Meta-Synthesis of Evaluations across the SIF: Interim Report

ICF INTERNATIONAL

Corporation for NATIONAL & COMMUNITY SERVICE ★★★

Meta-Synthesis of Evaluations across the SIF:
Interim Report

September 2015

Author
Xiaodong Zhang, Ph.D.

Prepared For
Lily Zandniapour, Ph.D.
Office of Research and Evaluation
Corporation for National and Community Service

Table of Contents

Executive Summary

The Corporation for National and Community Service (CNCS) runs the Social Innovation Fund (SIF), an initiative that combines public and private resources, to grow the impact of innovative, evidence-based solutions to improve the lives of people in low-income communities throughout the United States. As part of the national assessment of SIF, ICF International conducted a meta-synthesis of 26 final evaluation reports completed by evaluators for the SIF grantees and subgrantees in order to assess the state of SIF evaluation findings to date (June 2015). The findings from this report are mostly from the 2010 cohort. Because the 2010 cohort operated with less guidance and fewer expectations for rigorous evaluation than did the later cohorts, our ability to generalize to a larger population of later cohorts and future SIF grantees is limited.

Program and report features. The 26 evaluations were conducted by independent external evaluators, selected by SIF grantees and subgrantees who entered SIF with different levels of evaluation funding and capacity to evaluate the SIF-funded interventions. Six of the reports were prepared on behalf of the grantees that received the largest amount of SIF funding; 3 on behalf of grantees conducting a single evaluation across their portfolios of subgrants, and 17 by grantees that funded a range of programs and evaluations by regional subgrantees. Forty-two percent of the reports were in the healthy futures issue area, 38% in youth development (including early childhood), and 20% in economic opportunity. Funded programs employed a variety of programmatic strategies for target populations, including summer programs, one-on-one tutoring, social enterprises, workforce partnerships, community-based health care, and telemedicine. All the programs had multiple key components, consistent with the expectation that addressing social issues requires complex and multi-layered solutions.

Evaluation methods. Building a sound evidence base supported by rigorous evaluations is critical to the SIF. The evaluations reviewed included both implementation evaluations and impact/outcome evaluations. Overall, 22 of the 26 reports evaluated implementation, and 21 evaluated outcomes or impacts (Seventeen reported on both). Most of the 22 reports evaluating implementation addressed exposure and fidelity; a minority of these reports assessed program quality or responsiveness. Thirty-eight percent of the 21 impact/outcome evaluations employed such rigorous designs as randomized controlled trial (RCT) or quasi-experimental design (QED) with matching, which allows causal inferences. A third of the impact/outcome evaluations appeared to be adequately powered in all or at least one outcome. Most of these evaluations used valid and reliable outcome measures to capture the expected outcomes of interventions. Eighty-one percent of the impact/outcome evaluations employed inferential statistics to support the findings.

Overall, of the 26 final evaluation reports, 2 were categorized as having produced a strong level of evidence, 3 as providing a moderate level, and 21 (including 5 with implementation only results) as providing a preliminary level. The level of evidence refers to the rigor of the evaluation rather than whether or not the evaluation generates positive findings. The levels are based on how well a particular evaluation addresses concerns about internal and external validity, with more effective evaluations categorized as attaining strong or moderate levels, and less effective evaluations as attaining preliminary levels of evidence. Targeting and achieving moderate or strong evidence was not required of grantees in the 2010 cohort. For later cohorts, SIF's oversight of grantee and subgrantee evaluations evolved significantly. It is anticipated that a greater proportion of the evaluations in later cohorts will meet the criteria for strong or moderate levels of evidence.

Evaluation findings. Evidence generated by SIF-funded interventions is likely to become one of the SIF's enduring legacies and will inform future innovation and scale up by other nonprofits across the nation. These evaluations, in addition to their contribution to the social innovation knowledge base, can be expected to contribute to knowledge about effective strategies in the fields of child/youth development, economic opportunity, and community health, thus complementing the evaluation efforts of other federal agencies such as the U.S. Departments of Education, Labor, and Health and Human Services.

The evaluation findings show that most of the SIF-funded programs (85%) for which evaluation reports are available were implemented with fidelity. The same percentage of the programs were found to have had a positive impact on all or some target outcomes. The four reports that investigated program costs all concluded that the programs provided low-cost solutions to addressing the target outcomes. Half the reports reflected on challenges,

limitations to the findings, and lessons learned. For example, program-related lessons included suggested improvements to program elements and to working relationships among stakeholders, and other practices related to participant recruitment and retention, length of intervention, modest start-up, cost implications, infrastructure, and expansion to different populations. Evaluation-related lessons concerned adequate funding for data collection, sizeable administrative datasets, tracking longitudinal outcomes, and others, such as statistical power, detail about the control condition, and selection of inexperienced sites.

The SIF has a fully operational pipeline of on-going evaluations that will generate more results in the coming years. The national assessment report planned for 2016 will include additional findings from evaluations completed since this review. In addition to thematic-based meta-synthesis, as more impact findings become available, we plan to use meta-analysis techniques to analyze quantitative findings across studies.

1. Introduction

The Corporation for National and Community Service (CNCS) runs the Social Innovation Fund (SIF), a program that combines public and private resources to grow the impact of innovative, evidence-based solutions to improve the lives of people in low-income communities throughout the United States. The program was launched in 2010 following the signing of the Serve America Act, one of six Obama Administration tiered-evidence initiatives embodying the principles of social innovation. To date, the SIF has received an annual Congressional appropriation of approximately $50-70 million.[1] The program leverages federal funds by granting money to intermediary grantmakers who find, fund, improve, and grow promising community-based solutions with evidence of successful outcomes in three core areas: youth development, economic opportunity, and healthy futures. As described by CNCS, the SIF is characterized by the unique interplay of six key elements:

1) It relies on intermediary grantmaking institutions to implement the program–they take on the role of finding, selecting, monitoring, supporting, evaluating, and reporting on the nonprofit organizations implementing community-based interventions.

2) It is a tiered-evidence initiative that requires all funded programs/interventions to demonstrate at least preliminary evidence of effectiveness, or funding "what works."

3) It requires that all programs or interventions implement a rigorous evaluation that will build on their level of evidence.

4) It charges intermediaries with scaling evidence-based programs—increasing

impact within their community or to communities across the country—and as such, grapples with the field-wide challenge of how best to do so successfully and efficiently.

5) It leverages public-private partnerships to effect large-scale community impact in ways that neither a traditional federal grant nor a philanthropic investment could achieve on its own. This strategy features a unique leveraged funding model to support nonprofit programs.

6) It is committed to improving the effectiveness of nonprofits, funders, and other federal agencies by capturing and learning about best practices and promoting approaches that will generate the greatest impact for individuals and communities.

To date, the SIF program has implemented four rounds of grantmaking (in 2010, 2011, 2012, and 2014) and selected 26 organizations to receive 27 grants. These intermediaries, herein referred to as SIF grantees, have in turn selected 252 nonprofit organizations through open and competitive processes to implement promising interventions in youth development, economic opportunity, and healthy futures. Exhibit 1 summarizes the number of awards by year:

Exhibit 1. *SIF Grantees and Subgrantees by Cohort*

Award year (cohort)	Number of Grants to SIF grantees	Number of SIF subgrantees funded
2010	11	149
2011	5	48
2012	4	15
2014	7	40
Total	27	252

Extracted from CNCS website.

As a federal tiered-evidence initiative, the SIF is committed to using rigorous evidence both to select grantees for funding and to assess the

---

[1] This excludes grantees funded under SIF Pay for Success.

outcomes of the program models. To achieve this goal, CNCS expects SIF grantees and subgrantees to commit significant time and resources to conducting formal evaluations of program models that receive SIF funding. The SIF program office, together with the staff of CNCS's Office of Research and Evaluation and its contractor JBS International, work closely with grantees and subgrantees by providing them with technical assistance in designing, implementing, and monitoring the SIF Evaluation Plan (SEP),[2] collecting best practices to share with the broader social innovation sector, and disseminating evidence of effectiveness for each program model within the SIF. The tiered framework is presented as follows (CNCS, n.d.):

- **Preliminary evidence** means the model has evidence, based on a reasonable hypothesis and supported by credible research findings. Examples of research that meet this standard include: 1) outcome studies that track participants through a program and measure their responses at the end of the program; and 2) third-party pre- and post-test research that determines whether participants have improved on an intended outcome.

- **Moderate evidence** means evidence from previous studies with designs that support causal conclusions (i.e., studies with high internal validity) but have limited generalizability (i.e., moderate external validity) or vice versa—studies that only support moderate causal conclusions but have broad general applicability. Examples of studies that would constitute moderate evidence include: 1) at least one well-designed and well-implemented experimental or quasi-experimental study supporting the effectiveness of the practice strategy, or program, with small sample

sizes or other conditions of implementation or analysis that limit generalizability; or 2) correlational research with strong statistical controls for selection bias and for discerning the influence of internal factors. Moderate evidence requires third-party or external and impartial evaluators.

- **Strong evidence** means evidence from previous studies with designs that support causal conclusions (i.e., studies with high internal validity), which, taken together, 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). Examples of studies that would qualify as strong evidence include: 1) more than one well-designed and well-implemented experimental study or well-designed and well-implemented quasi-experimental study that supports the effectiveness of the practice, strategy, or program; or 2) one large, well-designed and well-implemented randomized controlled, multisite trial that supports the effectiveness of the practice, strategy, or program. Strong evidence requires third-party or external and impartial evaluators.

As part of the national assessment of the SIF, ICF International is conducting a synthesis of evaluations prepared by SIF grantees and subgrantees to capture the state of evaluation findings from the SIF. The cross-evaluation synthesis focuses on the following research questions:

- What are the key programmatic features of the SIF-funded programs?

- What are the methods used to evaluate the SIF-funded programs?

---

[2] Each SIF grantee/subgrantee evaluator is required to prepare a SEP that details the program model to be evaluated and justifies the evaluation approach selected.

- What is the state of evaluation findings from the SIF-funded programs?

This report presents initial findings from 26 final evaluation reports completed as of June 1, 2015.[3] The subsequent report planned for 2016 will also include findings from the evaluations completed since this review, and will use meta-analysis to summarize impact findings.

2. Methodological Approach

This chapter describes the methodological approaches used in this synthesis, including strategies for identifying evaluation reports and synthesizing findings across studies.

To describe the SIF projects and evaluations, we used meta-synthesis technique, i.e., meta-analysis without the effect size transformation, to summarize data and themes that emerged from the evaluation reports across projects. The full meta-analysis technique, including effect size transformation, will be employed in the updated synthesis report planned for 2016.

2.1 Final Evaluation Reports Synthesized

A total of 26 final evaluation reports were included in this synthesis, after their review by CNCS and its technical assistance contractor(Exhibit 2).[4] Four out of the 26 report on evaluations conducted at the grantee level (i.e., UniSEP, which is used when one intervention is implemented across multiple subgrants). The remaining 22 report on evaluations conducted at the subgrantee level (i.e., MultiSEP, used when different interventions are implemented across the subgrants within a grantee's portfolio), which were funded by 6 grantees. In other words, UniSEP reports each covered multiple program interventions whereas MultiSEP reports each examined individual programs. The reports come from programs funded during the first

two SIF cohorts: 7 grantees with 20 subgrantees from the 2010 cohort and 3 grantees with 3 subgrantees from the 2011 cohort.

It is important to note that the SIF had not distinctly articulated the requirement that grantee evaluations achieve a moderate or strong level of evidence until the 2011 cohort. The SIF permitted the 2010 grantees and subgrantees to end their grants with a preliminary level of evidence. As a result, although all the 2010 cohort grantees and subgrantees conducted evaluations, not all conducted impact evaluations and, among those that did, some did not design evaluations to achieve a moderate or strong level of evidence, due to funding limitations, data restrictions, or other factors. Additionally, the experience of the initial 2010 SIF cohort appears, based on anecdotal evidence, to differ in important ways from that of later cohorts. For example, many SIF grantees and subgrantees in the first cohort entered the program with little understanding of the SIF's evidence expectations or of related requirements such as Institutional Review Board (IRB) reviews (Lester, 2015).

To the extent that the mix of organizations or the experience of later cohorts differs, the ability to generalize findings from this report to the larger population of later cohorts and future SIF grantees is limited. The synthesis report planned for 2016 will draw on additional evaluation findings, allowing further analysis, especially of evaluations with more rigorous designs.

---

[3] For REDF, we reviewed a separate implementation report not included in the final evaluation report that focuses on the impact.

[4] Reports completed by grantees/subgrantees but still under review were not included in this analysis.

Exhibit 2. *SIF Final Evaluation Reports Completed and Reviewed by June 1, 2015*

Cohort	Grantee	Subgrantee	Evaluation organization	Program name	Priority area	CNCS first-year award* (number of years)
2010	Edna McConnell Clark Foundation	Building Educated Leaders for Life (BELL)	MDRC	BELL Middle School Model	Youth development	$2,000,000 (3)
2010	Edna McConnell Clark Foundation	Center for Employment Opportunities	MDRC	Center for Employment Opportunities (CEO)	Economic opportunity	$2,250,000 (3)
2010	Edna McConnell Clark Foundation	Gateway to College National Network	MDRC	Gateway to College	Youth development	$2,000,000 (3)
2010	Edna McConnell Clark Foundation	Reading Partners	MDRC	Reading Partners	Youth development	$2,000,000 (3)
2010	Foundation for a Healthy Kentucky	Cumberland Family Medical Center, Inc.	University of Kentucky	Cumberland Family Medical Center	Healthy futures	Not available
2010	Foundation for a Healthy Kentucky	Home of the Innocents	REACH Evaluation	Open Arms Children's Health program	Healthy futures	$250,000 (3)
2010	Foundation for a Healthy Kentucky	King's Daughter Medical Center	University of Kentucky, Prevention Research Center	Mobile Health Services for Rural Kentucky (MHSRK) program	Healthy futures	$124,548 (3)
2010	Foundation for a Healthy Kentucky	Meade Activity Center	University of Louisville	Meade Activity Center	Healthy futures	$250,000 (3)
2010	Foundation for a Healthy Kentucky	Montgomery County Health Department	CHES Solutions Group	Community Health Worker (CWH) program	Healthy futures	Not available
2010	Foundation for a Healthy Kentucky	Norton Health System GHAP	University of Kentucky, Prevention Research Center	Get Healthy Access Program (GHAP)	Healthy futures	$100,000 (3)
2010	Foundation for a Healthy Kentucky	Oldham County Health Dept./Hope Health Clinic	University of Louisville	Hope Health Clinic	Healthy futures	$100,000 (3)
2010	Foundation for a Healthy Kentucky	St. Elizabeth	Innovations	St. Elizabeth Telepsychiatry program	Healthy futures	$100,000 (3)
2010	Foundation for a Healthy Kentucky	St. Joseph Health System	St. Joseph Health System	Community Based Delivery Model: Virtual Care	Healthy futures	$250,000 (3)
2010	Jobs for the Future	National Fund for Workforce Solutions UniSEP	IMPAQ International	Workforce Partnership Programs	Economic opportunity	$3,200,000 (3)
2010	Mayor's Fund	Bronx Works	MDRC	Jobs-Plus	Economic opportunity	$563,250 (3)

Cohort	Grantee	Subgrantee	Evaluation organization	Program name	Priority area	CNCS first-year award* (number of years)
2010	Missouri Foundation for Health	Social Innovation for MO UniSEP	Washington University of St. Louis, Center for Tobacco Policy Research	Social Innovation for Missouri	Healthy futures	$1,000,000 (3)
2010	REDF	REDF UniSEP	Mathematica	Social Enterprises	Economic opportunity	$1,500,000 (4)
2010	United Way of Greater Cincinnati	Cincinnati Arts and Technology Center	Innovations	Bridging the Gap	Youth development	$150,000 (3)
2010	United Way of Greater Cincinnati	Cincinnati Museum Center	Innovations	Early Childhood Science Inquiry Training for Educators (ECSITE)	Youth development	$110,000 (3)
2010	United Way of Greater Cincinnati	Covington Public Schools	University of Cincinnati, Evaluation Services Center	Holmes 180	Youth development	$220,000 (3)
2010	United Way of Greater Cincinnati	Easter Seals	University of Cincinnati, Evaluation Services Center	Transitional Employment Services Model	Economic opportunity	$240,000 (3)
2010	United Way of Greater Cincinnati	Resilient Children Project	University of Cincinnati, Evaluation Services Center	Resilient Children Project	Youth development	$325,000 (3)
2010	United Way of Greater Cincinnati	University of Cincinnati	University of Cincinnati, Evaluation Services Center	UC Degrees Gen-1 Project and Higher Education Mentoring Initiative	Youth development	$275,000 (3)
2011	Mile High United Way	Colorado Parent and Child Foundation	University of Denver	Parents as Teachers (PAT), and Home Instruction for Parents of Preschool Youngsters (HIPPY)	Youth development	$371,097 (2)
2011	U.S. Soccer Foundation	USSF UniSEP	U.S. Soccer Foundation, Healthy Networks Design and Research, Child Trends	Soccer for Success	Healthy futures	$1,000,000 (2)
2011	United Way for Southeast Michigan	Detroit Public Television	Michigan Public Health Institute	Pre-School-U	Youth development	$152,845 (2)

*SIF program has an approximately up to 3-1 matching requirement for federal dollars.

2.2   Review and Synthesis Method

We used meta-synthesis to summarize the evaluation reports. Specifically, we reviewed and synthesized information from the final evaluation reports listed in Exhibit 2, incorporated findings from a systematic assessment by JBS of the quality of the reports, and reviewed other reports published by CNCS and external sources to contextualize the findings.
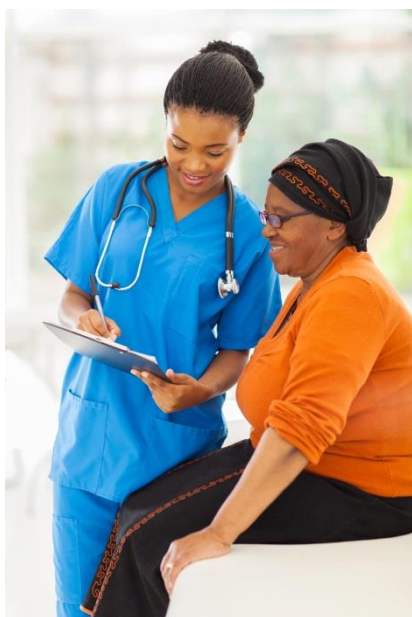
We first coded each report according to the three research questions (programmatic features, evaluation methods, state of evaluation findings), sorting the information further into specific sub-topics. We then summarized emerging themes across reports and illustrated each with examples. To the extent possible, we drew these examples from the SIF's three priority areas (youth development, economic opportunity, and healthy futures). In 2016, in addition to synthesizing the results of additional evaluation reports, we plan to include a meta-analysis with effect size transformation and moderator analysis (see further details in section 4).

The unit of analysis is the evaluation final report, regardless of whether it concerns a grantee or subgrantee, because each report targets a unique program model or intervention. In discussing the substantive findings, the reports are identified by these program names rather than as grantee or subgrantee reports.

3.   Findings

This synthesis summarizes findings across reports according to the three research questions or categories (programmatic features, evaluation methods, and findings), and then sorts each category further into sub-topics. The summaries in this chapter illustrated by examples from specific programs, capture the main themes.

3.1   Programs and Report Features

Information about the characteristics of the funded programs and reports helps describe the nature of the interventions, and provides the context needed to understand the evaluation methods and interpret the findings. In this section, we present data about the grantees, subgrantees, and evaluation organizations; outcome areas and target populations; programmatic strategies; key program components; and number of individuals served.

3.1.1   Grantees/Subgrantees/Evaluation Organizations

As illustrated in Exhibit 2, independent external evaluators prepared the 26 evaluation reports on behalf of various configurations of grantees and subgrantees. Some grantees required their subgrantees to submit their own reports, some submitted a grantee report covering the results of several independent subgrantee evaluations (MultiSEP), and some submitted a grantee report that looked collectively across subgrantee results (UniSEP):

- **Grantees with Multiple Subgrantees Submitting Their Own Subgrantee Reports (MultiSEP)**

  o Edna McConnell Clark Foundation (EMCF) (4 subgrantees)
  o Foundation for a Healthy Kentucky (9 subgrantees)
  o United Way of Greater Cincinnati (6 subgrantees)

- **Grantees Submitting Reports That Cover Independent Subgrantee Evaluations(MultiSEP)**

  - Mile High United Way (1 subgrantee)
  - United Way for Southeast Michigan (1 subgrantee)
  - Mayor's Fund (1 subgrantee)

- **Grantees Submitting a Single Report Across Multiple Subgrantees (UniSEP)**

  - REDF (10 subgrantees)
  - Jobs for the Future (21 subgrantees)
  - Missouri Foundation for Health (6 subgrantees)
  - U.S. Soccer Foundation (13 subgrantees)

Exhibit 2 also shows that the evaluators for these reports came from a wide range of backgrounds. Thirteen (50%) reports were authored by external research and evaluation organizations; 11 (42%) by universities or affiliated units, and 2 (8%) by internal staff. Some evaluators were responsible for more than one evaluation report.

3.1.2 Grantee Resources

A major factor influencing the success of the SIF projects is the availability of resources sufficient to handle the program's challenging evaluation demands. These resources include the pre-existing evaluation capacity of grantees and subgrantees, and the amount of evaluation funding relative to the scope. Exhibit 3 presents information about the evaluation funding associated with the grants and subgrants reviewed. As we interpret the findings, it is important to keep in mind the varying level of resources associated with the grants, subgrants, and evaluations. A recent report (Lester, 2015)

categorized SIF grantees and subgrantees into three broad categories:

- **Better-funded grantees**. A few organizations—such as the EMCF, the Mayor's Fund to Advance New York City, and Jobs for the Future—received the largest grants, with total federal funding of$12-30 million and substantial evaluation budgets;

- **Grantees with unified portfolios**. Nonprofits such as REDF and the U.S. Soccer Foundation, although they received smaller grants, achieved efficiencies by creating portfolios of projects that delivered substantially similar interventions and investing in evaluation;

- **Grantees that funded numerous, smaller, regional subgrants[5]** typically received SIF grants at or near $1 million per year and provided subgrants at or near the $100,000 annual minimum. This funding structure produced a relatively large number of subgrantees with relatively low funding and therefore low evaluation budgets. None of these subgrantees has yet achieved a moderate or strong level of evidence.

Of the reports evaluated for this synthesis, 6 were prepared by better-funded grantees; 3 by grantees with unified portfolios, and 17 by grantees that funded numerous, smaller, regional subgrants.

A CNCS report (Zandniapour and Vicinanza, 2013) pointed out that many SIF grantees and subgrantees faced the challenge of budgeting accurately for evaluation. Exhibit 3 presents information about program and evaluation budgets for 70 interventions supported by the SIF.

---

[5] Lester used the term "Small, Regional Intermediaries."

Exhibit 3. *Annual Budget for the SIF Grantee/Subgrantee Program and Evaluation (Years)*

	Minimum	Maximum	Mean	Median
Program Budget	$100,000	$5,460,618	$1,104,649	$593,309
Evaluation Budget	$12,000	$1,346,342	$216,838	$81,471
Evaluation-to-Program Budget Ratio	12%	25%	20%	14%

Source: Zandniapour and Vicinanza (2013)

The same report showed that stronger levels of evidence come at a cost. Among the SIF evaluations, the average cost of randomized controlled trials (RCTs) per year is almost four times that of quasi-experimental (QEDs) or non-experimental designs.

3.1.3 Target Populations and Outcome Areas

Of the 26 programs for which evaluation reports were available, 11 (42%) were in the SIF priority area of healthy futures, 10 (38%) in youth development, and 5 (20%) in economic opportunity. The target populations served and the outcomes the programs sought to achieve included the following:

- **Healthy futures**. All 11 programs served low-income and uninsured populations with 8 targeting rural areas, 2 targeting urban areas, and 2 targeting generally underserved communities. These programs aimed to improve various outcomes such as health care utilization, health behavior, satisfaction with health care, self-efficacy, and health outcomes.

- **Youth development.** All 10 programs served needy and at-risk populations, defined by low income and low performance. Four of the programs were in early childhood education, 5 in K-12 education, and 1 in higher education. These interventions were designed to improve outcomes, including academic achievement, engagement, school readiness, school attainment (e.g., course enrollment, graduation, and retention), job readiness, life skills, and employment.

- **Economic opportunity**. All 5 programs were designed to serve economically

disadvantaged populations, including individuals or families that were unemployed, hard to employ, low-income, high school drop-outs, parolees, homeless or on housing assistance, or contending with health issues. These programs targeted multiple outcomes, including job readiness, job skills, certifications, employment, job retention, earnings, and life stability outcomes such as stable housing, reduced criminal recidivism, and improved health.

3.1.4 Programmatic Strategies

The evaluation reports describe the programmatic strategies employed by grantees and subgrantees to tackle the target outcomes. These strategies, summarized below, varied by SIF priority area. The appendix offers additional detail about programmatic strategies .

- **Healthy futures**. Programs used various approaches to increase access to and quality of health care, including patient navigation systems, mobile or telemedicine, community-based care, afterschool programs, and mobilizing support from a variety of stakeholders (e.g., volunteer doctors, faculty members, or community care workers).

- **Youth development**. Programs employed a variety of models, including summer programs, dual-enrollment, one-on-one tutoring, inquiry-based curriculum, professional development, comprehensive school reform, residential programs, training for parents, and caregiving on site or through home visits.

- **Economic opportunity**. Programs involved transitional jobs programs, workforce

partnership, training/financial incentives/community support, social enterprises, and work experience and skill training.

3.1.5 Key Components of Interventions

Programmatic strategies are put into practice through concrete interventions, and those interventions usually consist of several key components. The necessity for multiple key components is consistent with the assumption that the complex and layered problems tackled by SIF grantees require complex and multi-layered solutions.

One of the challenges evaluators face is the question of how to account for participants who do not receive all of the key components of a planned intervention. For example, some patients in a SIF-funded health initiative may be exposed to only one service, or may leave the program before services are completed. This not only reduces the intervention's effectiveness for that individual, but also complicates efforts to evaluate the effectiveness of the entire program.

Because key components are intervention-specific, it becomes difficult to summarize them at the level of the SIF program or even the priority area. Below are examples illustrating how grantees combine multi-layered solutions within a priority area.

- **Healthy futures**. Mobile Health Services for Rural Kentucky has three components: 1) implementing a mobile service to provide free health screenings and education to the target population; 2) providing fee-based cardiac testing for individuals found to be at-risk during screenings; 3) referring individuals screened and tested who need additional care to physicians for follow-up care.

- **Youth development**. UC Degrees Gen-1 Project and Higher Education Mentoring Initiative provides a variety of services, including social-emotional support (highly structured/supportive housing environment with curfews and mandatory house meetings), student academic support (house classes, academic conferences, studying/tutoring, time management guidance, and "intrusive" advising), community building (ceremonies, awards, events, celebrations, and retreats), and career development.

- **Economic opportunity**. Jobs Plus is an employment program designed to improve the economic well-being of public housing beneficiaries through employment and training services (job placements, job search training, and GED/ESL programs) as well as financial incentives (rent-based incentives, e.g., Earned Income Disallowance program), and community support (e.g., employing a subset of residents as community coaches to raise program awareness and buy-in).

3.1.6 Number of Individuals Served

The number of individuals served by the SIF projects ranged widely, from 76 to 5,382 with a mean of 1,322 and a median of 1,143.[6] Several caveats are worth noting in interpreting these numbers:

---

[6] This computation excludes an outlier from Social Innovation for Missouri that was designed to affect policy changes in seven counties, potentially reaching a population of 536,737.

- The scope of funding among grants and subgrants varied considerably. The SIF awards at the subgrantee level ranged from $100,000 to $2 million;

- The nature and intensity of services differed across projects. For example, a medical program may serve a large number of patients, while an intensive training program for caregivers may work in-depth with a small number of participants;

- Although most programs targeted a single population, several worked with multiple populations. For example, the Resilient Children Project combined work with both teachers (280) and children (2,000) to improve school readiness by integrating education and health services; and

- Because the SIF is designed to build an evidence base and to support future scale-up, serving many individuals may be less important in the short term than gathering solid evidence about the program's implementation and outcomes/impact.

3.2 Methods to Evaluate the Programs

As one of the federal government's tiered-evidence initiatives, the SIF provides resources and support for developing and implementing rigorous evaluations with the goal of building the evidence base for replication and scale up. Although the SIF has multiple goals, including innovation, evidence, and scale, Lester (2015, p. 9-10) argues that "SIF's primary impact comes not from its ability to scale effective programs within the confines of a very small federal program housed in CNCS, but its ability to build the evidence base and to show how to do so effectively, using very specific strategies rooted in venture philanthropy. Evidence trumps scale."

SIF grantees conducted a combination of implementation, impact, and outcome evaluations. Impact evaluations tend to use more rigorous designs such as RCTs or QEDs

that involve counterfactuals that allow for attribution. In contrast, outcome evaluations tend to use interrupted time series designs (including pre/posttest designs) or posttest only designs that cannot demonstrate a causal relationship between program participation and the observed changes in outcomes. This report groups impact and outcome evaluations together. Although grantees and subgrantees in later SIF cohorts were required to provide impact evaluations, those in the 2010 cohort, which represents 88% of the evaluation reports synthesized for this report, were not.

Exhibit 4 shows the implementation and impact/outcome focus of the reports examined in this synthesis. Note the aggregate numbers of reports that addressed each type of evaluation: 22 of the 26 reports (85%) provided evidence about program implementation, while 21 (81%) provided evidence about program impact or outcomes.

Exhibit 4. *Focus of the Evaluation Reports*

Report focus	Total (N=26)		Aggregate Implementation/ Outcome Reports Combined with Single Focus Reports	
	n	%	n	%
Joint Focus: Implementation and impact/outcome	17	65%		
Single Focus: Implementation	5	19%	22	85%
Single Focus: Impact/outcome	4	15%	21	81%

In this section, we discuss methods that grantees and subgrantees used to evaluate the implementation and impact/outcome of the funded programs, examining a number of important (but by no means exhaustive) topics.

3.2.1 Implementation Evaluation Focus and Data Sources

Implementation evaluation assesses how well a program realizes what it sets out to put into practice. Rather than concentrating on outcomes
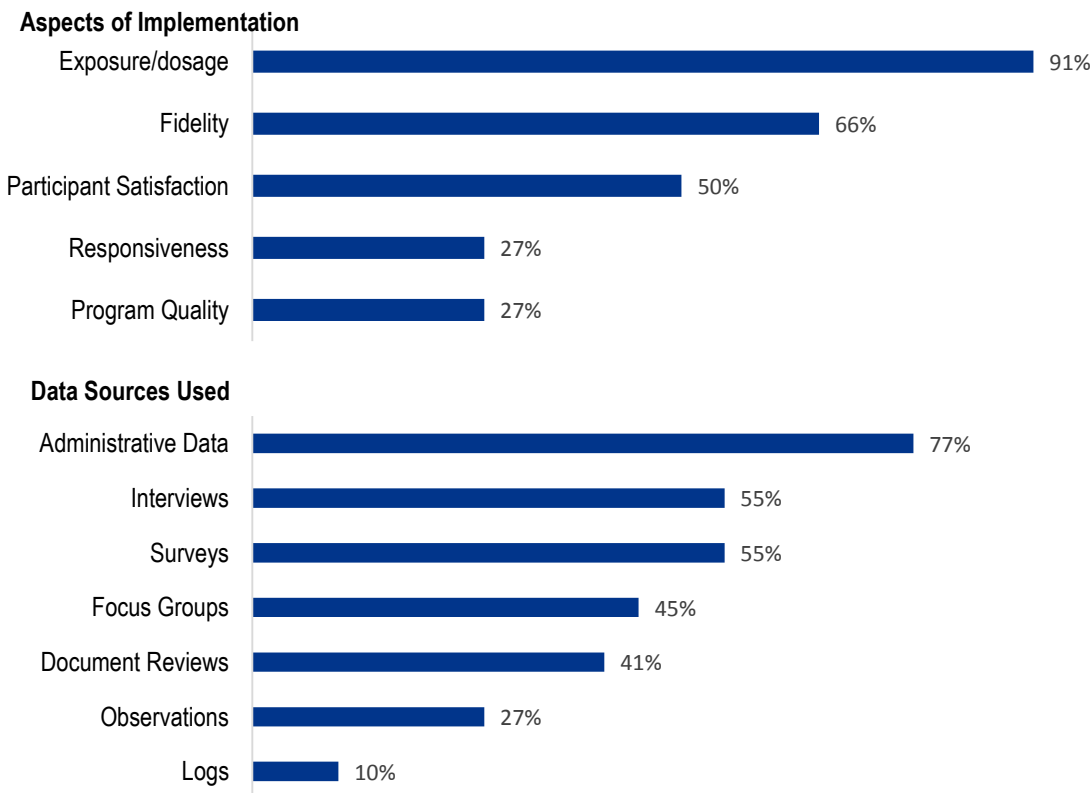
or impact, implementation evaluations examine the process by which a program provides services. Implementation studies are essential for any impact/outcome study. Implementation data are needed to explore how the intervention achieved (or did not achieve) its impact, to support or challenge claims of causality, to examine the relationship between implementation and intervention effects, and to identify challenges to future implementation and scale-up.

As noted in Exhibit 5, 22 reports provided evidence about program implementation (either as stand-alone implementation reports, or as joint implementation and impact/outcome reports). Exhibit 5 summarizes the aspects of implementation, as classified by Dane and Schneider (1998), that were examined in these reports. The analysis suggests that most reports addressed the exposure and fidelity aspects of implementation; half addressed participant satisfaction; only a minority addressed program quality or responsiveness.

In terms of data sources, at least half the reports relied on administrative data, surveys, and interviews. It is not surprising that observations and logs were used less frequently, because they require more resources and incur more response burden than other data collection alternatives.

Exhibit 5. *Implementation Evaluation Approaches (N=22)*

Aspects of Implementation

Aspect	Percentage
Exposure/dosage	91%
Fidelity	66%
Participant Satisfaction	50%
Responsiveness	27%
Program Quality	27%

Data Sources Used

Data Source	Percentage
Administrative Data	77%
Interviews	55%
Surveys	55%
Focus Groups	45%
Document Reviews	41%
Observations	27%
Logs	10%

Note: Percentages may not add up to 100 because an evaluation can include multiple aspects and data sources.

The following examples illustrate how grantees and subgrantees combined multiple aspects of implementation and multiple data sources into a single implementation evaluation.

- **Healthy futures**. The evaluation of the Open Arms Children's Health program focused on exposure and participant satisfaction, with data coming from administrative records, surveys, and focus groups.

- **Youth development**. Pre-School-U's evaluation focused on exposure, fidelity, participant satisfaction, and program quality. Data sources included administrative data, interviews, observations, and logs.

- **Economic opportunity.** The Transitional Employment Services Model measured exposure, fidelity, and program quality, and drew information from administrative data and focus groups.

3.2.2 Impact/Outcome Evaluation

The SIF program has devoted considerable effort to improving the quality and rigor of its impact evaluation. This section examines methodological aspects of the impact/outcome evaluations, including research design, sample and power, outcome measures, and analytic approaches to causal inferences.

Research design
Impact/outcome evaluation provides statistical evidence of how well an intervention achieves its purpose and what effect it has on participants. The quality of an impact study hinges on its research design, particularly on whether a comparison group is created through either random or nonrandom allocation. However, developing rigorous evaluations that lead to credible evidence of what works can be both costly and difficult. A CNCS report pointed out that, "Too often, effective nonprofits lack the expertise, resources, or infrastructure to evaluate their efforts, demonstrate impact, and take their programs from 'promising' to 'proven'" (Zandniapour and Vicinanza, 2013). For example, the 2013 study found that the SIF-funded RCTs cost an average of $437,110 per year—more than the entire annual subgrant for many SIF subgrantees. Projects that used community-based interventions, common in the healthy futures priority area, faced especially great challenges in aspects such as participant recruitment/retention, data collection, and making attributions, because participants may

only receive one service rather than exposure to a more thorough or longer-term intervention.

As shown in Exhibit 6, of the 21 reports focusing on impacts/outcomes, 38% employed more rigorous designs such as RCT or QED design with matching. Several additional reports indicated that the evaluation teams had originally proposed more rigorous designs, but could not implement them because of insufficient or reduced funding, difficulty collecting the outcome data, or problems in participant recruitment or retention—all common challenges in evaluation.

Exhibit 6. *Impact/Outcome Evaluation Design (N=21)*

Less Rigorous design (62%)   More rigorous design (38%)
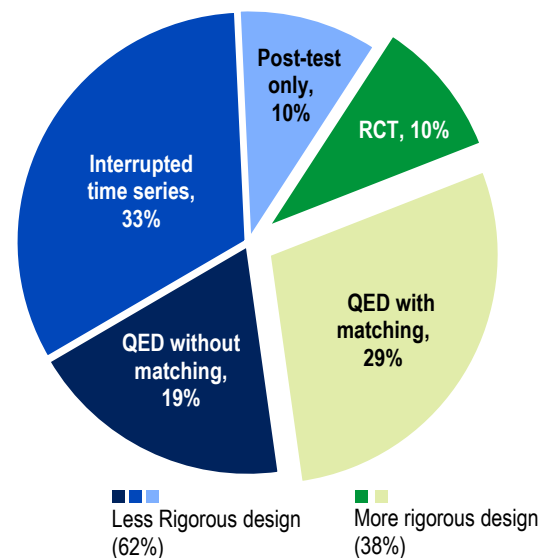
Note: Percentages may not add up due to rounding.

The SIF program began requiring grantees and subgrantees to design and implement impact evaluations that meet moderate or strong levels of evidence with the 2011 cohort (see section 1). Of the reports included in this synthesis, the following programs used the more rigorous RCT or QED with matching research designs.

- **RCT.** Bell Middle School Model and Reading Partners employed student-level RCT.

- **QEDs with matching** (most using propensity score matching): UC Degrees Gen-1 Project, Bridging the Gap, Resilient

Children Project, Workforce Partnership Programs, Social Enterprises, and Soccer for Success.

Sample and power
One of the key considerations in impact evaluation is statistical power. A power analysis is a calculation that estimates, given a specific sample size and analysis approach, how likely it is that a program effect will be significant. Adequate power is essential for statistical validity because insufficient power may lead to an incorrect conclusion that the intervention does not have a significant impact when it actually does. Although lack of power may not seriously threaten the internal validity of the findings, it will limit their generalizability.

Of 21 reports focusing on program impact/outcomes, the sample sizes vary considerably. Even within one study, the samples may vary depending on the outcome measures. Seven of the 21 impact/outcome evaluations appeared to be adequately powered in all or at least one outcome, while the remainder tended to be under-powered. Below are a few examples of studies demonstrating adequate power.

- **Individual-level RCT.** Reading Partners: The final sample for this RCT study included 579 treatment students and 567 control students. The unit of assignment is individual students.

- **Individual-level QED.** Workforce Partnership Programs: The final samples for this QED study using propensity score matching are separated into three sectors: health care (992 treatment, 46,701 comparison); advanced manufacturing (682 treatment, 42,293 comparison); and construction (379 treatment, 36,859 comparison). The propensity score matching allows multiple comparison subjects to be matched to one treatment subject.

- **Cluster-level QED.** Soccer for Success: For this QED study matched by sites, there are

712 treatment students from 16 sites and 522 comparison students from 14 sites. The unit of assignment is site.

Outcome measures
Ensuring that the chosen outcome measures are reliable, valid, and appropriate for a study is key to reducing threats to internal validity. The outcome measures in the SIF impact evaluations were closely aligned with the outcomes targeted by the interventions, and many of the studies employed valid and reliable measures to assess implementation and impact/outcomes.

Some of the outcome measures may be characterized as short-term or intermediate, others, as long-term. Many evaluations used a logic model to distinguish among these outcomes and conceptualize the relationships among them. Some used scales/instruments that are well-established in their fields, which helps address validity issues and allows comparison/integration with related research and evaluation literature.

The examples below illustrate the types of outcomes and specific outcome measures examined in impact evaluations in each priority area.

- **Healthy futures.** Healthy Futures impact evaluations tended to measure outcomes such as health outcomes, health behaviors, knowledge, self-efficacy, health care utilization, cost, and patient and stakeholder satisfaction. For example, the Get Healthy Access Program focused on health outcomes (e.g., blood sugar, blood pressure, Body Mass Index, and waist circumference) and health-related behaviors (e.g., smoking, physical activity, and consumption of fruits and vegetables).

- **Youth development**. Youth development outcome measures included student achievement or readiness, and student engagement or academic behavior/attainment measures such as attendance, course enrollment or

completion, retention, and graduation. For example, Parents as Teachers (PAT) and Home Instruction for Parents of Preschool Youngsters (HIPPY) measured school readiness for preschoolers using the Ages and Stages Questionnaire communication subscale and MacArthur Communicative Development Inventories.

- **Economic opportunity**. Economic opportunity impact evaluations examined such outcomes as employment, job retention, earnings, job readiness, skill certifications, and life stability. For example, Social Enterprises employed two types of outcomes: 1) economic self-sufficiency (measured by employment and income) and 2) life stability (measured by housing situation, recidivism, and health outcomes).

Statistical analysis for supporting causal inferences
Appropriate statistical analysis techniques ensure the rigor of evidence from the evaluation. The statistical analyses supporting the causal and correlational findings are often aligned with the study designs.

Descriptive analysis alone, while essential and informative, cannot establish causal or correlational inferences. Even among inferential statistics, some are more appropriate and rigorous than others, given the circumstances. For example, in an RCT, where the unit of assignment is cluster, a multi-level modeling approach is more appropriate than a single-level analysis, because it corrects for the cluster effect. In a group comparison, a multiple regression sometimes is a better approach than t-test or ANOVA, as it accounts for the other confounding variables.

Of the 21 projects with impact/outcome focus, 81% used inferential statistics such as multi-level modeling, t-test, ANOVA, or correlational analysis to support causal claims (Exhibit 7).

Exhibit 7. *Analysis Approach to Causal Inferences (N=21)*

Inferential Statistics (81%)   Descriptive Analysis (19%)
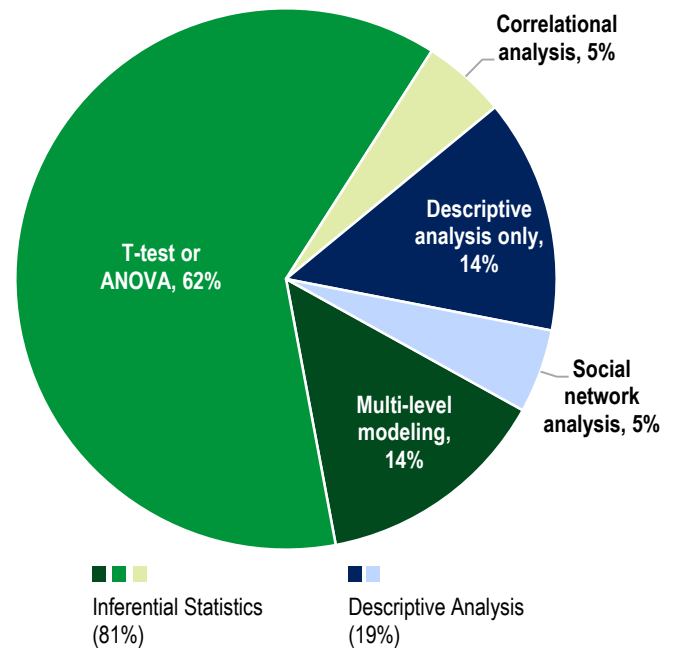
3.2.3   Level of Evidence

In order to support evaluation design and subsequent technical assistance, CNCS developed criteria to distinguish three tiers of evidence: preliminary, moderate, and strong. Although other aspects of the evaluation—such as measures, sample size and retention, and data analysis—contribute to the rating, the level of evidence is primarily determined by evaluation designs, as discussed in section 1, as well as whether their manner of implementation addresses threats to internal and external validity.

We can consider the level of evidence as an overall rating of the evaluation methods examined in this section. Exhibit 8 shows that, of the 26 evaluation reports, 2 were categorized as strong (Reading Partners and Soccer for Success), and 3 were rated as moderate (Bell Middle School Model, Workforce Partnership Program, and Social Enterprises). The remaining 21 reports (including 5 with implementation-only results) were classified as preliminary.

Since the 2010 SIF cohort, oversight of grantee and subgrantee evaluations has evolved substantially. In addition to articulating the expectation for achieving moderate and strong evidence, CNCS has become more involved in approving evaluation plans, overseeing subgrantee competitions, and increasing the evaluation capacities of grantees and subgrantees through technical assistance and knowledge sharing. It is anticipated that a higher percentage of the evaluations completed by later cohorts will achieve the strong or moderate levels of evidence.

Exhibit 8. *Level of Evidence (N=26)*
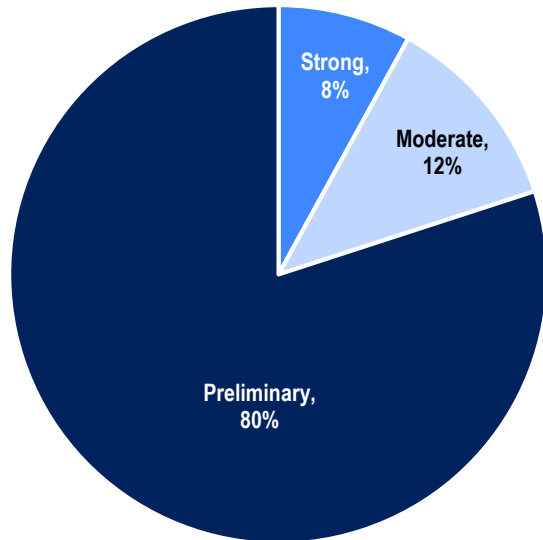
3.3 Summary of Evaluation Findings

Evidence, especially from evaluations that yield a strong or moderate level of evidence, is likely to become one of the SIF's enduring legacies, because sound evidence will inform future innovation and scale-up by other nonprofits across the nation. This section briefly summarizes the substantive findings from these 26 reports, as well as information about implementation, impact/outcome, and cost. It also summarizes the assessments made by these reports of limitations, challenges, lessons learned, and contributions to the field.

3.3.1 Implementation Findings

The evaluation findings show that most SIF-funded programs were implemented appropriately with respect to the key aspects of implementation summarized in Exhibit 5. Of the 22 reports focused on implementation, most (86%) reported evidence of a high or adequate degree of implementation in terms of exposure/dosage, fidelity, and/or participant satisfaction. Below are three examples of projects with findings about implementation.

- **Healthy futures**. Implementation of the St. Elizabeth Telepsychiatry program met the implementation goals, with 146 staff trained and reporting that they felt competent to provide services; 1,120 eligible patients who consented to receiving services; and 100% of patients in telepsychiatry assessment who received a plan of care, a referral for outside services, and/or follow-up care.

- **Youth development**. For Gateway to College, students consistently expressed a strong appreciation for the program and valued the mature learning environment it offered. The program succeeded in maintaining a model that has a shared culture of student-centered support and a solution-focused pedagogy. The sites implemented the core model as designed. However, flexibility in implementation allowed local programs to be responsive to student needs and build strong relationships with postsecondary host institutions, it created challenges in maintaining fidelity.

- **Economic opportunity**. The report for the Transitional Employment Services Model found that, overall, the program was implemented with fidelity. Screening was important for creating a good fit between programs and participants. Case management was the most important factor for program success. The majority of participants completed the transitional employment requirement. Program participants who viewed lack of industry-

embedded training as a barrier prior to the program, no longer held that view following the program.

Some projects were not implemented according to plan. Below are examples of the difficulties encountered.

- **Participant recruitment.** For the Pre-school-U program, none of the sites reached the targeted recruitment and retention goals. The program did not seem to reach its target population either, as most participants were already competent in their parenting skills and quickly lost interest after discovering that the program was not equipped to help them.

- **Data collection.** For St. Joseph Health System, the evaluator concluded that the data were insufficient to determine implementation fidelity, because one site's data had not been collected at the time of the report.

3.3.2   Impact/Outcome Findings

Evaluation findings show that most SIF projects led to positive effects. Of the 21 reports concentrating on impact and/or outcomes, 91% found some positive effects. Exhibit 9 further reveals that 10 (48%) found positive effects on all outcomes, 9 (43%) produced mixed but mostly positive results, and 2 did not find any effects. None of the reports found negative effects.

The subset of evaluations with a strong or moderate level of evidence followed a similar pattern. Of the 5 evaluations rated as providing a strong or moderate level of evidence, 2 found positive effects on all outcomes; 2 yielded positive results on some but not other outcomes; and 1 concluded with no effect.

Exhibit 9. Impact/Outcome Findings

Impact/outcome findings	All evaluations (N=21)		Evaluations with strong or moderate evidence (N=5)	
	n	%	n	%
Some Positive Effect	**19**	**91%**	**4**	**80%**
Positive effect on all outcomes	10	48%	2	40%
Positive effect on some outcomes but not others	9	43%	2	40%
No Effect	**2**	**10%**	**1**	**20%**

Note: Totals do not total 100% due to rounding.

Listed below are illustrative examples of impact/outcome findings for three evaluations that achieved a preliminary level of evidence.

- **Positive effect on all outcomes**. The impact report from Early Childhood Science Inquiry Training for Educators (youth development) found the following: 1) treatment students scored higher than comparison students with no documented preschool attendance; 2) treatment teachers reported increased comfort and competency in the preparation of science and math lessons; and 3) parents reported greater comfort in having conversations with children about science, as well as increased awareness and ability to support science learning.

- **Positive effect on some outcomes**: Evaluation findings for the Meade Activity Center (healthy futures) suggested 1) statistically significant and positive impact on physical activity self-efficacy, physical activity behavior, sedentary behaviors, cardiorespiratory fitness level, and body mass index; 2) no impact on healthy eating behaviors.

- **No effect**. The impact evaluation of the Community Based Delivery Model/Virtual Care program (healthy futures) revealed no apparent changes in utilization of primary and specialty care, and no perceived improvements in health or self-efficacy.

When interpreting impact/outcome findings, it is important to note the associated level of evidence. For example, outcome findings from a report at the preliminary level of evidence should be interpreted as descriptive rather than causal. Exhibit 10 summarizes findings from the 5 impact evaluations that had a strong or moderate level of evidence and showed strong or moderate evidence of a causal relationship.

Exhibit 10. *Summaries of Findings from Impact Evaluations with Strong or Moderate Levels of Evidence*

Project name	Evaluation findings
Strong level of evidence	
Reading Partners (Youth development)	1) Positive and statistically significant impact on all three primary impact measures: reading comprehension (d=0.10), reading fluency (d=0.09), and sight word fluency (d=0.11), equivalent of 1.5 to 2 months of growth in a school year; 2) Overall impact did not differ by grade level, English language learner status, gender, or implementation fidelity; 3) Little or no statistical differences observed in attentiveness, assignment/homework completion, attendance, and teacher-reported academic performance in reading or math; and 4) Treatment students were more likely to move out of the lowest quartile of the national reading scores than those in the control group.
Soccer for Success (Healthy futures)	Positively significant effects were observed in participants' BMI, waist circumference, and aerobic capacity.
Moderate level of evidence	
Bell Middle School Model (Youth development)	The program had no effect on reading (d=0.01) or student engagement (d=0.01) and a potentially positive but not statistically significant impact on math (d=0.07).
Workforce Partnership Programs (Economic opportunity)	All three programs in health care, advanced manufacturing, and construction were effective in helping unemployed participants obtain employment in the entire 12-month period. The health program also effectively promoted participant employment in its focus industry, and the other two programs had a modest effect on this outcome. The manufacturing program led to significant impact on job retention and earnings. In contrast, there was no evidence that the construction program led to higher job retention and earnings.
Social Enterprises (Economic opportunity)	1) Employment increased by 51%; 2) Significant improvement in physical health; 3) No impact on number of arrests and convictions; 4) Support after leaving the social enterprise (SE) is associated with increased self-sufficiency and life stability; 5) Larger and more established SEs might produce value to society and a net benefit to taxpayers in the first year after SE employment begins, although the smaller and newer SEs do not appear to produce such benefits.

3.3.3 Cost Findings

At a time of limited resources and competing priorities, findings about program cost are important for replication and scale-up, particularly for interventions yielding positive impact with strong or moderate level of evidence. Although cost analysis was not required as part of the SIF evaluation, it was included in 4 of the 26 reports (15%). Each of these programs was found to offer low-cost solutions to addressing the target areas:

- **Reading Partners** (youth development). The program is a low-cost option for under-resourced schools because in-kind contributions from community volunteers pay most of the costs. On average, schools bear about 20% ($710 per program student) of the total cost of the resources required, more than half of which are in-kind contributions from the school of space and staff time. The evaluation, supported by a strong level of evidence, found positive impact on all outcomes.

- **Jobs Plus** (economic opportunity). The cost-per-resident-household of operating the program at the targeted housing units in year 3 was $672 in New York and $503 in Texas. Financial resources to support program staff's implementation and engagement among clients and community members was sizable, i.e., over 70% of

funds. The evaluation, supported by a preliminary level of evidence, focused only on implementation.

- **Social Enterprises (SE)** (economic opportunity). The return on investment (ROI) for the SE program was at least as large, if not larger, than estimated ROI from similar programs. For every dollar SE spent, the SE returned $2.23 (outcome study) or $1.34 (impact study) in total benefits. The evaluation, supported by a moderate level of evidence, found positive impact on most but not all outcomes.

- **Community Based Delivery Model/Virtual Care** (healthy futures)**.** For specialty care, patients averaged a savings of $38 per consultation by utilizing telemedicine. The evaluation, supported by a preliminary level of evidence, did not find any impact on health outcomes.

3.3.4   Challenges in Program Implementation and Evaluation

Half the reports discussed challenges encountered during program implementation and evaluation, and in some cases, multiple challenges. Challenges related to program implementation included delays in schedule and activities, difficulties with participant recruitment and retention, concerns about implementation fidelity, and other issues such as transportation, procurement restrictions, staffing, and grant management (e.g., complexity in invoicing). Evaluation-related challenges include issues with data collection resulting from limited funding, difficulty achieving desired response rates, challenges in managing tracking systems, and other unexpected roadblocks, including delays in SEP approval, evaluator turnover, and use of an inappropriate evaluation instrument. Below are examples.

- **Participant retention and recruitment**. Gateway to College's (youth development) biggest challenge was retaining students in

the program during the initial Foundation term. Fewer than half the students passed all their Foundation courses and successfully transitioned to mainstream community college. The biggest stumbling block was English Language Arts. Retention was affected by many pressures and problems that existed prior to enrolling in the program, such as health, family issues, and conflicts with work.

- **Data collection**. Cumberland Family Medical Center (healthy futures) experienced significant challenges with the electronic health record (EHR) system and the process of implementing a research program in a rural health care delivery system. The evaluation team could not obtain the information needed for the evaluation from the EHR—a patient-focused system not designed to support research.

- **Limited funding**. King's Daughter Medical Center (healthy futures) encountered funding limitations, which curtailed data collection and prevented the evaluators from completing the evaluation as originally proposed in the SEP.

3.3.5   Limitations of Reported Evaluations

Half the reports explicitly discussed limitations, most of which concerned evaluation approaches. These limitations included lack of counterfactual, data collection problems with measures, participant recruitment and retention, data quality issues such as missing data and low response rates, lack of statistical power, inability to address important questions, conducting the evaluation at sites implementing the program for the first time, and significant program changes occurring after the evaluation data collection. Some examples are:

- **Lack of counterfactual**. Bridging the Gap (youth development) could not implement propensity score matching as proposed. There was selection bias between treatment

and comparison groups as well as confounding factors in attribution.

- **Participant attrition and missing data**. Early Childhood Science Inquiry Training (youth development) for Educators found that uncertainty in the lives of those in the treatment group resulted in considerable attrition, reduction of sample size, and missing data.

- **Target populations**. Although the Center for Employment Opportunities (economic opportunity) has been shown to be most effective for those within three months of release from incarceration, the sites were not so successful at targeting those recently released or at highest risk. Unlike New York City, with its extensive public transportation system, the limited transit in several communities limited their ability to recruit members of the target population.

- **Program changes and statistical power**. Building Educated Leaders for Life (youth development) evolved significantly after the evaluation data collection, rendering the results less relevant; the study was conducted in school districts implementing the program for the first time, raising questions about fidelity; the study was underpowered from the outset.

3.3.6 Lessons Learned and Recommendations

About half the reports specifically discussed lessons learned and made recommendations. These reflections were often based on perceived successes, setbacks, or ideas for future exploration.

- **Program-related lessons** included the need to:
  - Prioritize efforts to improve working relationships among stakeholders
  - Plan strategies to maximize recruitment and retention of participants

  - Seek to provide a longer duration of intervention, to have sufficient time to achieve a noticeable impact
  - Keep the program small at start-up to allow for rapid adjustments and program improvements and scale up once the initial complications are worked out
  - Consider cost implications explicitly when making program decisions
  - Assess the infrastructure improvements needed for effective implementation and tackle them early in the program
  - Find ways to expand successful programs to different populations, remaining alert to adjustments that may be needed to account for differences (e.g., age differences, cultural differences)

- **Evaluation-related lessons** include:

  - Provide appropriate levels of funding for data collection
  - Obtain staff with the skills to handle large datasets
  - Plan for tracking longitudinal outcomes
  - Select sites familiar with the program to enhance fidelity
  - Collect detailed information about the control condition to allow for appropriate comparisons
  - Obtain methodological guidance early and devise sound data collection methods to ensure that the study is well-powered.

3.3.7 Contributions to the Evidence Base

About half the reports explicitly addressed how it contributed to the evidence base in the field. All the implementation reports discussed issues relating to the implementation, replication, or scale up of a particular intervention in a target area. Consistent with their levels of evidence, impact/outcome evaluations made very concrete contributions, such as the following:

- **Offering a low-cost option**. The Reading Partners (youth development) report

explores the low-cost option of using community volunteers to provide one-on-one tutoring. One-on-one tutoring by certified teachers has been demonstrated to be effective in improving the reading proficiency of struggling readers but it is both time- and resource-intensive. (Strong level of evidence)

- **Providing empirical evidence to a field with little pre-existing research**. The Social Enterprises study (economic opportunity) expands the information available about the effectiveness and costs/benefits of social enterprises, about which little impact research exists. (Moderate level of evidence)

- **Replicating a model in a new context.** The Hope Health Clinic (healthy futures) replicated the model developed by the Church Health Center in Memphis that had operated for 25 years but was never systematically evaluated. The model builds on the literature concerning the benefits of providing low-cost and accessible primary care to uninsured persons, and the use of lay patient navigators, community health workers, non-clinicians, and clinic volunteers to improve patient self-efficacy and access to services. (Preliminary level of evidence)

4. Next Steps

This interim report has examined and synthesized findings from 26 evaluation reports, completed as of June 1, 2015, in terms of programmatic features, evaluation methods, and findings. The SIF has a fully operational evaluation pipeline that will generate more results in the coming years. The follow-up report, scheduled for completion in 2016, will include additional findings from evaluations completed since this review. In addition to an

update of the thematic-based meta-synthesis presented here, the 2016 study is expected to apply meta-analysis techniques to analyze the quantitative findings across studies (Zhang, X. et al., 2013). The technique currently envisioned will include the following features:

- **Inclusion criteria.** The study will include all impact reports using designs with counterfactuals (e.g., RCTs, QEDs, and interrupted time series). It will exclude studies relying on posttest outcome data.

- **Effect Size Calculation.** The study will use Hedges' g, a standardized mean difference between two groups, as the effect size index for this meta-analysis.

- **Unit of Analysis.** Meta-analysis requires that each study represents an outcome domain from an independent sample. We will use a shifting unit of analysis approach (Cooper, 2010) to further address possible dependencies among effect sizes.

- **Moderator Analysis.** The study will include testing for moderators on the groups of effect sizes that have a high degree of heterogeneity (Cooper, Hedges, and Valentine, 2009). These moderators are related to either the intervention (e.g., sector, type of intervention, programmatic strategies) or the study design (e.g., research design, study population, level of evidence).

References

Corporation for National and Community Service. (2013). *Getting results, transforming lives: the Social Innovation Fund, 2010-2012 investment report*. Washington, DC: Corporation for National and Community Service.

Corporation for National and Community Service. (n.d.) *Social Innovation Fund: Evaluation plan guidance-A step-by-step guide to designing a rigorous evaluation*. Washington, DC: Corporation for National and Community Service.

Cooper, H. (2010). *Research synthesis and meta-analysis: A step-by-step approach* (4th ed.). Thousand Oaks, CA: Sage.

Cooper, H., Hedges, L.V., and Valentine, J.C. (Eds.) (2009). *The handbook of research synthesis and meta-analysis* (2nd ed.). New York: Russell Sage Foundation.

Dane, A.V., and Schneider, B.H. (1998). Program integrity in primary and early secondary prevention: Are implementation effects out of control? *Clinical Psychology Review*, 18(1): 23-45.

Lester, P. (2015). *Social innovation fund: early results are promising*. Social Innovation Research Center.

Zandniapour, L., and Vicinanza, N. (2013) *Budgeting for Rigorous Evaluation: Insights from the Social Innovation Fund*. Washington, DC: Corporation for National and Community Service, Office of Research and Evaluation.

Zhang, X., Tran, H., Chen, E., Steketee, M., Hershey-Arista, M., Katterfeld, K., Ristow, L., and Frechtling, J. (2013). *Quality Review and Meta-analysis of Impact Studies for the National Science Foundation's Discovery Research K-12 (DR K-12) Program*. (Prepared under contract to the National Science Foundation). Rockville, MD: Westat.

Appendix: Programmatic Strategies

Program name	Programmatic strategies
Youth Development	
Bell Middle School Model	**Summer program** for middle school students to improve student academic achievement and engagement
Gateway to College	**Dual-enrollment program** for at-risk students to improve high school graduation and college readiness
Reading Partners	**One-on-one reading tutoring** to struggling readers by community volunteers
Early Childhood Science Inquiry Training for Educators (ECSITE)	**Pre-school curriculum and professional development**, which integrate science experiences across other curriculum domains through enriched environment, planned lessons, and incidental conversations with children.
Holmes 180	**Comprehensive high school reform** designed to connect high school students with meaningful career opportunities through instructional and policy changes.
UC Degrees Gen-1 Project and Higher Education Mentoring Initiative	**Residential program** providing low-income first generation college students and foster youth with a variety of academic, supportive, and other services
Parents as Teachers (PAT), and Home Instruction for Parents of Preschool Youngsters (HIPPY)	**Home visiting programs** designed to increase the frequency with which parents engage in age-appropriate language and other activities with their children
Pre-School-U	**Program to train parents and caregivers** with knowledge and skills to best prepare their children for school, implemented in faith-based settings
Bridging the Gap	**Program providing work experience and professional development in job and life skills** to reduce dropout and improve adult employment for high school students through partnership with employers and community service providers
Resilient Children Project	**School readiness program** integrating educational and mental health services through early childhood mental health consultation services designed to increase teacher and caregiver competence
Economic Opportunity	
Center for Employment Opportunities (CEO)	**Transitional jobs program** for parolees
Workforce Partnership Programs	**Program establishing local workforce partnerships** to help low-income workers to obtain skills in needed areas
Jobs-Plus	**Employment and training, financial incentives, and community support** to public housing beneficiaries
Social Enterprises	**Social enterprises**--businesses sell goods and services that the market demands in order to intentionally employ individuals who would otherwise have difficulty obtaining employment: providing employment, employment support while employed, life stability supports while employed, and post-employment support
Transitional Employment Services Model	**Employment and training** to allow hard to employ individuals to simultaneously access vocational skills training, work readiness training, and job seeking assistance.
Healthy Futures	
Cumberland Family Medical Center	**Nurse-managed community health center and a rural community-based training program** providing affordable and accessible health care
Open Arms Children's Health program	**Partnership between a comprehensive child-caring agency and a research university to offer on-site dental service and expanded health care access**
Mobile Health Services for Rural Kentucky (MHSRK) program	Increasing access to health care by providing **free screenings and affordable cardiac testing,** increasing knowledge of risk factors for heart disease, reducing unhealthy behaviors, and improving health outcomes among low-income, uninsured and underserved rural residents
Meade Activity Center	**Afterschool program** to promote physical activities for children in a center
Community Health Worker (CHW) program	**Employing community health worker** to provide health care to reach a wider community

Program name	Programmatic strategies
Get Healthy Access Program (GHAP)	**Program offers cardiovascular screenings**, including blood pressure, blood sugar, cholesterol, and BMI screenings and, for those eligible, enrollment in the GHAP program that offer health education and navigation to primary care providers.
Hope Health Clinic	**Embedded patient navigation system** offers low-income, uninsured clients, families and caregivers assistance to help overcome health care system barriers and facilitate timely access to services.
St. Elizabeth Telepsychiatry program	**Remote access to psychiatric assessments via telemedicine**
Community Based Delivery Model: Virtual Care	**Primary and specialty care through tele-health** to poor and uninsured populations
Social Innovation for Missouri	**Community Health Improvement model emphasizes collaboration across multiple sectors and among diverse stakeholders** to achieve a strong community-wide initiative via policy, environmental, and community changes to address tobacco use and obesity.
Soccer for Success	**Afterschool program** provides children in underserved communities with access to quality physical activity programs and nutrition education to support their physical and personal development, improve fitness, and reduce obesity