Philanthropy

Measuring Impact Isn't for Everyone

Collecting data to demonstrate your organization's impact is great to do when you should, wasteful when you should not.

By Mary Kay Gugerty & Dean Karlan | Apr. 2, 2014

Many organizations claim to implement programs and policies that benefit the world's poor without evidence of impact beyond anecdotes. As an example, microcredit organizations touted their programs as a solution to global poverty for years; with credit, people could start new lives and new businesses, increase their income, and send their kids off to school. Yet recent randomized evaluations of microcredit programs from around the world (including Mexico, Morocco, India, the Philippines, Bosnia and Herzegovina, Mongolia, and Ethiopia) suggest that the picture is more mixed. For example, an evaluation in Mexico led by Innovations for Poverty Action found that women who received loans were generally happier and weathered hard times better, but their overall financial standing in terms of income or investment in new businesses did not change.

Ultimately, impact evaluations should be comparative. It isn't just about whether something works, but also how to do the most good with scarce resources. Take school attendance for example. Anecdotal claims suggest that providing basic necessities, such as uniforms and scholarships, makes it more likely that children will attend school. But for the money, it turns out the that giving kids deworming pills is 28 times more effective than school uniforms and 56 times more effective than scholarships in increasing school attendance. However, getting credible information on impact is not easy—many organizations struggle to measure the results of their work, and often use methods and data that paint unreliable pictures of program success. We applaud the focus on impact, when feasible. But sometimes impact simply is not measurable in a credible way, and yet people (organizations, or perhaps their donors) push to measure it anyhow.

The Value of Strategic Planning & Evaluation

In this ongoing series of essays, practitioners, consultants, and academics explore the value of strategy and evaluation, as well as the limits and downsides of these practices.

🔔 FOLLOW THIS
SERIES

Even though randomized impact evaluations deliver invaluable evidence about which programs to implement, these evaluations are not possible for all programs. There are two main cases when organizations should not seek evidence about impact:

1. When that piece of evidence already exists

2. When generating evidence on impact is simply impossible to do well

Stated more broadly: We should conduct an impact evaluation only when the evaluation plan will narrow a knowledge gap.

For example, the first requirement fails (luckily) in the case of vaccines. Is there a knowledge gap on the efficacy of the measles vaccine? Perhaps we're ignorant about the medical literature, but we believe the answer is *no*. Thus, an NGO vaccinating children need not run a randomized trial to measure the impact of the measles vaccine; it violates the principle of "equipoise," which argues that researchers should run experiments only when there is real uncertainty over impacts. In fact, it would be an unethical expenditure, as the money could go to pay for more vaccines!

The second requirement falls apart when rigorous evidence simply isn't feasible or appropriate to collect. Sometimes this is the result of the question under examination (macroeconomics-level policies such as trade agreements, for instance); sometimes it is about the particular setting, size, stage, or scope of the activity. Yet this requirement is not as restrictive as many think. Often we find that settings deemed implausible are indeed plausible with a bit of creativity, and many advances over the past 10 years came from learning new approaches to conducting randomized trials on social science questions. But in many cases it still is not viable to answer the impact question well.

Unfortunately, many organizations still collect data on impact, even when it is not possible or feasible. An insistent focus on measuring impact in these cases can be costly, both in terms of money spent collecting that data (which could have better uses) and time (management's focus on bad data vs. running their program).

Instead of this wasteful data collection, organizations should work to build appropriately-sized data-collection strategies and systems that demonstrate accountability to funders and provide decision makers with timely and actionable operational data.

For a forthcoming book, called *The Goldilocks Problem*, we developed a set of principles that all organizations—regardless of their ability to assess impact—can use to build strong systems of data collection. We call these principles the CART—credible, actionable, responsible, and transportable data collection.

- **Credible:** Collect only data that accurately reflect what they are intended to measure. At a larger scale, credibility means accurately measuring the impact of a program through rigorous evaluation. At a smaller scale, credible data collection also refers to appropriateness and accuracy of chosen indicators.

- **Actionable:** Collect only the data that your organization is going to use. To make data actionable, ask if you can use the information to change the course of action at your organization—if not, do not collect it. Put simply: If all possible findings lead to the same decision, it is a waste of time and money to collect that information.

- **Responsible:** Match data collection with the systems and resources your organization has to collect it. Think about the resources you have. It is tempting to collect as much information as possible, but if overreaching will compromise the quality of data you collect and your ability to analyze it, the data will not help anyone.

- **Transportable:** Apply what you learn to other programs and contexts—either your own program in future years or in other locations, or those of other organizations working on similar problems. For transportability, you need to know something about why a program works, and be open and transparent about sharing learning with others.

Researchers and organizations are working hard to widen the set of programs that we can evaluate rigorously. Even though learning about impact may become more feasible over time, in every case, organizations will better serve their mission by focusing on cost-effective, decision-driven data collection, rather than a rigid focus on impact.

Mary Kay Gugerty is the Nancy Bell Evans Professor of Nonprofit Management at the Evans School of Public Affairs at the University of Washington. She is a research network member at Innovations for Poverty Action and faculty affiliate at the Center for the Evaluation of Global Action.

Dean Karlan (@deankarlan) is professor of economics at Yale University and president and founder of Innovations for Poverty Action.

If you like this article enough to print it, be sure to subscribe to SSIR!