

Stanford SOCIAL INNOVATION^{Review}

Features

The Power of Lean Data

By Sasha Dichter, Tom Adams, & Alnoor Ebrahim

Stanford Social Innovation Review
Winter 2016

Copyright © 2015 by Leland Stanford Jr. University
All Rights Reserved

➔ For years, the complex and costly nature of impact measurement has kept many social enterprises from doing it—or from doing it well. But a series of recent projects that incorporate lean design principles show that it's possible to gather high-quality impact data quickly and inexpensively.

The Power of Lean Data

BY SASHA DICHTER, TOM ADAMS, & ALNOOR EBRAHIM

Illustration by DAVID PLUNKERT



ZIQITZA HEALTH CARE LIMITED, a social enterprise in India that provides ambulance services, aims to make those services accessible to all segments of society. The company, founded in 2004, operates nearly 1,000 ambulances in six Indian states and answers more than 2.5 million calls per year. But for the first 10 years of its history, Ziqitza lacked reliable data on who its customers were and on whether it

was reaching the poorest people in its areas of operation.

This problem—an inability to gather usable impact data—is remarkably common in the social sector. But it need not be so. Acumen, a nonprofit organization that promotes innovative ways to alleviate poverty, has undertaken a series of projects that involve a new approach to impact measurement. Two of us, Sasha Dichter and Tom Adams, work at Acumen and helped lead these projects. The third co-author, Alnoor Ebrahim, is a professor at Harvard Business School who has worked with us to study these efforts.

In early 2014, our team at Acumen worked with Ziqitza to train call center employees in two states, Punjab and Orissa, to pose a set of 10 questions to customers. We drew those questions from the Progress Out of Poverty Index (PPI), a survey developed by the Grameen Foundation.¹ The PPI survey uses straightforward, neutral questions, such as “How many members does your household have?” and “What is the main source of lighting fuel for your household?” in order to gather data on poverty levels among a given population. The simplicity of the questions makes it easy to administer the survey during a short phone interview. Training Ziqitza’s call center operators took just one day, and Ziqitza was able to integrate the survey effort seamlessly into its operations. Within a month of that initial training, the company had surveyed 1,000 of its customers.

The results showed that three-quarters of Ziqitza customers were

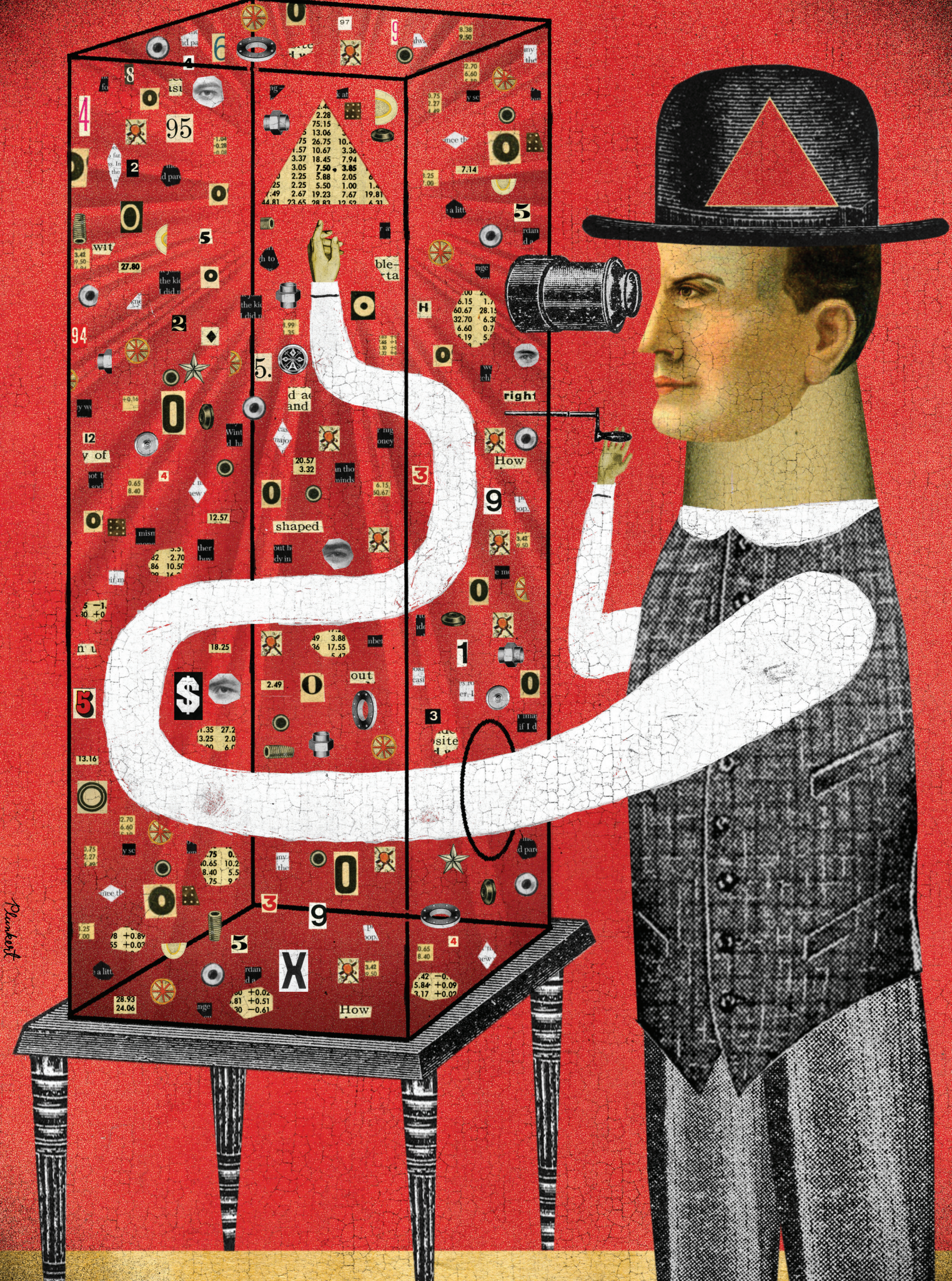
living below the World Bank poverty line of \$2.50 per day and that the company was serving women—pregnant women in particular—at a disproportionately high level. The survey also revealed areas for improvement. In rural Orissa, for example, Ziqitza’s penetration among those below the poverty line fell short of the state average by 11 percentage points.

The Ziqitza survey project was a pilot initiative in the use of lean data, an approach that Acumen has developed to meet the measurement needs of social enterprises in its investment portfolio. (Acumen has developed the lean data approach with grant support from the Aspen Network for Development Entrepreneurs and the Omidyar Network.) Lean data involves the application of lean experimentation principles to the collection and use of social impact data.² The approach incorporates two main features: first, a shift in mindset away from reporting and compliance and toward creating value for a company and its customers; and second, the use of methods and technologies for data collection that favor efficiency and speed while maintaining rigor.

Lean data embraces the uncertainties and complexities that are inherent in building a social enterprise. (Our work has targeted social enterprises, and in this article we focus on that type of organization. But the lean data method is relevant to any organization that operates in a dynamic, resource-constrained environment.) The lean data approach tackles head on the common belief that assessing social enterprise performance is inherently burdensome and expensive. In fact, the direct cost of implementing lean data projects is relatively low, and the payoff can be significant: In many cases, these projects not only yield high-quality data but also help companies build data collection systems that will become integral to their future operations.

THE IMPACT MEASUREMENT IMPASSE

Nearly all impact investors—95 percent, according to a survey conducted by JPMorgan Chase & Co. and the Global Impact Investing Network (GIIN)—say that they measure and report on the social



Plunkett

impact of their investments.³ But a close look at the impact investing field reveals that the state of practice is far from robust. Most impact reporting focuses on output measures such as “number of lives reached” or “number of jobs created.” Few investors or entrepreneurs seek to understand, in a deep way, how customers experience the goods or services that an enterprise provides. Nor do people in the field give much attention to demographic factors such as the income levels or the gender make-up of customers. As a result, we have little information on whether social enterprises are reaching those who most need their goods or services.

To be sure, the impact investing field has made progress in building standardized performance metrics. Acumen, for example, played a pivotal role—along with GIIN, the Rockefeller Foundation, and B Lab—in developing the Impact Reporting and Investment Standards (commonly known as IRIS). IRIS provides a catalog of standardized metrics that any impact investor can choose to track.⁴ Yet impact investors typically collect data only on the financial or operational metrics in the catalog. They seldom devote resources to tracking the social metrics.

In part, this paradox—a stated interest in impact measurement, combined with a pattern of weak measurement practice—reflects a justifiable concern about both the costs and the benefits of rigorous impact assessment. The problem is that neither the tools of emerging-market investment nor the measurement practices of traditional international development are appropriate to early-stage social enterprises. Standard business metrics (numbers of customers, market penetration, revenue totals, and so on) reflect the financial performance of a company but do little to capture its social value. And the monitoring and evaluation (M&E) methods commonly used by international aid agencies involve multiyear data collection efforts that are feasible only for well-established organizations that have substantial measurement capacity. Take the use of randomized control trials (RCTs), which many experts deem to be the gold standard of impact measurement. RCTs can certainly provide a wealth of valuable data. But they are costly, take years to complete, and require substantial expertise to execute properly. They also require conditions—such as the ability to establish both a “treatment” group and a “control” group—that most start-up enterprises are ill equipped to provide.

All too often, traditional M&E approaches result in elaborate reports that merely sit on funders’ desks. Rarely do enterprises use those reports to inform their ongoing work.

THE SOCIAL ENTERPRISE CONTEXT

The flaws in the traditional approach to impact measurement have led to an accountability gap. Social entrepreneurs have fallen into the habit of conducting evaluations that meet the needs of *upward* accountability: They collect data to meet the requirements of their investors. (And investors, in turn, often set those requirements in response to the reporting expectations of their limited partners.) What is often missing is a commitment to *downward* accountability—to making sure that social enterprises are using data to improve the lives of their intended beneficiaries.⁵

It’s hardly surprising, therefore, that social entrepreneurs have become increasingly frustrated with the conversation around measuring impact. They rightly lament that practical tools do not exist to help them measure, analyze, and improve the impact that they are delivering to customers. They bristle at the prospect of conducting

SASHA DICHTER is chief innovation officer at Acumen.

TOM ADAMS is director of impact at Acumen.

ALNOOR EBRAHIM is an associate professor at Harvard Business School.

Disclosure: Ziqitza, Edubridge, SolarNow, and other social enterprises discussed in this article are part of the Acumen portfolio. Sasha Dichter and Tom Adams have worked directly with these companies in that context.

large-scale impact assessment efforts that do not align with the day-to-day reality of their business. For most social entrepreneurs, the following attributes help to define that reality.

- **A dynamic environment.** As the economist William Easterly notes, start-up enterprises usually function as “searchers”: They are constantly testing and iterating their business models in order to build better solutions for their customers.⁶ They make decisions about their impact models within a context that is constantly changing as well.
- **Financial constraints.** A typical start-up social enterprise operates with at most a few million dollars of funding. So any project that it undertakes must be relatively inexpensive.
- **Limited human capital.** Newly formed social enterprises must focus on recruiting seasoned managers who can run a business. Hiring people with deep expertise in traditional forms of impact assessment is necessarily a low priority for them.
- **Poor data management systems.** Few early-stage social enterprises have the resources to invest in systems that would allow them to record, store, and manage impact data.

What social entrepreneurs and those who invest in them need is an approach to impact measurement that reckons with these attributes. Such an approach will have several core properties—properties that we encapsulate in an easy-to-remember acronym: BUILD.

- **Bottom-up.** It nurtures the habit of listening to customers in order to provide actionable insight on their needs and interests.
- **Useful.** It yields data that is of sufficient quality to support decision-making.
- **Iterative.** It allows for learning, adaptation, and replication.
- **Light-touch.** It uses low-cost tools and technologies that require a minimal investment of time and money.
- **Dynamic.** It enables rapid data collection within a fast-changing environment.

THE LEAN DATA WAY

Lean data reframes data collection and impact measurement in a way that corresponds to a real-world social enterprise context. Two important developments have paved the way for this new approach.

First, the near ubiquity of mobile phones makes it possible to communicate quickly and directly with customers even in far-flung rural areas. Cheap text messaging and capabilities such as interactive voice response (IVR) provide robust, efficient means by which to contact customers. (IVR technology enables automated phone communication and allows customers to answer questions at the push of a button.)

Second, customer feedback tools, such as the PPI and the Constituent Voice survey (a feedback tool developed by Keystone, a nonprofit social measurement firm), allow organizations to collect meaningful data while making limited demands on customers’ time

and attention.⁷ To be sure, such tools aren't new. A generation ago, for example, researchers developed "participatory rural appraisal" methods—methods that rely on oral communication, along with symbols and pictures—to survey people in rural areas.⁸ But the growing availability and increasing sophistication of such tools, combined with the use of mobile technologies, have made it much easier to engage in data collection efforts that have the core properties we have described.

By design, the lean data process is simple and clear. In many cases, after people in a company have gone through the process once, they will be able to repeat it or to adapt it without depending on extensive outside support. (See "How Lean Data Works" below.) A lean data project starts with the development of an *impact question* that an enterprise seeks to answer. In this phase, leaders of the enterprise define the specific thesis that they want to test. As part of that process, they gather feedback from customers about the impact of a given product or service.

Next comes the *design* phase, in which the leaders identify an *enabling technology* and an *enabling instrument* that they will deploy in their project. The enabling technology might be SMS, IVR, or a call center, for instance. The enabling instrument might be a pretested survey (the PPI, for example), or it might involve using a carefully structured focus group.

In the all-important *execution* phase, the enterprise develops a concrete plan for gathering data from people in its target market. During this phase, those who manage the project train staff members how to use the enabling technology and the enabling instrument, and they test their plan via rapid prototyping.

Once leaders of the enterprise have data in hand, they enter the *learning* phase. They analyze the data, extract lessons from the project, and determine how to apply these lessons to company operations.

The last phase centers on *action*. At this point, leaders of the enterprise decide how they will use their newly acquired knowledge. As part of this phase, they also decide whether and how to apply the lean data process to other impact questions.

PROJECT HIGHLIGHTS

Over the past two years, Acumen has developed and executed lean data projects at 12 companies that operate on multiple continents. We now have several additional projects under way, and we aim to complete as many as 20 engagements by the end of 2015. (See "Lean Data in Action" on page 40.) Two of our projects, in particular, illustrate the power of the lean data approach.

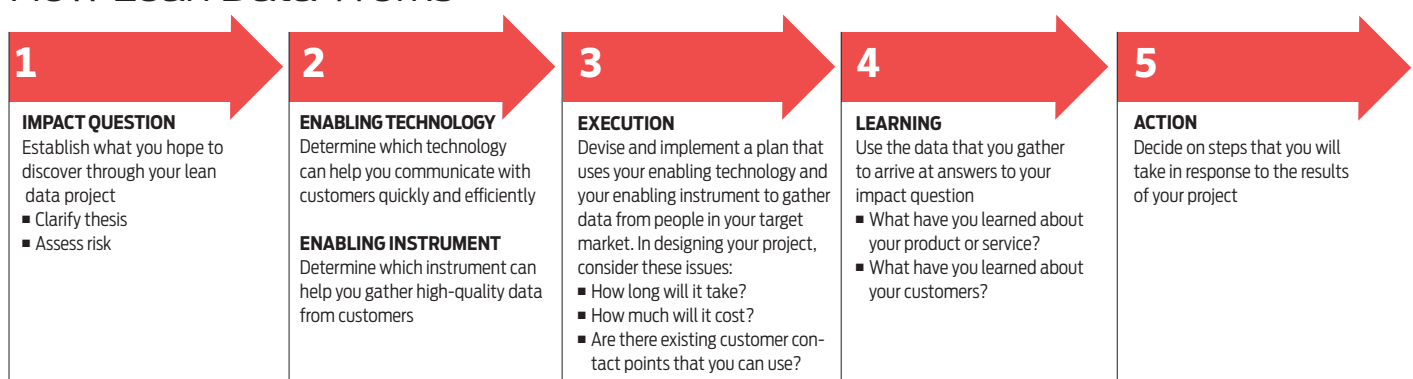
Training ground | Edubridge is a vocational training company that seeks to improve the labor market outcomes for workers in India who are migrating from rural to urban areas. Girish Singhania, CEO of Edubridge, had been puzzling over a question that is critical to his company's theory of change: How do "successful" trainees—those who obtain and accept job placements immediately after they undergo Edubridge training—differ from trainees who don't? Singhania didn't have the luxury of time. To guide the growth of his company, he needed an answer to that question in a matter of weeks.

Acumen, an early-stage equity investor in Edubridge, proposed a phone-call-based survey that would leverage Edubridge's existing call center employees, who were fluent in four Indian languages and who already knew how to build rapport with trainees. Edubridge had a database of phone numbers that enabled it to build a sample that included several discrete populations: people who had expressed an interest in Edubridge courses but had never signed up for one, people who had completed an Edubridge course but had not accepted a job offer that they had received afterward, and people who had both completed a course and accepted a job offer.

From the initial conversation between Singhania and his partners at Acumen to the presentation of survey results, the Edubridge lean data project took just four months. Call center operators set aside one hour of their time per day for survey calls and were able to meet their usual responsibilities in the remainder of their shift. They completed 650 calls in all, and each call lasted seven to eight minutes.

The results provided rich insight into Edubridge's customer base. Singhania had hypothesized that trainees with close friends in urban areas would be more likely to accept jobs than other trainees. That turned out to be true: Trainees who had friends in a city where a job was located were 21 percent more likely to take that job than trainees who didn't have friends there. Members of the Acumen team expected that trainees from higher-income families would be more likely to accept jobs than trainees from lower-income families. That hypothesis turned out not to be true. Those who had accepted jobs were 8 percent poorer than those who had not. (We are still working to make sense of the latter result. It could be that poorer trainees have comparatively fewer alternatives and are therefore more likely to accept the job offers that they receive.) Singhania is now using data from the survey to shape Edubridge's customer segmentation strategy as the company prepares to expand its operations to 100 training centers over the next several years.

How Lean Data Works



Survey power | SolarNow, based in Uganda, markets solar energy systems to off-grid households and micro-entrepreneurs. Willem Nolens, managing director of SolarNow, wanted to know how the company could make its systems more accessible to poor customers. SolarNow systems are more powerful but also more expensive than alternative energy solutions. To make its products more affordable, SolarNow had established an in-house financing service. It had also leveraged a government subsidy, funded by the World Bank, that gave consumers \$250 for the purchase of a home solar system that was at least 100 meters (about 330 feet) from the main power grid in their community. Nonetheless, it was not clear whether SolarNow’s target customers could afford its products. Early in 2014, when the World Bank withdrew its subsidy, the issue of affordability became even more salient.

Nolens and his team took numerous steps to improve affordability. To cut costs, SolarNow established a direct purchase agreement with a manufacturer that allowed the company to avoid working with local middlemen. SolarNow also extended the duration of its financing plan from 12 months to 18 months. Drawn by Nolens’s commitment to reaching the poorer segments of the Ugandan population, Acumen decided to invest in SolarNow in June 2014.

But the company’s data on customers, and particularly on customer incomes, remained spotty. So Acumen developed a 10-minute survey that uses PPI questions to collect (among other metrics) data on the poverty levels of SolarNow customers. Then, in just two days, the Acumen team trained SolarNow’s call center employees to conduct the survey. The results showed that nearly half of SolarNow’s customers—a considerably larger proportion than the company had expected—live on less than \$2 per day. This finding illuminated the demand among

poor customers for SolarNow products and affirmed the effectiveness of the steps that Nolens had taken to increase affordability. The survey data also provided insight into which price points would make the purchase of a SolarNow system affordable to poor customers and how that purchase might affect the household economics of buyers.

EMERGING INSIGHTS

Today, nearly two years after launching the Acumen lean data initiative, we are in a position to draw some preliminary lessons. First, the collection of meaningful data—data that early-stage enterprises can use immediately to inform strategic decisions—begets a culture of measurement. People in a social enterprise typically view impact measurement through the lens of compliance: They see it as an obligation to their funders. But once it becomes relatively easy for them to gather high-quality impact data, their attitude toward measurement changes dramatically. They become eager to collect and use data related to social impact.

Second, the insights about customers that arise from lean data efforts can help a company close the accountability gap. Lean data opens up a channel for listening to customers, and the opportunity to gather customer feedback on a large scale can be immensely powerful. SolarNow learned that its efforts to increase affordability have attracted far more low-income customers than it had expected to reach. Similarly, Ziqitza learned that pregnant women make up one of its core customer segments; that knowledge has given the company a point of focus as it works to reach new markets.

Third, entrepreneurs can conduct lean data projects quickly and at low cost. In our work with Acumen portfolio companies,

Lean Data in Action

COMPANY INDUSTRY LOCATION	IMPACT QUESTION	LEARNING	ACTION	METHOD (COST)	QUALITY ASSURANCE
Burn Cookstove sales Kenya	Do buyers of Burn cookstoves reduce charcoal usage to the expected degree (by 60 percent) as a result of using the product?	Customers’ use of charcoal decreases to the same degree as laboratory tests had predicted it would	Working to understand which distribution channels are most effective at reaching poor rural customers	SMS texting, call center (\$3,000 over the course of four weeks)	Using call center interviews to validate SMS data
Edubridge Vocational training India	Which factors account for the difference between “successful” and “unsuccessful” Edubridge trainees?	Customers with better urban social networks are more likely to get and keep a job; poorer students are more likely to seek out jobs and keep them	Using data to segment customers and working to improve the match between training services and trainee needs	Call center (\$1,500 over the course of four months)	Using Acumen personnel to train call center staff
KZ Noir Coffee processing Rwanda	What is the poverty level of KZ Noir farmers? Do the earnings of these farmers improve because of their participation in the KZ Noir premium-sharing program?	Preliminary results show that 59 percent of KZ Noir farmers are extremely poor; data on changes in farmers’ income are forthcoming	Using results (such as a high incidence of bank account usage among customers) to improve premium-sharing program	SMS texting and in-person tablet-based surveys (\$15,000 over the course of two months)	Enlisting researchers at a third-party firm (IDinsight) to implement survey
SolarNow Solar energy systems Uganda	Does in-house financing improve the ability of SolarNow to reach the poor? Does household expenditure on energy decline as a result of using a SolarNow product?	Reach among the poor is better than expected (49 percent of customers live on less than \$2.50 per day); most customers show only a small savings in energy expenditure in the first three to four years after purchase	Starting to conduct quarterly surveys to track customer segmentation and customer satisfaction	Call center (\$2,000 over the course of four weeks)	Using a third-party call center as well as remote field staff; using SolarNow’s call center for follow-up survey to test for consistency of responses
Ziqitza Health Care Limited Ambulance services India	Does the Ziqitza ambulance service succeed in reaching low-income customers? How does usage vary between male and female customers?	Most customers (75 percent) are below the poverty line; pregnant women are a core market	Improving access in hard-to-reach regions; working with local governments in remote areas to increase public awareness	Call center (\$14,000 over the course of four months)	Working with Grameen Foundation to train call center staff and to conduct in-person validation of sub-sample

Each of these companies is part of the Acumen portfolio. Cost figures do not account for Acumen staff time.

the direct cost per engagement has ranged from \$500 to \$15,000, and the duration of data collection has ranged from 10 days to 4 months. (Those cost figures do not take into account the cost of Acumen staff time.) In many cases, companies have been able to collect data through existing customer contact points. Both SolarNow and Ziqitza, for instance, were able to collect new data via standard follow-up calls. KZ Noir, a company that buys raw coffee beans from smallholder farmers in Rwanda, has used a combination of questionnaires administered by its sales force and SMS surveys to gather data.

Fourth, the lean data process doesn't always run smoothly. It requires iteration to ensure data quality. We're learning a great deal about the best ways to ask questions through SMS, IVR, and other platforms, and we have a long list of failed questions to show for it. Sometimes the problem relates to the format—using text messages alone can lead to a loss of essential nuance—and sometimes it is the questions themselves that create unexpected confusion. In any event, because these technologies lend themselves to rapid testing, we are able to figure out quickly which questions work or don't work in a given target market.

We are learning that one way to ensure the quality of lean data is to supplement SMS and IVR questions with in-person verification surveys. Doing so allows us to gauge the reliability of various data collection approaches. Reliability, we have discovered, often varies by question type. Take the example of LabourNet, a vocational training company in India. In our work with LabourNet, we used SMS and IVR to pose questions to former trainees about their current wages and employment status. Afterward, we enlisted call center staff members to verify selected trainee responses. In this instance, we found that the reliability of data gathered through SMS and IVR was lower than we had expected.

Fortunately, instances in which verification has resulted in concerns about data quality are fairly rare. But those cases point to the need to generate more knowledge about lean data approaches. We need to hone our understanding of which types of questions work best in which format (SMS, IVR, call center); how to draft and structure surveys for each technology in a way that will deliver reliable responses; and how to combine various technologies and instruments to achieve optimal results. Here's an example of how we are refining the lean data method: In working with Guardian, an India-based microfinance provider, we used an automated IVR message to tell customers that they would receive a survey call from an interviewer within the next few days. Doing so, we discovered, significantly increased survey response rates.

In short, we now know that the lean data process generates meaningful and timely results. But we need to keep testing different kinds of questions using different technologies in different settings. As we move forward, we may encounter innovations that allow us to solve persistent data-collection challenges. Recently, for example, we started experimenting with the use of sensor technology to collect real-time data. Through sensor technology, we can remotely measure patterns in usage for fuel-efficient cookstoves and other products.

BEYOND THE METRICS MYTH

The lean data approach is still in its early days of development. At Acumen, we continue to learn new ways to implement lean data techniques, and every project generates new insights. But if our experience

SSIR ONLINE

Visit ssir.org to learn more about the lean data method.

- ▶ "Innovations in Impact Measurement" report
- ▶ "Lean Data Field Guide"

with lean data has taught us anything, it is that social entrepreneurs can break what Jed Emerson calls the "metrics myth."⁹ Emerson, in explaining what he means by that term, emphasizes the wide gap between the rhetoric of social impact measurement and the actual state of practice in this field.

Lean data can close that gap. It has the power to shift the impact measurement conversation away from experts and toward social entrepreneurs—away from the use of complex, costly methods and toward the use of simple, inexpensive tools. With high-quality data in hand, impact-driven companies can iterate faster and achieve their missions with greater efficiency. Lean data shifts power back toward social enterprises by helping them measure and deliver social value to their customers.

Consider again the example of SolarNow. The insights that the company gained through the lean data method were "a real eye-opener," Nolens says. Indeed, he reports being "shocked by how honest [customers'] answers were." Along with providing information on the use of the SolarNow product by very poor customers, the lean data survey delivered important feedback on when and why customers were unhappy with the product. The survey shed light on problems that were inhibiting customers from realizing the full potential of their solar-power systems. Nolens and his team are now dealing with those problems by ramping up after-sales support. To track improvements, he has directed members of his call center staff to repeat the lean data survey every quarter. "Other M&E organizations put pressure [on us] to do detailed impact surveys that are insightful to them but not insightful to us," Nolens comments. The questions used in SolarNow's lean data survey, by contrast, were "simple, relevant, and not intrusive," and they yielded an "ideal combination of customer insight [and] social performance data," he notes.

Ultimately, the power of lean data extends far beyond measurement. Lean data offers a way to increase accountability between an enterprise and its target customers. It also allows an enterprise to move beyond *proving* what worked (or didn't work) in the past so that it can focus on *improving* its impact right away. ■

NOTES

- 1 Grameen Foundation, Progress Out of Poverty website, <http://www.progressoutofpoverty.org>
- 2 Peter Murray and Steve Ma, "The Promise of Lean Experimentation," *Stanford Social Innovation Review*, Summer 2015.
- 3 "Spotlight on the Market: The Impact Investor Survey," JPMorgan Chase and the Global Impact Investing Network, May 2, 2014, <http://www.thegiin.org/binary-data/2014MarketSpotlight.PDF>
- 4 "The Landscape for Impact Investing in South Asia," Global Impact Investing Network and Dahlberg Global Development Advisors, April 8, 2015, http://www.thegiin.org/binary-data/Full_GIIN_South_Asia_Report-1.pdf
- 5 Alnoor Ebrahim, Julie Battilana, and Johanna Mair, "The Governance of Social Enterprises: Mission Drift and Accountability Challenges in Hybrid Organizations," *Research in Organizational Behavior*, 34, 2015.
- 6 William Easterly, "Planners Versus Searchers in Foreign Aid," *Asian Development Review*, 23, 2006, <http://www.nyudri.org/wp-content/uploads/2011/10/plannersvssearchers.pdf>
- 7 "Constituent Voice," Keystone Accountability website, <http://www.keystoneaccountability.org/analysis/constituency>
- 8 Robert Chambers, "Participatory Rural Appraisal (PRA): Analysis of Experience," *World Development*, 22, 1994.
- 9 Jed Emerson, "The Metrics Myth," BlendedValue website, February 10, 2015, <http://www.blendedvalue.org/the-metrics-myth>