigorous Evaluations, was equally important. According to a representative of U.S. Soccer Foundation, “The criteria that SIF put into place are good metrics to use for nonprofits across the board. It is a strict set of rules, but they are there to ensure that these studies are up to the bar for funders, key stakeholders, and
other agencies. As we move forward, the quality of those evaluations will be scrutinized more, as funders get more educated on what evaluations look like.”

Grantees also mentioned the value of TA provided by CNCS and its contractor. MHM appreciated receiving TA support during the process of developing and fine-tuning the evaluation RFP to ensure it asked the right questions in the search for an evaluator. HRIA valued the methodological suggestions as well as the understanding of local constraints. GreenLight Fund found the CNCS evaluation officer and the TA contractor to be incredibly helpful and accessible going through the SEP approval process.

Although many grantees appreciated the TA, some evaluators viewed the SEP process as extending over more steps and more time than necessary. The review and TA process involved extensive checklists that, although well-intended, could be applied somewhat mechanically. Rather than focusing on the integrity of the basic designs, the requirement that evaluation plans pass through all review items often led to requests for further clarification and additional paperwork, taking up time and resources from limited evaluation budgets without substantively strengthening designs. These evaluators felt that review and TA resources might have been more efficiently spent if targeted on more challenging evaluation and design issues.

In addition, the participants interviewed believe CNCS could have provided more support in a few areas, including more guidance around the cost of rigorous evaluation designs and dissemination of results. NFWS/JFF would have liked to receive more support from CNCS during its attempts to access state administrative data. The grantee and its evaluator believe that states are more likely to share data with other government organizations than with private organizations.6

**Recommendations for improving organizational readiness and capacities**

- Consider the potential cost of evaluation before committing to an evaluation design.
- Assess the organization’s internal capacity to oversee the evaluation and anticipate any additional human and financial resources required.
- Connect with grantees working on similar issues to learn from their experiences about what worked well and what did not.
- Take advantage of the full array of support, guidance, TA, and peer learning opportunities provided by CNCS and its TA contractor.

4. **Designing and Implementing the Evaluation**

Although there are best practices and common standards for designing and implementing rigorous impact studies, these rules of thumb should not be considered as “one-size-fits-all.” The best evaluation designs are rigorous, context-specific, and appropriate to the design and implementation model of the intervention. Finding an optimal design requires careful consideration and balancing trade-offs.

---

6 Data sharing may trigger Federal privacy laws and may not be feasible, depending on the specific circumstances.
4.1 Choosing a Rigorous Evaluation Design

The evaluation design has concrete implications for the level of evidence that can be claimed about the impact of an intervention. While RCTs generally are regarded as the “gold standard” in evaluation, practitioners often raise several concerns about the use of this methodology. One is that RCTs are often perceived as unfair to individuals assigned to the control group, who do not receive the service. Experienced evaluators and savvy practitioners can often formulate responses to such concerns. Another concern is the perception that RCTs are more costly than other evaluation methods. Although common, this view is not necessarily correct, as it can be more difficult and more expensive to execute a well-designed non-experimental study due to the difficulty of establishing a strong, credible counterfactual. At the same time, there are often sound strategies for lowering the costs of RCTs. In addition, RCTs may not be appropriate when the key questions the evaluation seeks to answer are not best addressed through an RCT. Finally, RCTs are well-suited to mature programs, those that provide direct services (as distinct from systems change efforts, for example), programs that utilize discrete interventions that conform to the evaluation timeframe, ones that have clear program participation and non-participation, and conditions in which the intervention is delivered with sufficient scale in terms of participant numbers to allow for meaningful estimation of causal impact based on the study sample.

In addition, some grantees and evaluators felt that CNCS could have provided guidance on conditions under which alternatives to impact evaluation would be acceptable. “Unforeseen events can make impact studies infeasible—either because the programs are not ready for evaluation, or due to unforeseen problems in implementing research designs. While it is important to salvage rigorous impact studies where possible, clearer guidance and support from SIF on the value and approaches to rigorous implementation studies would be useful,” noted the evaluator from Abt.

Year Up’s evaluation was designed to be an RCT of the overall program. However, the evaluators found themselves changing course midway, as they could not guarantee a large enough pipeline of participants to accommodate a treatment and control group. Instead, Year Up is using an RCT approach to test improvements in the program implementation by comparing one approach to another. U.S. Soccer Foundation was interested in conducting an RCT, but the budget did not allow that design. University of Texas Rio Grande Valley opted for a quasi-experimental design due to ethical concerns about denying access to services to patients in the control group.

4.2 Importance of an Implementation Study

All the organizations interviewed for this issue brief highlighted the value of an implementation study. The U.S. Soccer Foundation staff were happy that rather than roll out the impact study immediately, they focused on implementation during the first year. Their implementation study accomplished two things: it bought time to train subgrantees and allowed the sites to gain experience running the program while preparing for the impact evaluation. However, the impact evaluator noted that showing impact without “ramp up” time would have provided the most credible evidence to support the impact study.
NFWS/JFF did not conduct an implementation study as part of its evaluation due to insufficient funding, and instead opted to devote the resources to the outcome and impact evaluation. The grantee found it challenging to gather implementation data after the evaluation was completed. To inform the impact study and compensate for the lack of a formal implementation study, NFWS/JFF staff invested a great deal of time on the phone with administrators to gather detailed information about each program and its operations.

Early implementation findings for the Year Up program showed that the early launch of the program was unlikely to achieve the model fidelity necessary to warrant a study of the program’s overall impacts. In concert with CNCS, the Year Up team devised an alternative impact study approach, which uses an RCT design to estimate the impacts of program improvements by testing a randomly selected subset of participants. Year Up sees this design as a win-win for maintaining desired methodological rigor, while concentrating measurement so that results can help to generate a control group to inform and improve the program.

4.3 Challenges with Data Collection

Data collection challenges such as securing data access and recruiting and retaining participants are common when conducting rigorous evaluation. For NFWS/JFF, the research design proposed to use administrative data (such as wage records and unemployment insurance claims) from three states which had data that met the quality criteria for the study. Although the grantee was able to secure data from one of the states, there was a 12-month delay in securing data from the second state, and an 18-month delay in securing data from the third state.

Another related challenge is participant recruitment and retention. Health Networks Design and Research experienced such challenges, particularly with comparison group participants in after-school settings where families tend to be transient, a circumstance further complicated by challenges engaging parents whose children are not receiving the intervention. Offering compensation for participation is considered good practice. However, University of Texas Rio Grande Valley faced a conundrum on this issue. The staff felt strongly that their participants had the right to compensation for the use of their information as well as for their time and effort. However, in order to offer compensation, the university needed to obtain information about the legal status of their patients, many of whom were undocumented. After consulting with other departments at the university, the subgrantee decided to offer incentives to participants and did achieve its recruitment and retention targets.

4.4 Dissemination of Results

Having well-defined goals for disseminating evaluation results can help grantees and subgrantees determine how to get the most out of their evaluation results. Some grantees prioritized the use of results to attract more funding and scale their programs, while others used results as a learning resource to advance organizational and policy changes.
The U.S. Soccer Foundation initially did not have a clear idea of how to disseminate its evaluation results. With help from its evaluator, the Foundation decided as a first step to submit the study results for peer review by a professional association. The statistically significant results were accepted for a presentation at the 2015 American Heart Association (AHA) Scientific Sessions. Since the conference, the Foundation has developed a four-page overview of the evaluation results, citing the report presented at AHA, to incorporate into the organization’s communication materials for partners and funders: “We really feel like that is a credentialed document that unquestionably justifies the effectiveness of our program. That’s been a huge tool for us.”

NFWS/JFF has very actively disseminated its evaluation results. The organization published the results as soon as approved by CNCS and encouraged its evaluator to develop white papers and present at research conferences, with the hope that this would attract funders.

**Recommendations for designing and implementing the evaluation**

- Explore the full range of design options in light of questions of interest, rigor and feasibility.
- Recognize the importance of conducting an implementation study in addition to the impact study.
- Think through all aspects related to data collection, including data access, sampling, recruitment, and retention to anticipate potential challenges and develop solutions.
- Plan for dissemination from the start. Build in dissemination throughout the evaluation process. Determine ahead of time the most effective ways to disseminate your findings and the goals you want to accomplish with those strategies.

5. **Organizational and Field Impact**

The SIF appears to be having an enduring impact on organizational practice, evidence base, and the culture of the philanthropic community.

5.1 **Effects on Organizational Practice**

Participation in the SIF has led some grantees and subgrantees to change the way they do business and build rigorous evaluation into non-SIF activities. For example, when the U.S. Soccer Foundation reshuffled its staffing to hire a dedicated “data person,” that role extended across the organization and will last beyond the SIF. The staff believes that the lessons from the SIF experiences have set the organization up for success for future evaluations. The Foundation recently completed another study of the “Soccer for Success” program. Although the staff did not have to undertake the rigorous SEP process, they found themselves referring back to the SIF criteria to structure the new study.

In another example, MHM is planning to build into its budget the necessary funds to hire an external evaluator and conduct rigorous evaluation of other programs the organization funds, including those not federally funded. With more than 90 agencies to fund, MHM views itself as a long-term funder and believes it makes sense to incorporate this new approach across all its programs.
5.2 Effects on the Evidence Base

The SIF has jump-started efforts to expand promising evidence to a higher standard and in new contexts. Most SIF-funded programs represented in this issue brief started with promising evidence and expected to contribute to the evidence base for these interventions. For example, the evaluation of the Soccer for Success program is one of the first studies providing evidence of the positive effects of sports-based youth development programs in preventing childhood obesity in underserved communities. The evaluation of the NFWS/JFF program is one of the first studies that looks at an industry-focused training program implemented locally.

Another example is Year Up, whose SIF grant involves taking an intensive and expensive stand-alone program and adapting it to operate at lower cost in a college setting. By embedding the program in the college system, Year Up can take advantage of other resources and capacity to make the program financially sustainable and scalable.

University of Texas Rio Grande Valley hopes that SIF results will not only have an impact on the services it provides, but influence the healthcare field more broadly. The subgrantee plans to use the evaluation results to support the expansion of an integrated behavioral health approach to primary care clinics serving the Latino community in low resource areas. They also hope that contributing to the body of work around similar interventions will result in more serious interest in organizations that do meaningful work in low resource areas.

5.3 Effects on the Philanthropic Community

Experiences from SIF participation have not only affected recipient organizations, but have also spurred changes in the broader culture of the philanthropic community. Some interviewees observed that the SIF has contributed to a major cultural shift from head counting to focusing on results, outcomes, and what is actually working.

Despite these changes, there remains a critical need for greater engagement with funders about the meaning of rigorous evaluation, its implementation requirements, and its value. A GreenLight Fund representative urged that funders think about how best to support organizations to build their evidence base and understand what it takes to achieve evidence of impact, because “Everyone is talking about ‘we need to see impact and results’ but without any investment or clear vision around what it would take for an organization to get there.” Therefore, it is very important to understand the level of investment and preparation that a nonprofit organization has to undertake to prepare for a rigorous evaluation.

Recommendations for increasing organizational and field impact
• Develop the practices of nonprofit organizations to promote evidence-based programs and evaluations. Specific practices may relate to the organization's capacity (internal systems, staffing, and technical expertise) or the organization's culture and priorities (its vision of impact and success).

• Make the case for evidence-based programs and evaluations to funders. And request the appropriate budget for rigorous program evaluations.

• Consider how the evidence produced by the SIF evaluation could foster wider adoption of programs with stronger evidence in the field and the broader nonprofit world.
About This Issue Brief

The National Assessment is sponsored by the CNCS Office of Research and Evaluation and conducted by ICF International. This issue brief was informed by the following people from SIF Classic recipient organizations, their subgrantees, and evaluators:

- Zach Riggle, U.S. Soccer Foundation (2011 grantee), and Dr. Danielle Hollar, Healthy Networks Design and Research (evaluator)
- Navjeet Singh, National Fund for WorkForce Solutions/Jobs for the Future (2010 grantee), and Marios Michaelides, IMPAQ International (evaluator)
- Anne Connor and Michelle Brodesky, Methodist Healthcare Ministries of South Texas (2014 grantee), Deepu George, University of Texas Rio Grande Valley (subgrantee), Lisa Wolff and Rebecca Adeigbe, Health Resources in Action (evaluator)
- Kate Barrett, GreenLight Fund (2012 grantee), Garrett Warfield, Year Up (subgrantee), and David Fein, Abt Associates (evaluator)

About The Social Innovation Fund

The Social Innovation Fund, an initiative of the CNCS under the Edward M. Kennedy Serve America Act, is a new approach by the federal government to address urgent national challenges. The fund mobilizes public and private resources to grow the impact of promising, innovative community-based solutions that have evidence of compelling results in three areas of priority need: economic opportunity, healthy futures, and youth development.

The operating model of the SIF is distinguished by the following six elements:

Innovation | Evidence | Scale | Grantmakers | Match | Knowledge Sharing