Corporation for NATIONAL & COMMUNITY SERVICE

Levels of Evidence Tip Sheet

Introduction

The purpose of the Social Innovation Fund (SIF) is to grow the impact of innovative community-based solutions that have compelling evidence of improving the lives of people in low-income communities throughout the United States. The program models funded by SIF grantees must produce rigorous evaluative evidence that not only assesses their effectiveness, but also builds the evidence base for other programs addressing similar issues. This tip sheet aims to help grantees and subgrantees consider the issues that come up when developing a study design that meets the standards of rigor for SIF evaluations.

When designing your study, remember that:

- Evaluation designs that minimize threats to internal validity and maximize external validity are most likely to yield strong evidence;
- Designs that address key threats to internal validity, although they may not have strong external validity, can yield moderate evidence; and
- Studies using designs that will only qualify as moderate evidence (such as interrupted time series) must carefully eliminate all threats to internal validity.

Levels of Evidence

In determining the level of evidence your evaluation is likely to generate, you should ask yourself the following questions:

What level of evidence can I optimally achieve given the timeline, budget, staff, established infrastructure, and agreed-upon involvement from potential participants?

It is always important for the study design to maximize internal and external validity and to strive to attain the highest level of evidence possible. Consider the type and amount of previous research conducted on your program or similar programs; your program and target population characteristics; and how you might generate a counterfactual to allow for a causal assessment of program impact.

What program characteristics lend themselves to particular research designs?

Think about how the evaluation can be structured to play to your program's strength, while achieving the highest level of evidence. For example, a program with demand that exceeds capacity could use a randomized controlled trial that allows control group members to continue to apply (thus, still allowing everyone the opportunity to receive services).

Levels of Evidence Tip Sheet, continued

Is it possible to take steps to preemptively address problems that may crop up with internal validity, self-selection into the program, attrition, or external factors (i.e., community-level changes)?

Consider how aspects of the evaluation design can be used to mitigate threats to internal validity. The use of a counterfactual (e.g., a control group, a comparison group, or an extended time series in which data are collected for a long period from a diverse and representative population) works to minimize threats to internal validity. Additionally, having mechanisms that ensure data collection takes place for all study participants (both in and out of the program) can also help with threats to validity.

How might the evaluation process and evaluator decisions (such as choice of survey items) affect my results?

Using validated and reliable data collection instruments helps ensure greater internal validity and strengthens external validity. Also, consider the ways in which the evaluation process itself could potentially affect the observed outcomes. Be mindful, for example, of how evaluation activities might impact program operations, how data collection methods could affect responses (i.e., interviewer effects), and how uneven data collection between program participants and nonparticipants can be avoided.