

## Measurement vs action: how much money should we spend figuring out what works?

Few would argue against measurement in theory. But anyone involved in philanthropic measurement knows that the best evaluations are very expensive and time-consuming. Thus the question isn't whether measurement is good but whether rigorous impact analysis is good value. Pointing to a long and rich history of failures and unintended consequences, a new generation of donors and researchers contend that rigorous evaluations are desperately needed, and relying on studies that don't meet the highest standards is sure to continue the pattern.

But randomized controlled trials and impact evaluations don't do anything direct to help those in need, others claim, and they consume much-needed funds. Critics also maintain that the results of studies are rarely conclusive, simply suggesting more studies that need to be done at additional expense and time.

While simpler measures may not meet scientific standards, are they good enough, and cheap enough, to keep philanthropic efforts on track towards their goals and the money spent on the most important things? Is philanthropy doomed to repeat its failures unless it learns to spend money on high-quality impact studies? **TO**

### IN FAVOUR OF EXPERIMENTAL EVALUATION

# Why randomized experiments are worth the cost

**Deepti Goel and Nachiket Mor**

There is good reason why randomized experiments (REs) have a long history in medicine. They are one of the most effective ways to isolate the positive and negative effects of a drug. But are they appropriate for development work? We in India have some experience of working with these tools, and our answer is a categorical yes – REs are useful to correct or adjust our beliefs as they provide clear evidence of intervention effects.



**Deepti Goel** is an Assistant Professor at the Institute for Financial Management and Research, Chennai. Email [deepti.goel@ifmr.ac.in](mailto:deepti.goel@ifmr.ac.in).



**Nachiket Mor** is a former banker and currently the President of the ICICI Foundation for Inclusive Growth. Email [nachiket.mor@icicifoundation.org](mailto:nachiket.mor@icicifoundation.org)

The views expressed in this article are solely those of the authors.

However, attention needs to be given to the construction of theoretical models as the basis for interventions and to help understand *why* interventions work the way they do. Practitioners constantly update their beliefs based on what they see in the field. But as we all know, observation and anecdote are inexact tools which often lead to the wrong conclusions. REs are important because they provide clear signals to update theory

based on data, not on the biases of the observer. REs create two groups that are identical in all respects except for the intervention itself. Any differences in group outcomes can therefore be attributed to the intervention. Using REs, researchers can make clear, factual, quantitative statements about intervention effects.

#### Unique insights provided

We feel that the greatest contribution of REs is to provide insights on specific design aspects of an intervention. A good example is the study of a microcredit programme of Spandana in Hyderabad. The study was conducted by the Poverty Action Lab in collaboration with the Center for Micro Finance (CMF) at the Institute for Financial Management and Research. Spandana agreed to randomize the rollout of its services in 104 slum neighbourhoods in Hyderabad. A baseline survey was conducted of all the neighbourhoods, while the rollout of services was staged over time. By doing a follow-up survey 18 to 24 months later we were able to learn more than ever before about why borrowers take loans and how they use the funds. For instance, we found that the availability of credit changed spending patterns significantly among those with pre-existing businesses (though less so among borrowers who started new businesses), resulting in lower overall consumption of 'temptation goods' like alcohol, lottery tickets and cigarettes, and increased investment in durable goods – both for business and for the home. This may indicate that, rather than resulting in profligate spending, as some might expect, the availability of credit encourages poor households to focus on more valuable investments and 'smart' consumption.

Another study, in West Bengal,<sup>1</sup> aimed to help micro-finance practitioners learn about clients' repayment behaviour. Using data from a field experiment that

randomly assigned clients to a weekly or monthly repayment schedule, it found repayment frequency had no significant effect on client delinquency or default. This finding suggests that more flexible repayment schedules could significantly lower transaction costs for MFIs without increasing default rates. Each of these findings will allow MFIs to improve the services offered to clients. They were only possible with the use of REs.

### Some criticisms

A common criticism of REs is the high cost. In fact, collecting data, even non-experimental survey data, is costly, so this is not an argument unique to REs. We believe that the costs are justified by the insights gained.

Another criticism of REs is that they delay programme implementation, at least to the comparison group. In properly designed REs this need not be the case. Financial or administrative concerns often dictate that an NGO phase in a programme over time anyway. In these instances randomization may be the fairest way of determining the order of phase-in. It also allows practitioners to evaluate the programme in the initial stages and make improvements as it is scaled up. IFMR Trust, whose mission is to ensure that every individual and enterprise has complete access to financial services, is planning to do precisely this. In creating a network to meet the huge, unmet demand for financial services in parts of India, it plans to follow a randomized rollout of the 300 local branches, with 15 branches in the first phase, followed by a set of branches starting every six months. This will allow them to evaluate the services being provided in each rollout and to experiment with improved services in successive rollouts.

Another challenge facing experimental studies is to measure the long-term effects of interventions. The logistics involved often make these studies prohibitively expensive. We believe, however, that this is a case of upfront investment paying off over and over again. As more programmes adopt REs, long-term data can be created that many studies can use. For instance, there is currently a CMF initiative in collaboration with the Yale Economic Growth Center to do a long-term panel survey in Tamil Nadu. The survey includes all persons in a random sample of 10,000 households located in

<sup>1</sup> Erica Field and Rohini Pande (2007) *Repayment Frequency and Default in Microfinance: Evidence from India*, Institute for Financial Management and Research, CMF Working Paper Series No 20.

## AGAINST EXPERIMENTAL EVALUATION

# Proving or improving?

David Bonbright

Few would defend current practice with regard to evaluating the difference that we make through development interventions. Let's call this difference 'impact'. The first point to make about developmental impact is that it is mind-bogglingly complex to detect accurately.



David Bonbright is Chief Executive of Keystone. Email david@keystoneaccountability.org

Let's spend a minute considering this complexity. There is the challenge to separate out all the other factors affecting the thing that you are measuring besides those factors associated with your intervention. Then there is the issue of time. Most developmental changes occur after some time, often unpredictably. If you measure at the wrong moment, you will miss the impact.

There are two extremes to the spectrum of responses to this inherent complexity. One is to abandon all efforts to get precise measures of impact and steer one's course by intuition. The other is to apply the scientific method of controlled experimentation. While I believe there are places for both of these approaches, I

200 rural villages. Collecting such data helps to cut costs by saving on baseline surveys every time a new intervention needs to be evaluated.

### More work to do

There are challenges for REs. A programme that works well in India may not work at all in other areas, but we will never know unless REs are conducted in multiple contexts. International donor agencies and philanthropists should encourage replication of studies that show promising results. Further, while it is important to know which programmes work and which do not, to add substantially to our existing knowledge of development activities it is also crucial to understand the reasons why. We therefore feel that REs should be guided by theoretical models, which provide hypotheses to test against experimental data. Without a theoretical model it is difficult to place the findings of experiments in a broader context or to make recommendations for incorporating successful intervention in the mainstream. @

believe we need a different approach that takes into account all the different purposes that we might have when we measure.

### Shortcomings of the experimental approach

Before coming to my preferred approach, I want to critically examine the experimental model. First, it really is expensive. This might be OK if there were not much cheaper alternatives. Second, while a properly conducted experimental evaluation will prove that X programme will produce Y outcome, it does not tell you why. Consequently, it is not particularly helpful for scaling up X programme to achieve Y in other settings (context turns out to matter, as many studies have shown).

Nor is it helpful for project improvement and learning purposes. When all is said and done, the circumstances in which experimental methods are cost-effective are very few – say 5 per cent of cases. Given that we want to get a much better handle on the difference we make for more than 5 per cent of cases, what should we do?

### In search of the 'right' metrics

I suggest three things:

- ▶ Adopt the system of measurement that best contributes to meeting the six main goals of evaluation.
- ▶ Constantly try out new ways to assess your work, but stop collecting data that you do not use.
- ▶ In addition to publishing what you have found through your evaluation efforts, publish what those most affected by your intervention say about what you say you have achieved.

Any organization can do these three things at an affordable cost.

The six goals of evaluation are: to improve interventions, to demonstrate impact, to inform strategy, to be accountable, to build capacity, and to educate society. If your aim is to contribute to meeting all six goals, you will need to take account of all the different producers and users of the data.

The perfect metric would of course contribute to all six goals. My favourite comes from Alcoholics Anonymous. It is produced the first time an addict speaks at an AA meeting: 'I have been sober for 532 days.' This metric simultaneously aids improvement (identifies where individuals are relapsing), demonstrates impact (sobriety being the goal that is directly measured) and builds capacity (in speaking it, addicts strengthen their resolve and self-determination).

### Charles Keidan

#### What we need is a philosophy of evaluation



Both writers agree about the need for measurement to inform action. The debate is about what kind of measurement provides the most useful data at the lowest cost. As this debate intensifies, it is important to remember two things: first that the need for evaluation should apply equally to funders and grantees. Second, that a focus on measuring results should reinforce, rather than distract from, understanding the reasons why we choose to intervene in the first place. What we need is a philosophy of evaluation as well as a science of evaluation. David Bonbright's more nuanced approach to evaluation hints at the possible outlines of such a philosophy.

### Bernhard Lorentz

#### Intelligent monitoring and evaluation does more than collect data



In the German philanthropic landscape rigorous evaluation has until now been an exception. That is why we would like to see more systematic monitoring and evaluation, like randomized trials, for example. There is no general answer to the question of how much resources should go into measurement: much depends on the field of intervention and policy area. In some cases, we would prefer to have too elaborate an evaluation rather than having no data or insufficient data, especially in the field of education and research grants, where there is not much nuanced data available.

Having stated that, intelligent monitoring and evaluation serves more goals than mere data collection: it is first and foremost an organizational learning tool and should inform strategic grantmaking. This is why we should develop more creative methods, which do not necessarily cost much. It is not that (German) philanthropy cannot know what it achieves with its programmes because it lacks resources. Resources are almost always scarce. More often it does not know because it does not put enough effort into the initial programme design and fails to develop flexible measurement systems right from the start.

My organization helps design and implement planning, evaluation and reporting systems. Our work strongly suggests that most evaluation data goes unused, and that there is insufficient experimentation to discover those few measures that really are useful. It is fascinating to see how hard it is for organizations to stop collecting data that they admit they don't use. This brings me to the last point – that collecting the views of those who are meant to benefit from our work really would be useful.

### The feedback principle of public reporting

In the world of business, billions of dollars are spent actively listening to customers. Much of what customers think is published widely in comparative formats. If we are serious about understanding how the people

## Basic education vs higher education: where should support be going?

Educational experts agree that education systems are broken at both the top and the bottom around the world. At the bottom, too many children do not receive a basic elementary education, regardless of whether they ever enrol in school. Many countries claim high enrolment while failing to provide those enrolled with the education they seek.

At the top, many countries have either a non-existent or deeply flawed secondary and tertiary educational system. The number of highly skilled secondary school and university graduates, particularly from impoverished or racial/ethnic minority populations, is far less than it should be in practically every country in the world.

What is the proper role of philanthropy in this area? Funders of education rarely advocate supporting education for education's sake – it is a means to an end. Where will investment provide the most leverage for systemic change? Should philanthropic funds be directed to ensuring that every child receives a high-quality basic education? Or can philanthropy have a greater impact at the top end of the scale, ensuring that at least some among disadvantaged populations have the necessary education for global competitiveness and can kick-start further development in their communities? Supporting a small number of high achievers might make more difference to national economies but can philanthropy forget the people at the bottom? **TO**

CONTINUED FROM PAGE 50

who are meant to benefit see our work, we need to follow this example. By definition, unlike paying consumers, our primary constituents don't pay and don't have consumer choice. By definition, almost, they have little structural power over our organizations. By collecting and publishing their views of the organization's progress measures, we create the conditions in which the entire ecosystem can be visible to all – from society-at-large to the funders, the implementers, and those meant to benefit. Please note that I am not saying the primary constituent is always right. I am arguing that what the primary constituent says should be cultivated, responded to, and published (with interpretation if necessary).

A friend who runs a homeless shelter told me a story that illustrates this point. One day one of his regular clients – someone he feeds and shelters several times a week – came in and asked him for a marker pen. He took the pen, opened out a cardboard box into a sign and wrote on it, 'This place mistreats the homeless.' He then went outside and sat in front of the shelter holding the sign. The point is this. If we are going to solve the problem of homelessness, we need to learn more about the homeless. We need to hear their voices.

Our resources are limited. Our management and measurement systems need to be cost-effective, nimble and interactive. It is a time of creativity and productivity in the world of impact measurement. Significant productivity gains (ie increased impact for the same amount of effort) can be realized by following my three suggestions. @

## IN FAVOUR OF BASIC EDUCATION

# Betting on the bottom-up Rohini Nilekani

Just recently, I happened to be standing outside a rather new-looking government school, high up on a scraggy hillock in Kolar district in South India. Two young boys, aged about 12, were our curious onlookers and we picked up conversations. One I knew to be a student, because he wore a school uniform, often the only decent clothing accessible to poor rural children, even today. The other, through the entire agony of his body language, clearly was not.



Rohini Nilekani is Chairperson of the Arghyam Foundation and of Pratham Books. Email [rohini@arghyam.org](mailto:rohini@arghyam.org)

His name was Anil and he had had unspeakable experiences. No mother, often absent father, not enough to eat, no school exposure ever. I gave each of them a story book. Predictably, the eighth grade student immediately began to read the first paragraph, but haltingly. Anil simply held the book as if it were a precious glass vase.

This scene provides the premise for my argument. Until Anil can bridge the social distance between himself and the school, until his school-going friend learns to go beyond reading words into articulating contextual ideas, we as philanthropists are not done with our basic social engagement. However long it takes, ensuring universal quality elementary education is the urgent work of our time. In the modern nation state, effective citizen engagement and empowerment is increasingly predicated on basic literacy and, with it, access to the primarily textual language of power and knowledge.