



**STANFORD**  
GRADUATE SCHOOL OF BUSINESS

**STANFORD** SOCIAL INNOVATION *review*

# **First Person**

## **But Does It Work?**

By Jed Emerson

Stanford Social Innovation Review  
Winter 2009

Copyright © 2009 by Leland Stanford Jr. University  
All Rights Reserved



**STANFORD**  
GRADUATE SCHOOL OF BUSINESS

**Stanford Social Innovation Review**  
518 Memorial Way, Stanford, CA 94305-5015  
Ph: 650-725-5399. Fax: 650-723-0516  
Email: [info@ssireview.com](mailto:info@ssireview.com), [www.ssireview.com](http://www.ssireview.com)

## But Does It Work?

How best to assess program performance BY JED EMERSON

MORE THAN 20 YEARS AGO, when I was running a community organization that provided critical services to homeless youth, our foundation funders never asked for data on outcomes. Our public funders, meanwhile, felt satisfied with a reporting system that largely consisted of my caseworkers and me running up and down the hallway, paper and pencil in hand, polling each other and pulling files from various drawers to report on the number of youths we served that month. No one discussed the difference between how many kids you served and how many kids' lives you actually changed.

Things are different today. Funders and nonprofit managers from all fields have spent years involved in deep, meaningful debates about how best to document impact—how to create sound program evaluations and dynamic management information systems capable of tracking impact over time. These discussions have resulted in the development of effective reporting systems that managers can use to track how, if at all, their efforts actually reach the goals that are set.

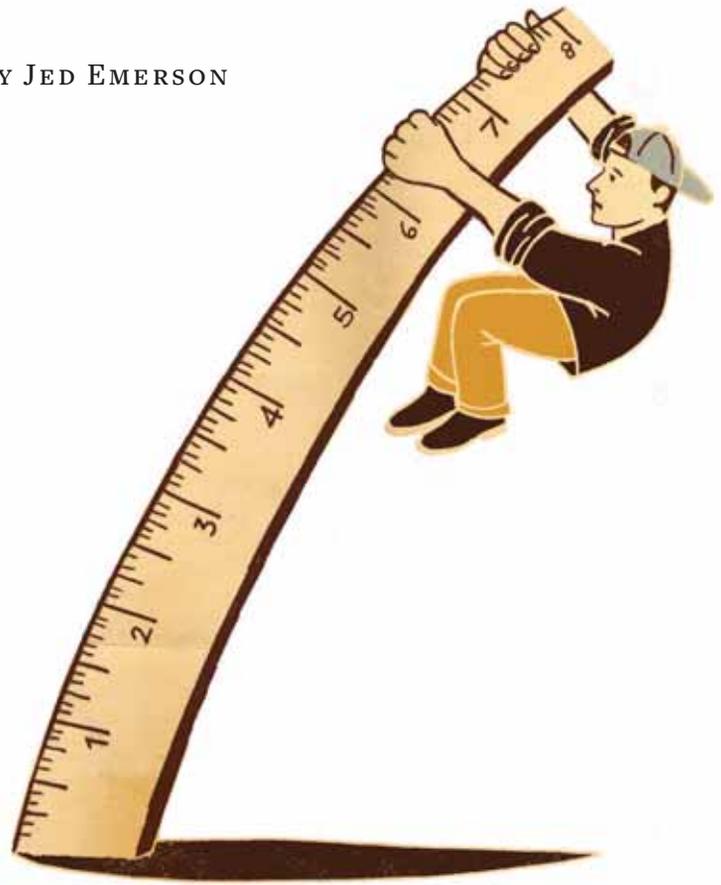
But there are times when the “best practice” within a given area will still fall short of what is considered scientifically valid evidence of effectiveness. What’s more, despite discussion of outcome funding and the increasing use of evidence-based approaches to the allocation of dollars, many funds today are not allocated strictly for impact, but rather for an attempt at “doing good.” You can see this in the annual rankings of donors and foundations: The winners were chosen not for value creation, but rather for institutional asset and grant size.

All of which brings us to the core question of evaluation: How do funders and nonprofit managers best assess program performance so that they can fund interventions that truly work?

At the Edna McConnell Clark Foundation (EMCF), which funds programs with documented evidence of positive impact on the lives of young people, we answer this question by understanding that such programs fall into one of three stages of effectiveness: “apparent,” “demonstrated,” and “proven.” The higher the level, the more confidence we have that the program is indeed effective and of possible interest to us.

### APPARENT EFFECTIVENESS

Organizations operating at the first level of measuring impact—apparent effectiveness—can assume that young people are benefiting as intended from participation in a specific program, given systematically collected data. The organizations’ performance tracking systems can answer such basic questions as who is accessing services, what programs do they participate in, and what outcomes have they achieved?



Specifically, every program participant is first given a unique identifier (such as a tracking or identification number). The organization then collects basic demographic data from program participants (age, gender, etc.). This information becomes the baseline for measuring changes over time (outcomes), and the organization will spell out and track these outcomes for all program participants. As an example of this last stage, EMCF grantee Green Dot Public

**JED EMERSON** is managing director for integrated performance at Uhuru Capital Management, a new global investment firm that will fund the expansion of programs supporting social entrepreneurs around the world. At the time of this writing, he was project manager for evaluation and performance at the Edna McConnell Clark Foundation.

Schools, a public charter school serving youth in some of the poorest neighborhoods in the Los Angeles area, tracks students’ attendance, how they perform in school, and whether or not they attend college or find a job after graduation.

Even though the number of organizations using this apparent effectiveness category is growing as more nonprofit managers try to understand the impact of their work, the vast majority of nonprofits active today still would not meet this most basic level of documenting performance.

**DEMONSTRATED EFFECTIVENESS**

Organizations at the second level of measuring impact—demonstrated effectiveness—systematically collect data that compare program participants with a similar population not receiving the same services. They can thereby substantially conclude that the youths involved are benefiting from program participation, and they can also positively answer the vital question of whether participants experience better outcomes than comparable people not in the program.

An example of such an organization is EMCF-funded Hillside Work-Scholarship Connection (HW-SC). It recently found that students in the program graduated from high school at twice the rate of their peers, and more than 75 percent of these graduates go on to postsecondary education, compared to 33 percent of the local popu-

An advocacy organization may not have the means to demonstrate from a scientific perspective that their strategy for advancing public policy is “proven.”

lation. HW-SC partners with private employers like Wegmans Food Markets so that young people develop marketable job skills and employers reduce training and turnover costs.

At this level, an independent, external evaluator carefully reviews the program and services, target population, desired outcomes, and indicators measuring success. The evaluator then compares outcome data to information and data for a carefully chosen comparison group whose members are not program participants but may be subgroups of that same population.

**PROVEN EFFECTIVENESS**

At the third and highest level—proven effectiveness (EMCF’s “gold standard”)—a program’s impact on participants has been scientifically confirmed through experimental research such as a randomized controlled trial, as is essential in medical research. Organizations here have shown a statistically significant difference in outcomes for their program participants vs. people in a randomized control group. EMCF grantee the Nurse-Family Partnership is a good example. Through years of random assignment evaluations, researchers have found that mothers in the program are healthier, don’t abuse and neglect their children as often, complete their degrees and obtain jobs, and have children who are healthier and who avoid criminal activity later in life.

To gauge proven effectiveness, the same kind of independent, external evaluator who would help establish demonstrated effectiveness randomly assigns participants in an outcomes evaluation either to a group that receives services, or to one that does not. The evaluator then collects and compares outcome data for both groups, with a sample sufficiently large to conclude statistically that the program is responsible for the difference in outcomes achieved by its participants versus those of the control group. (“Sufficiently large” depends on the population, context, and strategy.)

To help young people succeed, many organizations also implement program models already tested and proven by others (with

randomized controlled trials). Local affiliates of the Boys & Girls Clubs of America, for instance, are currently implementing Project: LEARN, an educational program proven to boost educational attainment. The key, of course, is ensuring that the program is implemented with fidelity and high quality so that outcomes correspond.

**WHERE WE GO NEXT**

We would be the first to acknowledge that not all organizations active in the nonprofit sector should strive for the last, proven level of effectiveness. Depending on the program type and strategy goal, such a focus on effectiveness may not be possible. An arts group or advocacy organization may not have the means to demonstrate from a scientific perspective that their strategy for advancing sound art or public policy is “proven.”

The present challenge is in understanding where an evidence-based approach fits and where it does not. At EMCF, we remain fully committed to striving for proven effectiveness for all the organizations we fund—and the reality is that a majority of our investments are made in programs and visionaries still working toward this goal. We do not use the gold standard to put pressure on our grantees; rather, it is a tool for exploring how we may improve practice, manage more effectively, and have greater mutual confidence that the impact created is the one we want.

Will all EMCF grantees attain proven effectiveness? Perhaps not. But as they try, we are still making significant financial investments in them as they try to improve their performance over time, test their assumptions, and build their capacity to pursue the goals they set for themselves. And we think that these are worthwhile investments even if we reserve our largest investments for those who do reach proven effectiveness.

The shortcoming of the current focus on evidence-based programs, of course, is the misapplication of the concept regardless of domain or program design. But rather than abandon the movement to document evidence of a program’s effectiveness, we should attempt to do even more in this regard—to seek out the limits of evidence-based frameworks, understand where these approaches to both program and funding have the greatest value, and work to determine the boundaries where these approaches should and *should not* be applied.

The bottom line is this: We must hold on to the ambitious objectives and vision of our work as we develop more effective ways of assessing its impact. Otherwise, we risk locking ourselves—grantee and grantor—into what I have previously called a “dance of deceit,” a relationship between those with the funds and those in pursuit of them that is grounded in a false understanding of the relative value of competing programs—a dance that has us confusing our aspirations with the realities of advancing innovative, yet truly effective, solutions to our world’s most vexing challenges.

This will not be an easy process. But the many thousands of people our organizations work with in communities across the nation— young people needing real options, those outside the economic mainstream, or people caught in the trap of substance abuse—make it worth every penny and ounce of effort on all our parts. ■