
Advancing Evaluation Practices in Philanthropy



Framing the Issue P. 2

By Jane Wales

A Focus on Culture P. 4

By Luis A. Ubiñas

Learning from Silicon Valley P. 8

By Matt Bannick & Eric Hallstein

Shared Outcomes P. 12

By Judith Rodin & Nancy MacPherson

Risky Business P. 16

By Paul Brest

Assessing One's Own Performance P. 20

By James E. Canales & Kevin Rafter

Framing the Issue

BY JANE WALES

In recent years, the philanthropic sector has neared consensus on the need to improve measurement and evaluation of its work. Although the philanthropies they lead use different methods, members of the Aspen Philanthropy Group (APG) have agreed that basic principles and practices can inform efforts to monitor performance, track progress, and assess the impact of foundation strategies, initiatives, and grants. They hope to build a culture of learning in the process.

Over the past two years these CEOs of private, corporate, and community foundations have supported a series of meetings on measurement and evaluation (M&E) with leaders of grantee organizations, issue experts, and evaluators. They have concluded that, when done right, assessment can achieve three goals. It can strengthen grantor and grantee decision-making, enable continuous learning and improvement, and contribute to field-wide learning. Below are broad observations from the workshop process, followed by articles from five APG authors describing the M&E philosophies of the institutions they lead. Their articles will be among those to appear in an edited e-volume, to be published by the Aspen Institute and continuously updated to capture evolving foundation practice and comments from voices in the field. This is what we learned.

Definitions Matter | APG members found that differing terminology can undermine efforts by grantors and grantees to collaborate effectively in the design and implementation of an M&E system. Many grantors and grantees use the terms “evaluation,” “impact measurement,” and “measurement and evaluation” interchangeably. In fact, M&E encompasses distinct activities with distinct purposes, methods, and levels of difficulty. In his article, William and Flora Hewlett Foundation president Paul Brest separates M&E into three categories undertaken at three stages: theories of change described and logic models devised during the initial design of a project or foundation initiative; tracking progress against the strategy set during the life of the grant or initiative; and assessing impact after the fact. The first of these is essential background for M&E, and the three together provide a useful means of organizing the various activities and purposes of M&E. The second enables a grantor and grantee to gain the information needed to make mid-course corrections to the strategy and learn throughout the process. The third activity—assessing impact—is the most daunting. Brest notes that in some undertakings, such as policy advocacy or Track II diplomacy, exogenous influences

make it hard if not impossible to attribute impact to any one actor or strategy. He argues for demonstrating “contribution” rather than claiming “attribution,” where contribution means increasing the likelihood of success, and notes that the true impact of such “risky” grants may not be possible to ascertain. Nonetheless, they are well worth pursuing.

Purpose Matters | At its best, M&E informs decision-making and provides for continuous learning. In his article, Matthew Bannick, managing director of Omidyar Network (ON), discusses why M&E is more likely to be used—and used to good effect—when it is designed collaboratively by grantor and grantee, and when data are gathered and organized around decisions that each needs to take. It is therefore critically important that they agree on their evaluation approach at the outset. Ford Foundation president Luis Ubiñas agrees, adding that “from the very beginning, grantees should have a clear sense of what benchmarks of success are expected of them at each stage of initiative development”—and why.

The Cost-Benefit Ratio Matters | Ubiñas points out the costs of M&E, arguing that in designing an evaluation system, careful consideration must be given to the burdens on each party. Failure to do so, he writes, can lead to “excessive data gathering” in which grantor and grantee gather as much data as possible in search of evidence of impact. The costs are fourfold. “First, it is a burden to grantees, creating surplus work for often tightly staffed and financially strapped nonprofits. Second, it undermines quality because grantees will provide the requested information to meet their grant obligations, but may not have time to supply the insight that is often more valuable than the data. Third, it inundates foundation staff with information but may leave them little time to use it effectively. Fourth, it may not provide the information that is actually needed to understand how effective our initiatives and grantmaking are.” According to Bannick, ON reduces the burden by using a limited number of easily collected metrics, and, as an alternative to the “time-consuming, costly, and complicated challenge of measuring impact,” ON often measures outputs as proxies. As for the price tag, Rockefeller Foundation president Judith Rodin notes the efficiencies gained by using technology to gather real time data. Brest notes (in *Money Well Spent*, co-authored with Hal Harvey) that “if you are a philanthropist with a long-term commitment to a field, it is well worth putting your funds—and lots of them—into evaluation.”

Culture, Context, and Capacity Matter | M&E requires a commitment to building capacity within foundations, grantee organizations, and the field of evaluation in general. The Rockefeller Foundation invests in M&E teams in both the developed and developing world

JANE WALES is vice president, *Philanthropy and Society*, of the Aspen Institute and director of its Program on Philanthropy and Social Innovation. She is the founding CEO of the Global Philanthropy Forum.

to monitor foundation initiatives and to act as “critical friends” to its grantees, establishing monitoring and learning systems where none existed. The goal is to facilitate learning among grantees and within the foundation aimed at improving performance all round. But most important, it is to leave behind greater capacity among local M&E professionals. Rodin reports that the foundation has supported regional institutions that train and mentor local evaluators and partner with similar institutions elsewhere, with the goal of building lasting capacity. Bannick speaks to the importance of providing technical assistance to grantees. And, within a foundation, James Irvine Foundation president James Canales notes, it is important for there to be leadership by trustees and senior officers, as well as a readiness to devote time, dollars, and expertise to assessing the philanthropy’s strategy, initiatives, and grants. Doing so “mandates full institutional commitment and cannot be the province of just the evaluation director.” Beyond these important tangible contributions lies the requirement for building what Ubiñas calls an “impact culture” in which continuous learning and adaptation are enabled, required, and rewarded. “Goals, theories of change, and operating approaches are all necessarily imperfect; only by learning from our successes and our mistakes can we build an impact culture,” he writes.

The Unit of Analysis Matters | The good—and the bad—news is that almost any activity can be evaluated. It is important to sort out the different units of analysis, as is done in the Bill and Melinda Gates Foundation’s *Guide to Actionable Measurement*, which notes three distinct areas of focus. At the level of foundation strategy, the focus should be on measuring outcomes over impact (as Bannick describes), on assessing contribution rather than attribution (as Brest recommends for certain grants), and on the degree of harmony that can be achieved among grantees pursuing a given strategy. At the level of foundation initiative, the foundation should use grantee reporting data on outputs and outcomes to signal whether the initiative is making progress; track the program team’s activities other than grants (such as convening and public speaking); use independent evaluation; and capture both intended and unintended consequences of the initiative. And at the level of the individual grant, the foundation should align expected grant results with strategic intent; work with the grantee to track grantee inputs, activities, outputs and outcomes at critical points to manage and adjust each grant appropriately; and measure the foundation’s input of human, financial and technical resources.

Timing Matters | Just as the units of analysis differ, so too do the time horizons required to measure and evaluate the performance of a foundation’s strategy and grants. Many short and medium-term metrics are useful in assessing how well an organization is managing processes or reaching target populations. Longer-term longitudinal studies are critical to gauging the impact of a program and to establishing the causal relationship between intervention and desired outcome. Such rigorous, long-term studies can be particularly useful for those seeking to scale up innovations. Canales notes, however, that the annual grant-cycle poses a “structural barrier” to the longer-term undertaking of evaluating and learning, noting that “program goals and aspirations rarely follow annual timelines, nor should they...if they are sufficiently ambitious.” He points to the importance of “creating the space for consideration of broader progress assessment.”

Feedback from Grantees and Beneficiaries Matters | M&E must incorporate into the process the viewpoints and observations of the funder, grantee, and ultimate beneficiary through all stages of work—identifying problems, co-creating solutions, and implementing with a shared vision of outcomes. Under the leadership of APG member Carol Larson, the David and Lucile Packard Foundation solicits feedback from the on-the-ground staff of the grantee and has established written standards to help its program team to communicate with grantees. Grantees, in turn, can better assess community needs—and their performance in helping to meet those needs—when program beneficiaries provide them quantitative and qualitative feedback. At Ford, Ubiñas notes that the foundation selects grantees “managed by those living and working closest to where targeted populations are located.” Noting the Rockefeller Foundation’s commitment to evaluation practices that include stakeholders’ voices, Rodin cites the consensus of the Africa Evaluation Association: “only when the voices of those whose lives we seek to improve are heard, respected and internalized in our understanding of the problems we seek to solve” will philanthropy achieve its purpose.

Transparency Matters | Although the goals of M&E are to inform decision-making and enable continuous learning by those immediately involved, there is a larger community to serve and a larger purpose to pursue. By publicly sharing the data gathered and conclusions reached, grantors and grantees can contribute to field-wide learning. APG members agree that this is an opportunity to seize. The philanthropic sector has helped build communities of practice that generate knowledge. These evaluators form professional associations that set standards for the field. Independent organizations assess foundation and grantee performance and publish their findings. Donor organizations and networks transfer knowledge among philanthropies, and between grant-making institutions and individuals. And academic programs provide the intellectual underpinning for much of this work. Supported by philanthropy, these and other institutions provide some of the early hardware for wider impact. Moreover, the gradual evolution of principles that guide and practices that enable rigorous evaluation can contribute to its software.

But has a true *system* for philanthropic impact been designed and widely adopted? Perhaps not. And so, in publishing an e-volume and opening a conversation, Aspen’s Program on Philanthropy and Social Innovation will seek the wisdom of the crowd and ask the questions: What might the components of such a system be? Where will the breakthroughs occur? What sort of venture capital will be needed to finance the prototypes? And what markets will bring these innovations to scale? Or, building upon Ubiñas’s apt phrase, under what conditions might an “impact culture” spread? What would it take for its language to be adopted, its standards embraced, its methods refined, and its potential realized? And if that culture is to be global, dynamic, and enduring, how might it be informed and advanced by the new cadre of evaluators to which Rodin refers?

None of the architects or beneficiaries of modern-day philanthropy would claim that they alone can create such a world, but they may well agree that it is one worth imagining. Doing so together would make for a powerful beginning. ♦

A Focus on Culture

How the Ford Foundation is engaging its global staff in building a shared culture of results

BY LUIS A. UBIÑAS

Philanthropic organizations confront some of the most exigent and enduring problems facing humankind. Our mission-driven organizations pursue broad societal goals, including reducing poverty, advancing human rights, fostering educational achievement, and strengthening democratic principles and processes. Pursuing these goals is a long-term commitment, requiring aggressive problem solving, sustained effort, and firm resolve.

Our philanthropic resources are exceptionally modest when measured against the depth of the social and economic challenges we tackle. And in a volatile economic climate, philanthropic leaders must be highly disciplined in managing limited resources. The need for strategic and effective philanthropic effort is further heightened by the unprecedented pace of change brought about by the seemingly limitless technological innovation of our time.

Now more than ever, results matter, and how we as philanthropic leaders define, promote and reinforce a commitment to results in the culture of our organizations can profoundly affect whether or not our collective mission to make a difference in people's lives succeeds. A results-focused culture must be predicated on an institution-wide commitment to clearly defined objectives pursued with strategic clarity and supported by dynamic resource allocation. Results can only be defined by end outcomes to the communities we are serving. Process or activity-based results are valuable on an interim basis, but the results that matter are the results that are felt by people in need. How do we understand, in appropriate timeframes, whether we are

making a difference by delivering the results we seek? To truly maximize impact, foundations must work to build results-focused cultures that embed internally the risks and demands faced by our grantees, and demand of ourselves the same level of performance demanded of grantees.

At the Ford Foundation, we continually work to answer these challenges. We make long-term investments, understanding that patient capital and well-reasoned risk are required to chart bold new solutions to complex social problems. Our goals are centered on social justice principles that recognize the inherent dignity of all people. In accordance with these principles, the foundation's work seeks to ensure that social systems and institutions give all people a voice in decisions that affect their lives, and the opportunity to reach their full potential. This is far-reaching and essential work, but it is also, by nature, difficult to evaluate.

Our challenge, then, has been to create a culture focused on results within the Ford Foundation without sacrificing either our ambitious objectives or our commitment to maintaining certain core principles as we pursue those objectives. In building and fostering a results-focused culture, the Ford Foundation is guided by five principles:

1. Create a clear and focused strategic vision.
2. Allocate resources on a differentiated and dynamic basis.
3. Build accountability based on clearly delineated roles and responsibilities.
4. Put a premium on deep and effective listening.
5. Implement a results-focused culture across the entire organization.

Realizing these fundamental principles in an organization's culture is no simple un-

dertaking. It requires comfort with uncertainty and risk, and openness to elements of accountability that at times call for significant shifts in perspective. It demands that we continually ask ourselves, at every stage of our work and at all levels of the organization: Are we making as great a difference as we can for the communities and people we are entrusted with serving? Is there anything we can learn—and change—to achieve the maximum results? And how do we define what constitutes results for the range of work and goals to which we are committed?

CREATE A CLEAR AND FOCUSED STRATEGIC VISION

A culture focused on results begins with a clear understanding of objectives and the development of strategies to achieve them. During my first year as president of the Ford Foundation, the global program team and I went through a comprehensive review of our grantmaking, motivated by the desire to bring focus and clarity to our work. We identified a core set of social justice issues that together constitute a cogent way to be true to the foundation's mission. For each issue we set clearly defined goals and strategies, developed theories of change for achieving them, and designed operating approaches for working day-to-day. The results of this effort can be seen on our website, which provides a visual map of our grantmaking objectives and strategies and directly links them to the nonprofit and nongovernmental efforts we are supporting.

The second phase of the effort was to determine the scale of resources required to achieve each strategy. Focusing our resources meant we increased the average funding of initiatives from just over \$1 million to more than \$10 million each. We brought the same kind of depth and concentration to our staffing, moving from indi-

LUIS A. UBIÑAS is president of the Ford Foundation. Before joining the foundation in 2008, he was a director at McKinsey & Co., leading the firm's media practice on the West Coast.

vidual pursuit of initiatives to a team-based approach. Perhaps most important, we were able to thoughtfully deploy the foundation's non-grant resources, including field leadership, public and private sector partnerships, communications, and convening capabilities. Direct financial support, our traditional staple, became only one part of the portfolio of tools the foundation could mobilize to address a problem or challenge once a strategic vision had been set.

One example of the results that can be achieved by combining non-grant resources with financial resources is JustFilms, a new Ford Foundation effort to advance social justice through the creative lens of emerging and established documentary filmmakers. Although the foundation has been a leading funder of social justice documentaries for decades, supporting landmark films such as *Eyes on the Prize*, this effort is harnessing the power of film to create a national and global dialogue on social justice issues.

Our first step was to identify clear, focused outcomes that could be thoughtfully assessed, such as growing a new cadre of filmmakers, bringing social justice films to market at strategic moments that maximize awareness and social change, helping films find an audience, and connecting new works to a global network of social change makers and movements. Evaluating how we were meeting these objectives required us subsequently to change our strategy, more than tripling our funding from under \$3 million annually to more than \$10 million annually to support our aspirations; establishing a creative collaboration with two leading film organizations, the Sundance Institute and the Tribeca Film Festival; and assigning two dedicated, full-time, high-level and experienced foundation staff members to direct the initiative. In its first year, JustFilms had more films in the 2012 Sundance Film Festival's documentary competition than any other producer, and one of our films won the award for Best Documentary.

ALLOCATE RESOURCES IN A DIFFERENTIATED AND DYNAMIC MANNER

With clear objectives, strategies, and operating approaches in place, the central question we asked was what scale of resources was required for each strategy to effectively achieve the results required. From the start, the foundation allocated



A culture focused on results begins with a clear understanding of objectives and the development of strategies to achieve them.

resources—financial, staff, and other, such as public voice—according to the nature and scale of its objectives. We change that allocation dynamically, on the basis of strategic need. As a result, the level of resources can vary greatly across initiatives and over time. Through ongoing assessments, we engage in a dialogue with program staff about whether our investments should be augmented to achieve significant results, or reduced to allow for future exploration and course correction. This process creates a dynamic internal marketplace of ideas and information that informs the evolution of our grantmaking strategies.

For example, using results-based strategic assessment and adjustment recently allowed us to launch a \$200 million investment to help shift the approach of urban development from one focused on isolated areas of urban need to one focused on regional solutions that link potential workers to centers of

employment and housing. The goal is now to tie economic opportunities to communities in need through better transportation, housing, and zoning policies.

Once initial resources are allocated in a differentiated way, we immediately face the question of how to reallocate resources as initiatives evolve, succeed, and, fail. A results-focused culture must integrate course-correction methods, acknowledging successes and failures and adjusting resources accordingly. This process begins by answering five important questions:

1. Does the initiative or field office have a clearly defined strategy and approach with explicit and achievable objectives that can be documented and evaluated?
2. Can additional funding for a defined period deepen or hasten desired results over time in an initiative or office?

3. Are the funds currently allocated being deployed strategically? Have peripheral activities been eliminated?
4. To what extent has the initiative team successfully engaged other partners, including other foundations, businesses, and government?
5. Is the initiative team strategically using the foundation's non-grant resources—intellectual, convening, and communications—to pursue the initiative's goals?

The answers to these questions help guide decisions about which initiatives are best positioned to deploy additional funds strategically, as well as which ones require additional funds to achieve more significant results. But the answers may also lead to a decision to change course or to end activities.

It is essential to use the same differentiated, dynamic approach in the allocation of staff and other non-grant resources that add value beyond direct financial support. One example is the strategic use of communications—a core resource that is often left out of the results equation. Each strategy must consider if and how the use of communications activities is critical to achieving its desired goal. At the Ford Foundation, we think about the deployment of external communications assets—how we allocate the foundation's voice—very carefully as an integral part of how we work.

Achieving results often pivots on our ability to communicate effectively and strategically about promising solutions to complex challenges that are effective and realistic. An early example came in November 2010, when a critical mass of people gathered in Cancun, Mexico, for the United Nations Climate Change Conference. The Ford Foundation, along with our grantees, saw this convening as an opportunity to raise the profile of community forestry in Mexico—a little known but significant success story in cutting greenhouse gases. A foundation-supported media campaign offered journalists a chance to visit community forestry sites in Mexico, generating important media coverage before and during the conference. To complement this effort, I authored an opinion article, which ran in US and Mexican media, on the need to invest in sustainable forestry programs that respect and promote community stewardship. Using the foundation's public

voice continues to be a critical asset toward achieving results in our work in community forestry and other grantmaking areas.

BUILD IN ACCOUNTABILITY BASED ON CLEARLY DELINEATED ROLES AND RESPONSIBILITIES

Clearly defining roles and responsibilities is central to a results-focused culture. For senior management, this kind of clarity—complementary to clarity in strategy and approach—supports the robust level of oversight and partnership that is vital to effectively align resources with goals. Clarity about responsibilities is an essential precursor to accountability.

At the Ford Foundation, we have expanded

often takes the form of increasing economic rewards. Nonprofit organizations and foundations have a different model but must be no less committed to creating a culture that rewards performance and contribution to forwarding the mission. It is critical to acknowledge high performers who achieve disproportionate results—innovative leaders who foster internal and external innovation, build movements, and establish national and international dialogues as a new norm. Our ability to reward such high performers is no less powerful than in the private sector because we can provide the added resources that these leaders need to deepen their work.

At the same time, any culture that focuses on performance results and standards of ac-

We believe that our early effort to define and instill a culture of results has relevance for other foundations, both large and small.

and defined the role of directors to give them explicit responsibility for initiatives. On average, a director oversees three initiatives, supervising their teams and budgets. Our representatives lead the work in our many regional offices, collaborating with primarily, though not exclusively, New York City-based directors to determine which of our initiatives are most needed in a specific geographic area and to ensure that initiatives are implemented effectively. Although many people are involved in grantmaking, responsibility for outcomes in a particular initiative or geographic area sits squarely with the team of directors and representatives involved.

Our expectations of grantees are high. We demand the best from them, and as foundation leaders we need to demand the best from ourselves as well. Every member of the Ford Foundation's staff, from senior leadership to the newest hire, is expected to recognize that being part of the organization means holding himself or herself to the highest standard in ensuring results. Each staff person must know that, like grantees, she or he is accountable for results.

Tying external results to internal outcomes is central to the development of a results-focused culture. In every organization, high performance can be driven by rewards and incentives. In the for-profit sector this

countability must have an equally strong organizational commitment to staff development. Philanthropic organizations must invest in building the skills and organizational tools needed to evaluate progress for each initiative and for each individual grantmaker—indeed, for every member of the staff, from administrative assistants to the president. Every willing staffer has the potential to be a change leader.

An important caveat, as we think about results, is the need to remain aware of the risky nature of our work. Many of our most important initiatives are extraordinarily demanding, and their outcomes are highly uncertain. We face the challenge of maintaining high standards of accountability in an environment where speculative innovation is encouraged and often necessary. In a results-focused culture, it is critical for leaders to make positive use of noble, well-planned, and well-executed efforts that nonetheless were not able to succeed, recognizing and understanding that failure is often the potent seed of future success.

PUT A PREMIUM ON DEEP AND EFFECTIVE LISTENING

A results-focused culture puts a premium on listening. At the Ford Foundation, we know that our grantees have the deepest

understanding of the issues that affect the people they serve. That is why we support initiatives managed by those living and working closest to where targeted problems are located. We recognize the critical need to maintain a rich dialogue with our grantees and others in the fields and communities in which we work.

The Ford Foundation's worldwide network of offices is designed to help us realize this aspiration by fostering both face-to-face and virtual meetings. In a typical year, in our New York City offices alone, we host more than 20,000 visitors who come together with our teams to work on existing initiatives and to explore new ones. Investing time in listening and learning enables us to make real-time strategic adjustments and to plan highly informed future directions. This emphasis on listening, in turn, enables us to improve how we are working, and demands of ourselves the same excellence we demand of our grantees.

Deep and effective listening is more about shared dialogue than information gathering. Traditional reporting and data collection can sometimes be excessive and ultimately uninformative or misleading, at times yielding adversarial relationships with grantees rather than a constructive collaboration. Excessive data gathering has several costs. First, it is a burden to grantees, creating surplus work for often tightly staffed and financially strapped nonprofits. Second, it undermines quality, because grantees will provide the requested information to meet their grant obligations, but may not have time to supply the insight that is often more valuable than this form of data. Third, it inundates foundation staff with information but may leave them little time to use it effectively. Fourth, it may not provide the information that is actually needed to understand how effective our initiatives and grantmaking are.

Data reporting and design do play an important role in building a culture of results, but just as resources must be strategically targeted, so too must reporting requirements be focused and shaped by certain basic principles. The information and indicators collected must be aligned with program goals and strategies to effectively show whether an activity is on the forecasted path or has deviated from it. From the very beginning, grantees should have a clear sense of what benchmarks of

success are expected of them at each stage of initiative development. Such information and indicators are not exclusively or always quantified or quantifiable. Therefore we must be creative in articulating what kind of information will allow us to make these assessments and subsequent decisions.

The skill of deep listening is vital not only to our individual initiatives, but to our organization as a whole. It is a central part of the culture anchored in our initial strategy-setting effort. At the Ford Foundation, we engaged more than 4,000 grantees and other experts in our strategy-resetting effort in 2008, and we believe every one of them played a role in reshaping and revitalizing our work. Goals, theories of change, and operating approaches are all necessarily imperfect; only by learning from our successes and our mistakes can we build an effective culture of results. Our objective, as we listen to grantees and others, should not be merely to gather data and evaluate others but to learn about ourselves and to evolve.

IMPLEMENT A RESULTS-FOCUSED CULTURE ACROSS THE ENTIRE ORGANIZATION

Leading a foundation is a complex task. At the Ford Foundation, we make more than \$500 million in grants and contribute to change in a range of other ways, from providing individual and team leadership, to bringing voice and convening, to building awareness and partnerships. But excellence in these dimensions does not adequately define a results-focused culture.

Our success in implementing and supporting program initiatives depends on the foundation's outstanding day-to-day operations and exceptional endowment management, and consequently we have worked hard to ensure that our culture of results reaches beyond the program team to include all staff. A philanthropic organization must bring to all of its operations the same level of rigor that it uses to define strategy, allocate resources, gather information, and evaluate results.

More than \$100 million of the foundation's 2011-grantmaking was the direct result of the changes we made in operations and endowment management in 2008 and 2009. For example, shifting \$40 million from internal operating costs to external grantmaking in 2009 enabled us to launch

and fully fund the programs discussed in this article. In addition, careful stewardship of our endowment—including changing our allocation of funds and the management of those allocations—has allowed it to fully recover the investment losses that resulted from the economic downturn. The decrease in our endowment that has occurred since the beginning of the recession is now exclusively the result of our grantmaking activity.

WHAT WE HAVE LEARNED

The Ford Foundation has a complex global footprint and a broad scope. Nonetheless, we believe that our early effort to define and instill a culture of results has relevance for other foundations, both large and small. The lessons we have learned over the last several years and the five principles we have put in place have charted a well-defined course for the foundation.

Although the five principles are now clear, nothing about our endeavor to build a culture of results has been easy. Listening means hearing criticism as well as praise; dynamic resource allocation often means making hard and even painful choices; and a commitment to accountability means taking responsibility for performance organizationally and individually. But when we consider the benefits of this carefully developed approach—a greater ability to help the people we serve—we think they far outweigh the challenges of implementation.

As the president of the Ford Foundation, leading a results-focused culture is a constantly evolving challenge. Very little is static. Strategies change and resources shift as strategic need and results demand. Accountability is maintained, based on individual and team performance. Grantees and others need to be continuously engaged and heard.

A culture focused on results is not always comfortable—it has many of the stresses of a marketplace that demands effectiveness and efficiency. In a time defined by technological change and economic discontinuity, however, we have no choice but to demand the most of ourselves. Cultural transformation is, undeniably, hard work, but the rewards of such an endeavor are great. Any leader trying to achieve strategic results at a foundation or nonprofit should consider the hard-won but invaluable benefits of building a culture that fosters, reinforces, and rewards that goal. ♦

Learning from Silicon Valley

How the Omidyar Network uses a venture capital model to measure and evaluate effectiveness

BY MATT BANNICK & ERIC HALLSTEIN

When visitors step out of the elevator at the Omidyar Network (ON) office in Redwood City, Calif., the first thing they see is a quotation on the wall: “Every person has the power to make a difference.” A belief in the fundamental value and power of each individual shapes nearly every facet of ON, the philanthropic investment firm created and funded in 2004 by eBay founder Pierre Omidyar and his wife, Pam. At ON, we strive to create opportunities for individuals to unlock this power.

ON’s approach to measurement and evaluation is inseparable from our approach to philanthropy. We start with the notion that unlocking the potential of individuals is an indispensable way to improve the world. Given opportunities, people will tap their talents to improve their lives and consequently, the lives of our families and communities, which ultimately benefits society more broadly.

Omidyar Network strives to scale up innovation by applying many of the best practices from the venture capital industry to philanthropy. In rapidly changing markets, we believe the truly game-changing innovations will emerge from entrepreneurs who are empowered to identify changing situations and rapidly adapt their organizations. Our investment model is to invest in and

then help scale up innovative nonprofit and for-profit organizations with the potential to create opportunity for hundreds of thousands or millions of people. Instead of funding others to implement programs related to our strategy, we seek to enable entrepreneurs to identify—and iterate on—their own strategies for creating social impact.

MEASUREMENT AND EVALUATION

ON has tried many approaches to measurement and evaluation over the last few years, and from that experience we are convinced that an organization’s approach to measurement and evaluation must flow from its investment model. ON’s approach is to conduct an up-front, detailed due diligence process to assess an organization’s potential and decide whether to invest. We then develop trusted partnerships with our investees to measure their progress in real time and provide additional, non-financial support as appropriate.

ON provides grants to nonprofit organizations and invests debt and equity capital in for-profit ventures. ON is an impact investor whose bottom line is social value. We take a systems view; we are driven by the problem that needs to be solved and have the flexibility to tailor the highest impact solution using the most appropriate type of capital.

Philanthropy through grants can benefit society in many ways. Grants can help nonprofits provide pure public goods. For example, the free legal tools offered by Creative Commons, which allow content creators to grant copyright permissions to creative work, increase the amount of content available to the public for sharing. Grants can help subsidize goods and services that produce

positive societal outcomes but are undersupplied by the market. For example, Endeavor is helping to transform emerging markets by supporting entrepreneurs. Grants can help disadvantaged populations that are unable to participate in the market, as Landesa and other property rights organizations do. And grants can spur investments in high-risk ventures, as they did with the initial development of the microfinance industry.

Philanthropy through for-profit ventures, in contrast, can benefit society by leveraging the power of markets. When the primary motive is to generate profits, businesses will strive to deliver value in excess of costs and scale up. Customer willingness to pay for a product or service not only indicates value creation, but also serves as a valuable feedback mechanism.

ON has found that grants and for-profit investments can play complementary roles in delivering social impact, even, or maybe particularly, in difficult socioeconomic environments. Socially responsible grant capital can help potentially high-impact innovation gain market traction, and for-profit equity and debt capital can then help that idea scale up.

The phenomenal growth of microfinance is a good example of such complementary roles. Since 2004, ON has invested more than \$100 million in 26 microfinance organizations: 15 nonprofits and 11 for-profits. In the 1980s and 1990s, most microfinance institutions were grant-funded NGOs. As microfinance’s impact and commercial viability became apparent, business investors—many with strong social motivations—invested heavily in commercial microfinance institutions, helping them grow rapidly. According

MATT BANNICK is managing partner at Omidyar Network, where he leads all aspects of the philanthropic investment firm’s strategy and operations. Before joining Omidyar, Bannick was president of PayPal and president of eBay International.

ERIC HALLSTEIN is a vice president at Imprint Capital Markets and a venture partner at CalCEF Clean Energy Angel Fund. He was previously director of investments at Omidyar Network.

to a 2009 Monitor Institute study, the percentage of the world's top 50 microfinance institutions that were for-profit banks increased from 22 percent to 62 percent between 1998 and 2008. In other words, grants sparked and nurtured microfinance, and for-profit capital helped it scale up.

ON takes equity stakes in for-profit companies and typically provides negotiated general operating support to nonprofit organizations. This support gives organizations great flexibility to change their tactics and reallocate their resources in response to new information or changing market conditions. It also enables them to invest in critical overhead functions, such as developing management talent and information technology, which may otherwise not get funded because these functions are not directly related to delivering specific programs.

A grant or investment is only the beginning of ON's relationship with our portfolio organizations. Although vital for growth, scale, and capacity, money alone cannot solve every problem or bring about the positive social impact we seek. Pierre and Pam Omidyar's vision was to create an organization that augments the impact of its financial investments by deploying human capital and leveraging the knowledge, expertise, and innovation of all those involved in solving the world's most challenging problems.

When requested by a portfolio organization, ON's human capital team provides strategy, governance, and leadership assistance tailored to meet the grantee's or investee's needs. The team also creates opportunities for its investees to tap the power of the network by regularly organizing forums for them to exchange information and share best practices.

SELECTION CRITERIA AND DUE DILIGENCE

ON's investment selection criteria and process provide a foundation for measurement and evaluation of our investments. To make investment decisions, we consider factors such as innovation, scalability, and viability of business or revenue models. ON encourages nonprofit investees to develop earned-income streams whenever possible. Earned income is often a powerful tool for ensuring that a nonprofit focuses on understanding its constituents (customers); benefits from clear, market-based signals about the value of its



product or service; and delivers products and services with values exceeding costs.

Like traditional venture capital firms, ON performs extensive due diligence before investing in an organization. Although traditional approaches to measurement and evaluation do not view due diligence as evaluation, organization-wide, up-front due diligence is central to our approach. In considering a funding investment, we carefully assess the organization's management team, operational planning and governance, target market and competition, technology, product or service, and financials. We meet with the organization's management team; interview customers, key channel partners, and board members; read industry and peer-reviewed reports; and analyze company financial and operating data.

The due diligence process is primarily about whether to invest, but it also affords an opportunity to explore strategically important issues that ultimately help ON support an organization. Detailed discussions with an organization's management team help to ensure strategic alignment between ON and the investee. During this process the management team and ON agree on the metrics that will be used to evaluate the

organization's progress. Due diligence also helps ON determine how best to use its human capital to support an organization. Thus the diligence process has a direct bearing on all subsequent measurement and evaluation.

REACH AND ENGAGEMENT

ON, like all philanthropic funders, considers social impact the single most important investment criterion. Although our specific metrics vary across sectors and organizations, we have developed two overarching metrics that help us understand the social impact of portfolio organizations: *reach* and *engagement*. Reach is a measure of how many individuals are touched by a product or service. Engagement is a measure of the depth of that interaction.

Take for example the nonprofit Wikimedia Foundation, which operates Wikipedia. Wikipedia engages hundreds of millions of people in creating educational content, sharing information, and learning online. In 2010, a measure of Wikipedia's reach was the roughly 400 million different people who visited its website. A measure of engagement was the 13 minutes that an average visitor spent visiting Wikipedia every month. Other measures of reach and engagement

for Wikipedia include the number of active contributors of content and the level of their activity on the site.

Neither the reach nor engagement metric directly measures Wikipedia's social impact, however. To determine impact we would need to understand the extent to which a product or service affects the end user. In this case, a potential impact metric for Wikipedia would be the improvement in students' test scores after they use the site. Accurate estimation of Wikipedia's impact would require an experimental or quasi-experimental approach, typically a randomized trial in the absence of a natural experiment, and might take months or even years to perform.

As an alternative to the time-consuming, costly, and complicated challenge of directly measuring impact, we often measure outputs—reach and engagement—as proxies for impact. Outputs are units of production, which often can be readily measured by an organization in the normal course of business.

Our use of a very limited number of easily collected metrics is a departure from some of our early evaluation frameworks in which we asked entrepreneurs to track as many as 18 different metrics. The earlier approaches not only were more costly to implement, but had less impact on investees' success because neither they nor ON was able to prioritize appropriately.

We have learned that an organization's social impact is best estimated using a combination of reach and engagement metrics tailored to the industry, the organization's business model, and its growth stage. Some organizations, such as Wikimedia, have wide reach and relatively limited personal engagement. Other organizations, such as Endeavor, have modest reach and deep engagement. Selection of appropriate reach and engagement metrics can be difficult and nuanced.

SCALABILITY AND SOCIAL IMPACT

ON invests in highly scalable solutions with the potential to reach millions of people, and in which costs per unit generally decrease as the volume produced increases. Because organizations need a reliable revenue source to achieve scale, financial sustainability is a focus of our evaluation process—but there are some important differences in how we think about financial sustainability in for-profits and nonprofits.

For ON to consider investing in a for-

profit company, the critical first filter is its potential for positive social impact, and the second is its ability to profitably and quickly scale up. We look at financial metrics that are standard for a profit and loss statement, such as revenue, operating margin, and net income. In particular, operating income is a leading indicator of impact because it demonstrates a company's ability to deliver value to customers above cost.

In the for-profit organizations that ON funds, financial metrics are indicators of financial sustainability *and* social impact. A company's revenue and margins are driven by the number of customers served and how highly those customers value the product, both of which provide a quantitative measure of the opportunities a company is creating.

To supplement standard profit-and-loss measures, ON uses metrics captured by the company—typically, key performance indicators that a for-profit company tracks to understand the underlying drivers of its financial performance. At Bridge International Academies, a Kenyan startup that operates a franchise-like network of low-cost, for-profit private schools, these metrics include the number of schools, the number of students and trained teachers, academic performance, and the turnover rate of students and teachers.

REVENUE STREAMS

In nonprofit organizations, social impact is the most important evaluation criterion, and we spend considerable time with investees to define reach and engagement metrics that measure it. We also encourage and work with nonprofit investees to develop diversified sources of revenue as well as to generate earned income and other revenue sources to increase their financial sustainability.

Earned income is valuable to nonprofits not only because it provides revenue to cover operating costs, but also because it is a conduit for clear, rapid, and market-based feedback about whether the organization is providing a highly valued product or service. In this sense, not all revenue is equal. We view revenue tied to an organization's products and services to be the highest value.

Similar to revenue and profit in for-profit organizations, market-based signals tied to products and services can be a powerful mechanism for getting nonprofit manage-

ment teams to prioritize the activities of highest value to the end user. Market-based feedback also is likely to be more useful and valuable for both ON and investees because it informs real-time decisions without waiting for infrequent formal reviews.

For example, consider GuideStar, a nonprofit in which ON first invested in 2007. GuideStar offers information on 1.8 million US nonprofits and private foundations and had more than 10 million website visitors in 2010. Although 98 percent of GuideStar data uses are free, users pay a fee for more sophisticated tools and services. Fee revenue covered 90 percent of GuideStar's operating costs in 2010. Charging fees for premium services has given GuideStar resources to grow quickly and enabled its management team to focus on delivering high-quality services to paying customers rather than raising money through donations. GuideStar now uses foundation grants primarily for special projects and opportunities.

Several of ON's portfolio organizations have developed innovative donation models that are a twist on the earned-income model. For example, DonorsChoose.org matches donors with school classroom projects in need of support. When donors support a specific project, DonorsChoose.org invites them to contribute a small amount toward the costs of running the organization. In fiscal 2011, DonorsChoose.org covered a remarkable 102 percent of its operating expenses through this optional fee. Voluntary donor support of operating fees is a type of immediate feedback indicating the value the donor places on a product or service.

COST-TO-SERVE

Cost-to-serve—the cost of providing a product or service to a single customer—is another metric that we typically track in for-profit and nonprofit investees. Cost-to-serve is an important measure of productivity—specifically, how efficiently an organization is delivering a particular product or service. In general, we find that despite its potential value for driving financially sustainable growth, few nonprofits focus on reducing cost-to-serve.

We calculate the cost-to-serve as the quotient of reach and the trailing 12 months of operating expenses. Cost-to-serve can vary significantly across organizations. For example, we estimate that in 2010, Wikimedia's

cost-to-serve each visit to its website was \$0.02, whereas DonorsChoose.org's cost-to-serve was \$5.23, or roughly 250 times greater.

Cost-to-serve can offer some insight when comparing two businesses with similar models in the same sector, but is less useful when comparing across sectors. Typically, we look at an organization's cost-to-serve to highlight opportunities for efficiency or asset productivity improvements. DonorsChoose.org, for example, is working on reducing its cost-to-serve by lowering new user acquisition costs and increasing the donor retention rate. Bridge International Academies—a new organization with relatively high startup-related overhead—has high costs relative to the number of students served. As Bridge opens more schools and benefits from the resulting economies of scale, we expect the cost-to-serve each student to drop dramatically.

COLLABORATIVE EVALUATION

A successful relationship between ON and each organization is highly collaborative and based on trust. Collaboration starts with the due diligence process, through which ON evaluates the leadership team's ability and commitment to practice exemplary governance, operational efficiency, transparency, and disciplined financial planning. This process builds relationships between ON personnel and the investee's management team, establishing the basis for a positive, long-term relationship.

The spirit of collaboration and trust carries over to the measurement and evaluation process. Our investment professionals work closely with each investee to develop milestones jointly and then measure the investee's progress against them. For-profit companies usually track the desired reach and engagement metrics during the normal course of business. For some consumer Internet companies, for example, the number of unique visitors to a site (reach) and number of page views per visit (engagement) are critical revenue drivers. The organization's performance relative to established milestones is an important criterion for ON when we consider subsequent investments, but specific impact metrics typically are not captured in any agreement between ON and a for-profit company.

In nonprofits, ON first engages the leadership team in a dialogue about the organization's potential social impact, and then

works with the team to specify appropriate, related reach and engagement metrics that are codified in the grant agreement. These milestones provide a mechanism to both hold the organization accountable for achieving its agreed-upon objectives and identify when it might need additional human capital support from ON to be successful.

ON's investment professionals regularly review each investee's performance against these metrics. ON holds board seats or board observer rights in about 40 percent of our portfolio organizations, and metrics are usually reported by the management team at board meetings. If ON does not hold a board seat, the grant agreement typically requires the investee to provide quarterly reports of its progress against the metrics.

To supplement ongoing performance tracking and comprehensively evaluate our investees' progress, ON holds an annual, internal portfolio review. The review covers the performance of each portfolio organization and the overall performance of all organizations by sector.

ON's human capital contributions keep us in frequent contact with our portfolio organizations' management teams. The close professional relationships with investees and the tight integration through these governance roles engender a sense of shared commitment between ON and each investee, building a foundation for effective, real-time evaluation and response. The frequent contact creates a continuous flow of information between ON and the investee, providing a substantial context for important decisions. Equally important, ON's investment professionals have the added benefit of being able to look broadly across the entire market and competitive landscape, and can highlight specific issues that a more narrowly focused management team might miss.

WHAT WE'VE LEARNED

ON's investment model fundamentally shapes how we approach measurement and evaluation. Developing and applying ON's approach has taught us to:

- **Foster partnerships.** Trust-based partnership is the key to establishing a productive approach to measurement and evaluation. ON fosters partnerships by listening carefully to our investees' needs from the beginning of the due dil-

igence process and investing heavily in our human capital function.

- **Exercise judgment.** Data is rarely a substitute for good judgment. Only so much can be measured, much that matters cannot be measured, and any amount of data alone cannot provide all the answers. ON invests in talented management teams and works with them to strike the right balance between the robustness and complexity of a company's approach to evaluation.
- **Be flexible.** Markets, competitors, and companies can change rapidly. What is important to measure today may not be meaningful tomorrow, so we cannot rely on static measurements in a dynamic environment.
- **Embrace feedback.** We manage what we measure. ON takes the time to understand what the data says about a company's product or service. We then help our investees use that information to improve their product or service to deliver even greater value to customers, who will ultimately determine their success or failure.

At ON we continue to discuss and debate many topics related to how we approach measurement and evaluation, including the appropriate role for independent impact assessments, the optimal resource allocation across sectors and companies, the best approach to meaningful information and data sharing across funders, and whether to evaluate ON's contribution to helping shift sectors.

Our view is that an organization's approach to measurement and evaluation flows directly from its investment model. ON regularly re-evaluates whether its investees are having a positive impact on the sector and whether other investments are necessary to break bottlenecks. ON's approach to measurement and evaluation thus remains a work in progress. We continue to iterate and innovate, jointly with our colleagues in the impact investing sector, with the objective of investing in and scaling organizations that have the ability to create opportunity for millions of people. ♦

The authors wish to thank Tie Kim and Gowri Pai for research assistance; Jeff Bradach, Suzanne Lippert, Paula Goldman, Annis Steiner, and Greg Pershall for editorial comments; and Arjuna Costa, Chris Bishko, Amy Klement, Bill Barmeier, and Todor Tashiev for valuable discussions.

Shared Outcomes

How the Rockefeller Foundation is approaching evaluation with developing country partners

BY JUDITH RODIN & NANCY MACPHERSON

Across every sector of society, decision makers are struggling with the complexity and velocity of change in an increasingly interdependent world. The context of decision-making has evolved, and in many cases has been altered in revolutionary ways. In the decade ahead our lives will be more intensely shaped by transformative forces, including economic, environmental, geopolitical, societal, and technological seismic shifts.¹

As part of its response to this global dynamism, the Rockefeller Foundation has translated its 1913 mission of promoting the well-being of humanity into two overarching goals: expanding opportunity through more equitable growth, and strengthening resilience to acute crises and chronic stresses, whether man-made or ecological. Our vision is a world in which globalization's benefits are more widely shared and the inevitable challenges that accompany a world that is fast changing, diverse, and complex are more easily weathered.

The Rockefeller Foundation structures its work around time-bound cross-sectoral initiatives that seek innovative solutions and support enabling environments to bring about change. The foundation's structure reflects its view that today's problems and solutions are multi-dimensional in scope and nature, and that they require multi-disciplinary responses at the intersection of fields.

Just as the Rockefeller Foundation's

JUDITH RODIN is president of the Rockefeller Foundation. Before joining the foundation, she was president of the University of Pennsylvania, and preceding that, provost of Yale University. Rodin is on the board of Citigroup, Comcast, and AMR.

NANCY MACPHERSON is managing director, evaluation, at the Rockefeller Foundation. Before joining the foundation she spent 25 years in development evaluation for international nonprofit organizations in Asia and Africa and for the United Nations.

approach to philanthropy has evolved, so too has its approach to evaluation. With its mission to improve the well-being of humankind, its focus on impact, and much of its grantmaking in developing countries, the Rockefeller Foundation is committed to evaluation practices that are rigorous, innovative, inclusive of stakeholders' voices, and appropriate to the contexts in which the foundation works. This article discusses how the Rockefeller Foundation integrates the views of developing-region

Development Philanthropy in a Changing World, Robert Picciotto, former vice president at the World Bank and now professor at King's College London, squarely tackled the role of evaluation in philanthropy and development. "The changing context and thinking on development has profound implications for development evaluation itself, and for the contribution evaluation can bring to the empowerment of people; and the effectiveness of development interventions by national governments and

The value of evaluation must ultimately be judged by its usefulness in helping to improve outcomes for target beneficiaries.

evaluators into its evaluation approaches, and highlights five key strategies:

1. Engaging stakeholders to develop shared outcomes.
2. Expanding capacity through use of non-staff monitoring and evaluation specialists to partner with grantees.
3. Sharing knowledge through learning forums and communities of practice.
4. Strengthening developing country evaluation practice and ownership of results.
5. Developing innovative methods and approaches to evaluation and learning.

RETHINKING, RESHAPING, AND REFORMING EVALUATION

In November 2011, the Rockefeller Foundation brought together leaders from philanthropy and development at the Future of Philanthropy and Development forum, held at its conference center in Bellagio, Italy. In one of the keynote papers, *Evaluating*

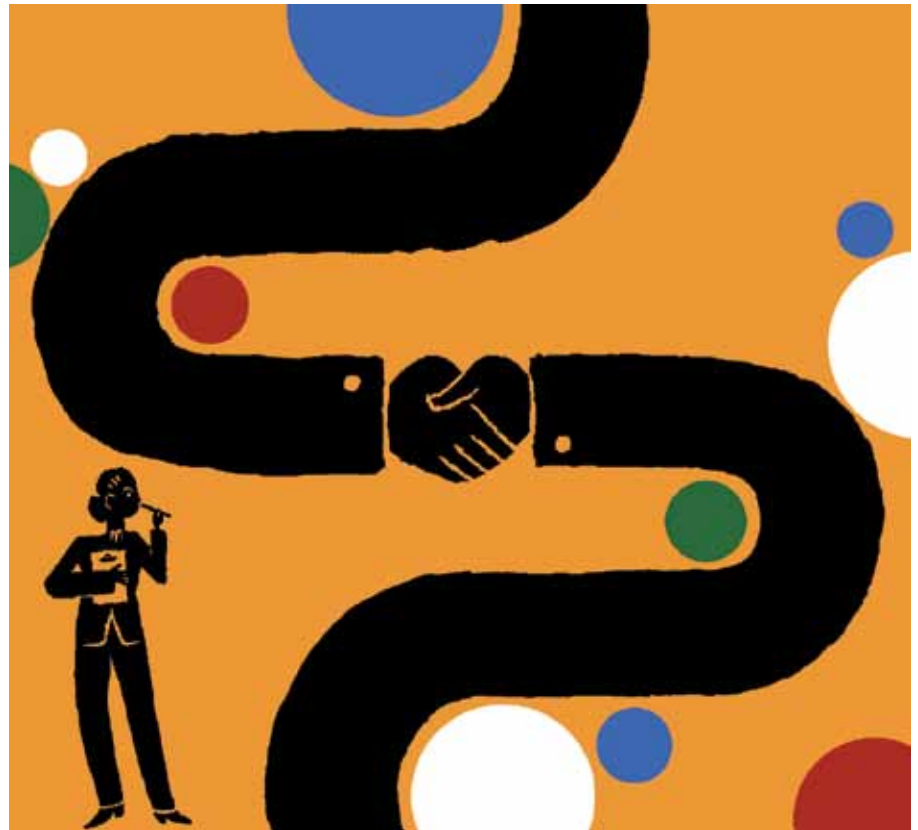
international partners and, increasingly, by non-state actors—foundations, philanthropists, and agencies that promote investing for impact."

Picciotto continued, "Pressing human needs are not being met by an official aid system that is short of resources, catering to multiple interests, and hobbled by massive coordination problems. By contrast, private giving for development is growing and has proven nimbler and more results-oriented than official aid. However, the philanthropic enterprise will not fulfill its potential unless it identifies and taps into its distinctive comparative advantage and coordinates its interventions with other development actors; embeds evaluation in its processes to achieve operational relevance, effectiveness, and efficiency; and demonstrates that it is accountable and responsive to its diverse stakeholders."

Developing country evaluation leaders have also articulated the need for a new approach to evaluation and the role it plays in

improving the wellbeing of humankind—in particular, the lives of the poor and vulnerable in developing countries. At the January 2012 gathering of the Africa Evaluation Association’s biannual conference in Accra, Ghana, African evaluation leaders and policy makers highlighted five steps foundations and development agencies must take if they aspire to play a meaningful role in social transformation.

1. Broaden the inclusion of key stakeholders in evaluation. Only when the voices of those whose lives we seek to improve are heard, respected, and internalized, will we be able to effectively evaluate what success should look like for the people we are most concerned about. Foundations and agencies need to take practical steps to include key stakeholders in the design of, conduct of, and learning from evaluation.
2. Regard evaluative knowledge as a public good and share it widely. Learning with our partners and stakeholders about what works and what doesn’t work should be seen as a global public good not limited to boards and program teams, but shared widely with grantees, partners, and peers.
3. Address evaluation asymmetries between developed and developing regions. The majority of human and financial resources for evaluation emanate from agencies and foundations based in the developed world. With evaluators from developing countries playing a minor role, if any, many of them do not get sufficient experience to move into leadership roles. Mentoring, coaching, and training can strengthen the role, capacity, and resources of developing-country evaluators so that they can play key roles in conducting and using evaluation results for social transformation and accountability in their own countries.
4. Broaden the objects of evaluation to learn more beyond the individual grant or project to a more strategic assessment of portfolios of investments, policy change, new financing mechanisms, and sector-wide approaches that tell us more about what works and what does not in different contexts.



Framing evaluation to take into account the drivers of unsustainability and causes of the challenge being addressed provides greater learning than narrower evaluations that focus only on the funder’s specific intervention.

5. Invest in the development and application of innovative new methods and tools for evaluation and monitoring that reflect multidisciplinary and systems approaches to problems and complexity; invest in methods that assess network effectiveness and policy change; and use and adapt new technology to enable stakeholders to provide close to real-time data and feedback.

THE ROCKEFELLER FOUNDATION’S APPROACH

With its long history of supporting developing country institutions, the Rockefeller Foundation has responded to the call to action from developing-region evaluators by adopting the following approaches to planning, monitoring, and evaluating its work.

Shared Outcomes | An important underpinning of the Rockefeller Foundation’s initiative-based approach is a fundamental recognition that the world’s greatest chal-

lenges can’t be solved alone. These challenges involve a complex mix of actors that are often globally interdependent across sectors and geographies. Networks, alliances, and coalitions of diverse stakeholders from governments, foundations, civil society, and business are increasingly seen by the foundation as a more powerful way to mobilize the vast range of resources and actions required to bring about sustained and transformational change on a significant scale.

Increasingly, the Rockefeller Foundation brings together grantees and partners from developed and developing countries to establish a common vision of the problem, outcomes, and indicators for success. Grantee agreements now include reference to the common vision of results and shared outcomes to which the grantee contributes, and foundation teams are expected to manage portfolios of grants and relationships with grantees towards that common vision. This shared-outcomes approach forms the basis for ongoing monitoring, evaluation, and reporting, and for learning dialogues with grantees and partners. (See “Shared Results Framework” on p. 14.)

Monitoring and Evaluation | Most foundations have capacity limitations on

the amount of time that can be devoted to monitoring and learning with grantees and partners, visiting field projects, and working collaboratively—activities that we know contribute to greater collaborative learning and effective relationships. Recognizing these limitations, the foundation awards grants to monitoring and evaluation (M&E) groups and specialists in developing and developed countries who act as monitoring partners, or what we call “critical friends,”² throughout the life of initiatives (typically, a five- to six-year period). They work with grantees to identify key learning questions, help to set up monitoring systems, and provide support in analyzing monitoring data. The most sig-

nificant feature of the critical friends is that they build trust with grantees and partners to ask tough evaluative questions, and they support grantees in seeking and using feedback to make improvements throughout the life of the initiative. Periodic evaluations are conducted by independent teams to provide an objective assessment of progress toward outcomes and impact.

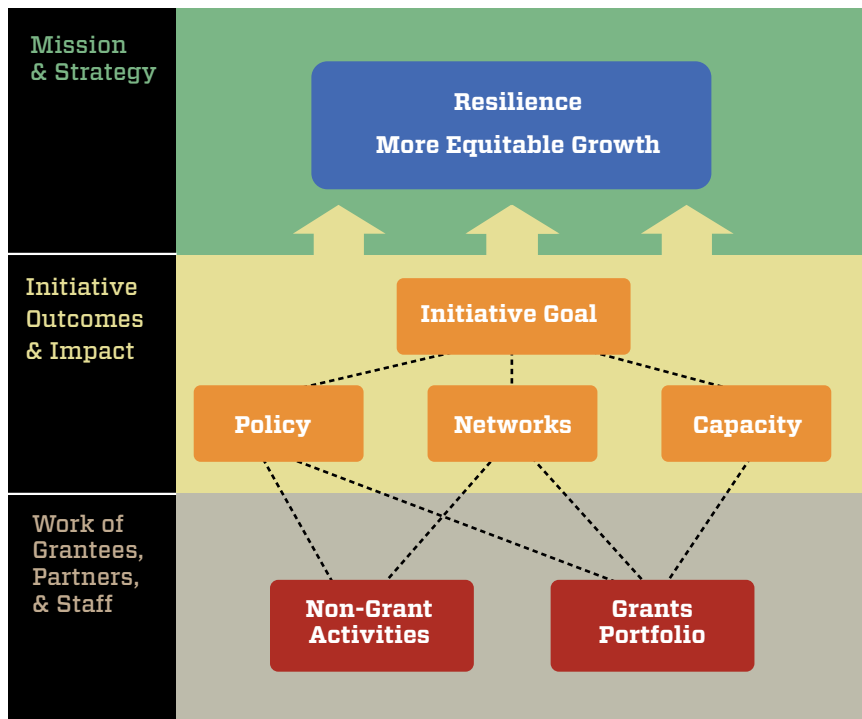
For example, the India-based nonprofit Participatory Research in Asia, in collaboration with the Ghana-based Institute for Policy Alternatives, works alongside Shack/Slum Dwellers International (SDI) which directly represents millions of urban poor slum dwellers in 33 countries. The aim of this critical

friend partnership is to strengthen the participatory learning, monitoring, and evaluation systems and abilities of the urban poor networks to better capture and systematize learning and strengthen accountability with the goal of empowering the urban poor to achieve wider positive impacts. The critical friend role underpins a belief that federations of the urban poor are capable of changing their own situation for the better. As a result of this partnership, SDI has strengthened its ability to democratize learning, monitoring, and evaluation—continuing to place the tools, responsibility, and ability for change in the hands of its members.

Learning Forums and Communities of Practice | Most of the Rockefeller Foundation’s initiative teams convene grantees and partners annually to review progress, highlight lessons and challenges, celebrate successes, and identify improvements needed. Through these forums grantees learn from others in the field, meet new resource people, and adjust their strategies going forward. Although this practice does not guarantee impact, it increases the likelihood of it by creating a greater sense of ownership and shared outcomes, and it increases leverage by connecting grantees with new resource people, funders, and mentors. Increasingly, M&E grantees produce high-quality knowledge products as a public evaluation good to highlight what works, what does not, for whom, and under what conditions. Our aim is to establish with grantees a body of collaborative knowledge, shared lessons, and a culture that values evaluation as a resource for learning as well as for accountability.

For example, the Rockefeller Foundation has aligned with the South East Asia Community of Practice in Evaluating Climate Change Resilience (SEA Change), facilitated by the nonprofit organization Pact, to work on urban climate change resilience in ten Asian cities. This community of practice brings together evaluators, program managers, grantees, and policy makers concerned with learning what works in interventions aimed at adapting and building resilience to the effects of climate change and extreme weather events in Southeast Asia. Resources and lessons are shared through online learning, onsite convening of SEA Change participants, and coaching, mentoring, and training provided by members of the community of practice through-

Shared Results Framework



This figure illustrates the framework around which Rockefeller Foundation staff, grantees, and partners develop a common vision of the results and impact that they seek to achieve collectively. The top frame represents the mission and strategy of the foundation—promoting the well-being of humanity in two overarching goals: expanding opportunity through more equitable growth, and strengthening resilience. The middle frame represents the medium-term outcomes that the foundation seeks to achieve during the life of the initiative (these change from initiative to initiative). The lower frame represents the work that grantees, partners, and staff do in their individual organizations to collectively bring about outcomes and ultimately improve the lives of beneficiaries. These shared results frameworks anchor the ongoing dialogue with grantees about progress toward achieving this vision and their contribution to the shared outcomes. It also serves as a framework for managing portfolios of grants and monitoring changes during the life of the initiative.

out the countries of Southeast Asia.

Addressing Asymmetries | The Rockefeller Foundation is supporting the formation and strengthening of regional development-country networks and the first-ever regional institutions to train, coach, and mentor evaluators, and to partner with evaluators from other regions. Through these platforms and networks, the foundation aims to help rebalance the asymmetries of choice and opportunity for developing-country evaluators to control the evaluation process in their own localities and to improve the quality of evaluation by partnering with evaluation leaders globally.

One example of this is the African Evaluation Association (AfrEA), a pan-Africa umbrella organization comprising more than 25 national M&E associations in Africa, and a resource for individuals in countries where national evaluation bodies do not exist. AfrEA, which has more than 1,000 evaluators from all regions of Africa, receives Rockefeller Foundation funding to enable the formalization of its organizational, operational, and management structure, and to build communities of practice among its membership to tackle the most pressing evaluation challenges on the continent.

The Centers for Learning on Evaluation and Results, located in Africa and Asia, are another example of an effort aimed at addressing asymmetries in evaluation in developing countries. Together with a consortium of funders committed to building developing country capacity for taking charge of the evaluation agenda in their regions, the foundation is supporting regional centers³ in East and West Africa and South Asia to strengthen their skills, networks, and experience in monitoring and evaluation and results-based management capacity of public, private, and civil society development in the global south.

New Methods and Approaches | Traditional evaluation methods and approaches to learning, accountability, and feedback have not kept pace with the advances in technology and social media. The majority of evaluation practice is still largely paper-based despite great strides in technology, interactive web-based platforms, and multimedia tools that make real-time feedback from grantees and beneficiaries possible and accessible. The Rockefeller Foundation and its partners learned a great deal from

Ushahidi, an open-source crowdsourcing project that allows users to send crisis information via mobile devices to map reports of violence or suffering. Inspired by the potential of these kinds of tools to democratize evaluation information, increase transparency, and lower the barriers for individuals to share information and stories, the foundation is supporting a number of innovative approaches to evaluation.

One example is GlobalGiving's Story Telling project, an innovative way to gather local feedback from people in developing countries and to share it with communities, implementing organizations, and donors to create real-time feedback. With support from

in at least four ways. We must:

1. Embrace a broader set of voices in framing our approaches to evaluation.
2. View collaboration and partnerships between developed and developing areas as mutually beneficial toward a common goal of expanding and sharing evaluation knowledge as a public good aimed at achieving better development outcomes.
3. Recognize the need to address issues of accountability, transparency, ethics, culture, and independence.
4. Address asymmetries in individual and institutional capacities for under-

We are privileged to work in an expanding field in which our evaluation findings can change lives for the better.

the Rockefeller Foundation, GlobalGiving successfully deployed a network of people in Kenya, Uganda, and Tanzania that has generated more than 20,000 tagged narratives from thousands of people. Some of GlobalGiving's partners are deriving actionable intelligence from these stories, and GlobalGiving is discovering patterns in the stories that inform its own operational and strategic decision-making processes.

Another example is BetterEvaluation, an online interactive community of evaluation practice developed by the Royal Melbourne Institute of Technology in partnership with the Institutional Learning and Change Initiative and the Overseas Development Institute, with support from the Rockefeller Foundation and Pact. BetterEvaluation provides advice, online support, and good practice examples to evaluators in developing and developed countries.

RESHAPING DEVELOPMENT EVALUATION

Philanthropists and development practitioners have a golden opportunity to join together with grantees and partners in developing countries to reshape evaluation to better respond to global change and to serve our missions and goals more effectively. To do this, we must be prepared to rethink and reshape our evaluation practice

taking, driving, and owning evaluation in developing regions by promoting opportunities for professional excellence, networks, and sustained global partnerships in the discipline of development evaluation.

The value of evaluation must ultimately be judged by its usefulness in helping to improve outcomes for target beneficiaries. The quest for impact is currently in the spotlight among foundations and development agencies as we seek collectively to maximize the positive benefits of our resources. We are privileged to work in an expanding field in which our evaluation findings can change lives for the better. Together with our peers, partners, and grantees, we can and should rethink, reshape, and reform the practice of evaluation to better meet that challenge. ♦

- 1 For more discussion of this topic, see *Global Risks 2012*, Seventh Edition, World Economic Forum, 2012.
- 2 The term "critical friend" refers to a partner that builds trust and engenders a reflective evaluative culture that is able to provide both negative and positive feedback to the grantee in a supportive way throughout the life of the work.
- 3 In South Asia The Jameel Poverty Action Lab South Asia at the Institute for Financial Management and Research, India, with the Center for Economic Research in Pakistan; In Africa the University of Witwatersrand with the Kenya Institute of Administration and Ghana Institute of Management and Public Administration.

Risky Business

How The William and Flora Hewlett Foundation approaches high-risk philanthropic ventures

BY PAUL BREST

The William and Flora Hewlett Foundation's Guiding Principles state that the foundation "focuses on the most serious problems facing society where risk capital, responsibly invested, may make a difference over time." The foundation's grantmaking strategies, with goals ranging from mitigating climate change to reforming California's fiscal policies, reflect the board's and staff's considerable tolerance for risk. This article outlines our framework for investing in strategies where the likelihood of success is small and often difficult to quantify. Let me begin with a little allegory.

You come across a small, determined group of villagers pushing a heavy boulder up a steep and craggy glacier. The boulder is threatening their homes, and they are trying to get it to the top and then roll it into an uninhabited valley on the other side. The glacier is shrouded in fog, but you can discern that there are many peaks, valleys, and crevices on the way to the top. It isn't evident that the group is up to the task—sometimes it's one step forward and two back—and every once in a while, an opposing group tries to push the boulder back down the slope. The villagers ask you to pitch in. You are persuaded that the mission is important, but you don't know their likelihood of success, other than that it is small. Before deciding whether to join, you would like to know whether your contribution will make a difference, but this is difficult to predict.

The metaphor of pushing a boulder up a glacier describes a variety of risky philanthropic strategies. Advocacy to change

public policy is paradigmatic. Other examples include public interest litigation; second-track diplomacy (such as informal meetings of Israelis and Palestinians to get productive peace talks under way); and support for yet-untested innovations in service delivery, technology, and medicine (such as an AIDS vaccine). In many of these cases the outcomes are subject to what economists term "uncertainty" rather than "risk," because the likelihood of success is not quantifiable—at least not within any satisfactory margins of error.

Moving from allegories to philanthropy, I'll use two hypothetical examples—a risky advocacy strategy and, for contrast, a relatively non-risky service delivery program. The risky strategy is an environmental organization's campaign to persuade a public utilities commission to adopt renewable portfolio standards, which require a certain amount of electricity to be generated from water, wind, or solar power. The non-risky example is a program to reduce teen pregnancies through a well-evaluated peer counseling program.

Every philanthropic grant has an intended outcome, or goal, such as the use of fewer hydrocarbons in generating electricity or reducing teen pregnancies. Philanthropists are interested in outcomes from three points of view: *ex-ante*—how likely the strategy is to have its intended outcome; *in progress*—whether the strategy is on course toward that outcome; and *ex-post*—whether the strategy actually achieved its intended outcome. As I will discuss later, philanthropists are ultimately concerned with impact rather than outcomes—with whether the activities they support actually cause or contribute to the outcomes. But it is useful to begin with outcomes, which are necessary, though not sufficient, for achieving impact.

EX-ANTE

The Theory of Change | Before investing in a particular venture, a philanthropist needs to understand how and why it is likely to achieve its intended outcome. Making that assessment requires a theory of change—an empirically based causal theory that links activities to outcomes. It is *causal* because it holds that if you do a particular activity, then a specific outcome is likely to happen—if you press on the gas pedal, the car will move. It is *empirical* because it purports to describe the way the world actually works. The causal theory may be based on an understanding of the underlying mechanism (the gas pedal is connected to the carburetor ...) or observation (every time I've seen someone press on the gas pedal, the car has moved). Although a theory of change is based on the analysis of the causal links of past interventions, it provides a basis for predicting the effects of future interventions as well.

A teen pregnancy prevention program might be based on any number of different theories of change—for example, that one can reduce pregnancies by counseling abstinence, or by educating teens about how to use contraceptives and making them available. The theory of change might posit that the best counselor is a peer, a religious leader, or someone with medical expertise in contraception.

What makes assessing the likelihood of succeeding in direct service interventions relatively easy is that their validity can be assessed by well-established methods of evaluation. The gold standard for evaluation is randomized controlled trials (RCTs), in which the target group of teenagers is randomly assigned either to a group receiving the counseling (the treatment group) or to a group that does not receive the intervention (the control group), and

PAUL BREST is president of The William and Flora Hewlett Foundation. Before joining the foundation in 2000, he was a professor at Stanford Law School, serving as dean from 1987 to 1999.

the outcomes (pregnancy rates) are compared.¹ Evaluators are interested in two fundamental questions: the magnitude of the effect of the intervention (what percentage of participants avoid pregnancy as a result of the program?) and whether the difference between the treatment and control group is statistically significant.

It turns out that although abstinence-only education has no effect, some programs that include information on contraception can make a difference.² For our hypothetical example, let's assume that in a high-quality study of a program involving thousands of girls, only 4 percent of those in the treatment group become pregnant, compared to 7 percent of non-participants—a 43 percent improvement, which is an extraordinarily good outcome for any social intervention. Because the likelihood of achieving the benefit is not only determinate but high, the teen pregnancy program is not risky from the philanthropist's perspective.

The effort to advocate for renewable portfolio standards also is supported by a theory of change, in this case from the domain of political science. In its most general sense, the theory links the organization's advocacy activities to the intended outcome of persuading the decision makers to adopt the regulation. More specifically, the theory of change specifies the conditions under which advocacy will be effective—and the paths to effectiveness—based on what motivates the decision makers and how to manipulate the (often indirect) levers to affect their behavior.

But this theory of change is not testable through methods such as RCTs, which rely on the comparison of large samples of very similar subjects. The political theory of change is a set of generalizations based on the observation of a number of unique events—advocacy concerning different issues in different contexts. Moreover, the inputs and outputs of such events are often ambiguous. As Steven Teles and Mark Schmitt write in “The Elusive Craft of Evaluating Advocacy” (summer 2011 issue of *Stanford Social Innovation Review*), “Sometimes political outputs are reasonably proximate and traceable to inputs, but sometimes results are quite indirectly related and take decades to come to fruition.”

Even when one can have some sense of the likelihood of success of an advocacy strategy, the margins of error are typically



so large as to put the enterprise in the domain of uncertainty rather than quantifiable risk. (Predicting the outcome of an advocacy strategy is somewhat analogous to predicting the counseling program's success in preventing one particular participant's pregnancy.)

The Logic Model | The theory of change for an intervention provides the basis for its logic model, which describes (among other things) the activities that an organization must undertake to achieve the desired outcome. For example, the logic model for the pregnancy prevention program involves the logistics of counseling. It includes activities such as recruiting the target group of teenagers, recruiting and training counselors, setting up counseling sessions, and ensuring that the counselors provide the requisite information and support. Although there is plenty of room for variation—for example, in the substance and dynamics of the counseling sessions—the logic model is essentially a cookbook recipe.

By contrast, an advocacy strategy seldom has a detailed recipe—only a number of dos and don'ts, whens and hows, from the accumulated knowledge of master chefs.³ For example, the strategy for achieving re-

newable portfolio standards might involve identifying the views and motivations of the public utility commissioners and approaching each one individually or persuading a constituent to approach them.

As Teles and Schmitt write: “[Advocacy is] inherently political, and it's the nature of politics that events evolve rapidly and in a nonlinear fashion, so an effort that doesn't seem to be working might suddenly bear fruit, or one that seemed to be on track can suddenly lose momentum. ... [T]actics that may have worked in one instance are not necessarily more likely to succeed in another. What matters is whether advocates can choose the tactic appropriate to a particular conflict and adapt to the shifting moves of the opposition. ... [S]uccessful advocates know that such plans are at best loose guides, and the path to change may branch off in any number of directions. ... Successful advocacy efforts are characterized not by their ability to proceed along a predefined track, but by their capacity to adapt to changing circumstances.”

Predicting a Program's Value | From the strength of the evidence underlying the theory of change and the details of the logic model, one can predict (with more

or less confidence) the value of a philanthropic investment in a particular program or strategy.

There are two related ways of assessing the value of the teen pregnancy prevention program, both of which are captured in this simple equation:⁴

$$\text{Value} = \frac{\text{Benefit}}{\text{Cost}}$$

In our example, the benefit is the reduction of teen pregnancies.

Cost-effectiveness analysis compares the impact of different programs seeking to achieve the same result. For example, if our program costs \$100 per participant, while a different program serving the same population achieves the identical results for \$75, our program is less cost effective.

Cost-benefit analysis takes cost-effectiveness analysis one (ambitious, if not heroic) step further by monetizing the value of an averted teen pregnancy. In principle, this allows a donor to compare the effectiveness of the teen pregnancy prevention program with, say, a program for preventing drug abuse.⁵

Even when one cannot undertake a formal cost-benefit analysis, a donor may have an intuitive sense of when a program is having enough impact to justify his or her charitable support: \$3,000 to prevent one pregnancy⁶ may seem like a bargain, whereas \$30,000 may seem excessive.

The framework for assessing risky strategies adds the element of risk to the cost-benefit equation in the form of likelihood of success. The value, or expected return, of the strategy takes into account the magnitude of the benefit if the strategy succeeds, the likelihood of success, and the cost of pursuing the strategy.

$$\text{Expected return} = \frac{\text{Benefit} \times \text{Likelihood of success}}{\text{Cost}}$$

The equation captures the fact that a risky philanthropic venture with a small likelihood of success is justified by very high benefits if it does succeed. That's the explanation for much policy advocacy, second-track diplomacy, early stage R&D, and, of course, joining the group pushing the boulder up the glacier. But, it is devilishly difficult to quantify the likelihood of success in these cases.

At the Hewlett Foundation, we have been working on approaches to reducing

the margins of error by keeping track of factors that commonly contribute to success. For advocacy, this includes the existence of technically and financially viable solutions, windows of political opportunity, and the presence of inside and outside champions for the outcome. The expertise of experienced advocates plays a role as well. But even experts lack reliable intuitions about the probability of unlikely outcomes, exhibiting more confidence than accuracy.⁷ Thus, thoughtful philanthropists gather as much information as possible about the paths to a successful outcome, make their best estimate, place their bets, and adjust as new information becomes available.

IN PROGRESS

Assessing Progress | The activities prescribed by a logic model provide the framework for assessing progress. Because the pregnancy prevention program's activities

of reliable information.

Even non-risky strategies can be derailed by exogenous events—consider the many social programs in New Orleans that faltered in the wake of Hurricane Katrina. But risky strategies tend to be even more vulnerable: unforeseen events may relegate an issue that was ripe for legislative action to the back burner, or key supporters of a policy measure may have their attention drawn to other matters or even defect.

The logic model for many social interventions is essentially linear: additional counselors counseling additional participants lead to fewer teen pregnancies. In contrast, most risky philanthropic ventures are nonlinear. There may be long periods during which no progress is apparent, and then the desired outcome occurs—or not. And even if the desired outcome occurs, other forces may try to thwart its effective implementation or try to reverse it.

Before investing in a particular venture, a philanthropist needs to understand how and why it is likely to achieve its intended outcome.

have a causal relationship to its intended outcome, the organization and its funders can assess progress in terms of, say, the number of counselors and teenage participants recruited, the number of counseling sessions held, the participants' views of the value of the sessions, and (perhaps) its effect on their behavior. A small program may not be able to obtain reliable information about the pregnancy rates of its teen participants, but basing the program on reliable documented studies gives rise to reasonable confidence that the activities will deliver the hoped-for results.

The logic model for a risky advocacy strategy provides a structurally analogous framework, but is much more dynamic and far less certain of success. If an essential aspect of the strategy is to communicate with uncommitted members of the public utilities commission, or with individuals or groups who could influence them, then it is possible to determine whether the communications were made, received, and acted on. But throughout the process, advocates must make tactical decisions in the absence

Paralleling these observations, a philanthropic donation to a well-tested service-delivery program is almost assured of having some impact. Although some risky ventures may have partial successes, others have all-or-nothing outcomes. For example, after years of advocacy by climate organizations, Congress failed to adopt a cap on carbon dioxide emissions.

Tactical retreats and pulling the plug | Changing circumstances during the implementation of a risky strategy sometimes call not merely for adjustments but for a tactical retreat until the environment improves. For example, after a multi-year initiative to reform public school governance and finance in California, the Hewlett Foundation concluded that it could not make significant gains until the state addressed more fundamental governance problems. Rather than abandon the effort entirely, the foundation has continued to support a group of organizations to engage in research, conduct policy analysis and advocacy, and be prepared to act when promising opportunities arise.

At some point, even a funder with a high tolerance for failure may decide that the opportunity costs of continuing a risky strategy outweigh its potential benefits. For example, most US climate advocates have shifted attention from Congress to the states. But it's hard to know when to give up. It is said that it took Thomas Edison 1,001 tries to come up with a workable light bulb, and that he commented: "I have not failed 1,000 times. I have successfully discovered 1,000 ways to *not* make a light bulb." But what if Edison had given up before the 1,001st effort?

Just as the expected return equation provides a framework for deciding whether to undertake a risky venture in the first place, it provides guidance in deciding whether to abandon an ongoing venture. Besides the difficulty of doing the numbers, however, the decision to pull the plug is complicated by the competing psychological phenomena of impatience on the one hand, and the fallacy of sunk costs on the other.

EX-POST

Learning from Success and Failure |

Evaluating the actual impact of a philanthropic strategy necessarily occurs after the strategy has been implemented. The evaluation provides feedback for improving the design and implementation of the strategy and deciding whether to continue investing in it.

For these purposes, one must look beyond outcomes to ask whether the strategy actually had impact. Although an organization and its funders may rightly take pleasure in seeing their intended outcome occur, the value of their work depends on whether the outcome would or would not otherwise have occurred. The point is nicely captured by the Sam Gross cartoon published in the Aug. 1, 1991, issue of *The New Yorker*, which shows a pack of wolves howling at the moon, with one saying: "My question is: Are we making an impact?"

The counseling program achieved its intended outcome to the extent that participants did not become pregnant, but lacked impact if they wouldn't have become pregnant in any event. The RCT that underlay the program's theory of change predicted its impact by establishing the baseline of pregnancy without the intervention and showing that the intervention had a statis-

tically significant effect.

Assessing the environment organization's impact in advocating for renewable portfolio standards is a quite different matter. Even if the desired outcome occurred, exogenous factors, such as political donations by a wind turbine manufacturer, may have contributed to the public utility commission's adoption of the standards. Of course, many exogenous factors contribute to a teenager's getting pregnant or not, but evaluation of the program through RCTs or similar means is designed to assess the program's contribution to the outcome by holding exogenous factors constant. The theory of change underlying the advocacy strategy is neither as specific nor as specifically evaluable.

From the evaluation of the teen pregnancy prevention program, one can say that the program contributed a certain amount to reducing teen pregnancies. By the same token, one can say that the outcome was attributable to the program. To the extent a donor supported the program, he or she can appropriately claim attribution as well.

Occasionally, but very rarely, the causal link between an advocacy strategy and its intended outcome is so clear that one can attribute the outcome to a particular organization. Suppose, in our example, that the public utilities commissioners were predisposed against renewable portfolio standards, that no other groups advocated for them, and that our organization persuaded the commissioners one by one.

But typically there are so many exogenous factors and so many other advocates that, as Teles and Schmitt say, "If it is hard to know whether advocacy played any part in a policy outcome, it is harder still to know whether any particular organization or strategy made the difference." In these cases, which are typical of risky philanthropic ventures, some commentators have used "contribution" in a different sense, meaning not that the organization's effort contributed a certain percentage to the outcome, but rather that its efforts increased the likelihood of achieving the outcome (though seldom quantifiably). It's like joining the group pushing the boulder up the glacier, but not knowing with much confidence whether the group would have succeeded without you.

Thus the success (or failure) of an advo-

cacy strategy provides little information about the soundness of its underlying theory of change. Second-track diplomacy has the same characteristics, and then some, because diplomatic negotiations are even more opaque than domestic politics. Although not a paradox, it is an irony of most risky grant-making that although one can make thoughtful bets ex-ante, one may not fully know how they eventuated ex-post. Kierkegaard wrote that "Life can only be understood backwards; but it must be lived forward." Alas, much risky philanthropy cannot be understood even in retrospect.

Donors who made risky grants with high potential benefits ex-ante may regret the decision if they do not succeed. Indeed, hindsight bias may lead a foundation's board or management to think that its staff should have anticipated that a risky strategy would fail. Without claiming that the Hewlett Foundation's staff and board are entirely immune to this pervasive psychological bias, we try to learn from our failures as well as celebrate successes, reminding ourselves that taking appropriate risks may be philanthropy's highest calling. ♦

Copyright © Paul Brest 2012. This work is licensed under the Creative Commons Attribution 3.0 Unported License. To view a copy of this license, visit <http://creativecommons.org/licenses/by/3.0/>.

I am grateful for improvements suggested by Ivan Barkhorn, Peter Belden, Iris Brest, Jeremy Brest, Jacob Harold, C.R. Hibbs, Steven Teles, and Fay Twersky.

- 1 There are other methods for evaluating such interventions—for example, matching participants in the program with a group of teens with similar demographic characteristics—that are less expensive but also usually less robust. And even strong findings in an RCT do not entail that an intervention will be equally effective with a different demographic group or under different circumstances.
- 2 Compare Christopher Trenholm, Barbara Devaney, Ken Fortson, et al., *Impacts of Four Title V, Section 510 Abstinence Education Programs*, Princeton, N.J.: Mathematica Policy Research Inc., April 2007, with Douglas Kirby et al. "Sex and HIV Programs: Their Impact on Sexual Behaviors of Young People Throughout the World," *Journal of Adolescent Health* 40 (2007): 206-217.
- 3 A classic work on this subject is John Kingdon's *Agendas, Alternatives, and Public Policies*.
- 4 For a good discussion of such approaches, see *Measuring and/or Estimating Social Value Creation: Insights Into Eight Integrated Cost Approaches*, <http://www.gatesfoundation.org/learning/documents/wwl-report-measuring-estimating-social-value-creation.pdf>
- 5 The Robin Hood Foundation has done ambitious work in this respect.
- 6 Actually the average cost is \$3,333, based on the cost of \$100 per participant and its success in avoiding 3 pregnancies for every 100 participants.
- 7 See Philip Tetlock, *Expert Political Judgment: How Good Is It? How Can We Know?* 2006.

Assessing One's Own Performance

What the Irvine Foundation has learned over the past six years about performance assessment

BY JAMES E. CANALES & KEVIN RAFTER

Private foundations enjoy a unique degree of freedom to pursue their missions without the constraints that face many other nonprofit organizations. Such freedom can also be a challenge insofar as it comes without the external forces that can drive organizations to achieve results. As Thomas Tierney and Joel Fleishman note in their book *Give Smart: Philanthropy that Gets Results*: “in philanthropy excellence is self-imposed.” That reality contributed to our decision to create a new framework—called Foundation Performance Assessment—to help us assess The James Irvine Foundation’s performance.

In this article we describe the Irvine Foundation’s approach to performance assessment and discuss some of the lessons we have learned over the past six years. We write from the dual vantage point of having been involved in creating and evolving the foundation’s approach to performance assessment as well as producing the annual report that describes our performance (report preparation is integrated into our assessment process). We define Foundation Performance Assessment as an effort to assess the organization’s performance by examining the various levers by which the institution can achieve its mission.

The Irvine Foundation was created in

JAMES E. CANALES has been the president and CEO of The James Irvine Foundation since 2003 and held numerous positions at Irvine since 1993. He is also a past board chair of Grantmakers for Effective Organizations.

KEVIN RAFTER is the manager of research and evaluation at The James Irvine Foundation and has been at Irvine since 2007. He received a PhD from the City University of New York, where he conducted research on the nonprofit and philanthropic sector.

1937 to benefit the people of California. In the 75 years since its inception, the foundation has awarded more than \$1 billion in grants to thousands of organizations serving Californians. The majority of our grantmaking falls into three program areas: youth, arts, and California democracy. The youth program helps high school students build a strong foundation for success in both college and career. The arts program promotes engagement in the arts for all Californians. And our California democracy program advances effective public policy decision-making that is reflective of and responsive to all Californians. All three

committee of board members and staff to help shape this work, and we spent several months studying the best practices of other foundations that were pioneering foundation performance assessment. Although a few foundations—most notably the Robert Wood Johnson Foundation—stood out, we concluded that few foundations were taking a comprehensive approach to performance assessment, and there was therefore an opportunity for the Irvine Foundation to contribute to the development of new ideas and approaches in this emerging area.

With the few existing models in mind and the knowledge that foundation per-

Our work in evaluation and performance assessment helps focus program staff on the goals and outcomes for their grantmaking.

areas are guided by our mission of expanding opportunity for the people of California.

THE PERFORMANCE ASSESSMENT FRAMEWORK

A major strategic planning process in 2003 led the Irvine Foundation to streamline its program focus and create a new performance assessment approach. As the result of a comprehensive review and planning process, we updated our mission statement and trimmed six diverse program areas to three.

Once we had identified where we would focus and what we sought to achieve, we turned our attention to how to measure and assess our progress toward achieving our goals. We assembled an ad hoc

performance assessment was still nascent, we created the Performance Assessment Framework. With the Irvine Foundation board of directors as the primary audience, the framework addresses six areas—the first three focusing on our programmatic work, and the remaining three providing a more institution-wide view.

We evaluate our performance in these six areas by asking the following six questions. The results of this inquiry provide the basis for our yearly report to the board.

1. What is the context in our program fields? This section of the performance assessment outlines information that helps our board understand how the Irvine Foundation’s work fits in a broader context. For

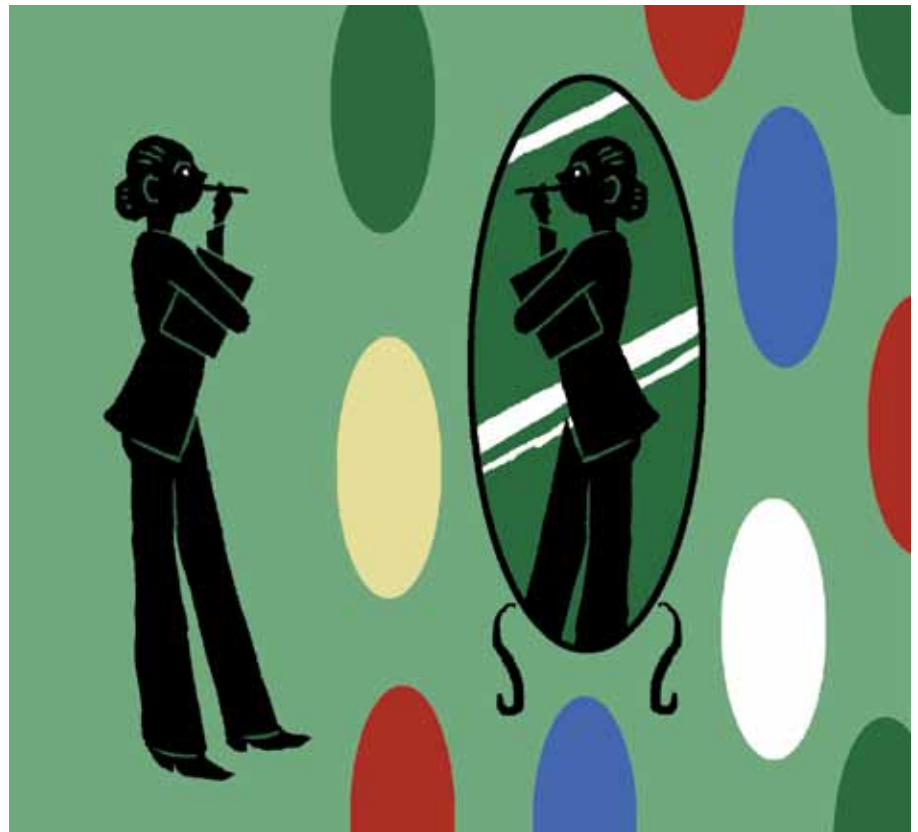
example, we include external indicators and new research findings that are relevant to our program goals. We also report on grantmaking by peer foundations in similar areas. Board members have indicated that this context is valuable to their deeper understanding of the challenges and opportunities we face in each program area.

2. What progress are we making toward our program goals? This section reports on evaluation findings and program progress indicators that track the impact of our grantmaking. The progress indicators are developed by each program team and cover a range of information, both quantitative and qualitative. The indicators are organized by the goals and priorities in each of our programs. In many respects, this section covers much of what is traditionally considered to be “evaluation” of a foundation’s work.

3. How do lessons from our program work improve our approach? This section discusses how we have used our grant monitoring, evaluations, and other engagements in the field to inform and refine our strategies and implementation. Because the Irvine Foundation’s philosophy of evaluation is guided by continuous improvement and refinement, much of our work with the board during the year focuses on this set of questions. In this section of the annual performance report we summarize our activities and reinforce our commitment to ongoing improvement.

4. How is the foundation exercising leadership in the field? This section, which shifts the focus from specific programs to the broader organization, assesses the ways we can use our leadership platform and voice to extend our impact and advance the foundation’s mission. We do not presume that leadership is conferred upon us by virtue of our resources, but we are also mindful that foundations can play important leadership roles, especially when done with humility and through authentic partnership. Here we assess leadership activities undertaken by the foundation as well as ways we help to frame discussion, often via publications. We are beginning to integrate social media measures into this work.

5. How do key stakeholders perceive us, and how do their perceptions inform our work? This section reports feedback we have solicited about how important constituents view the foundation—typically



collected by third-party surveys, confidential interviews, or other organized fora. We typically include results from Grantee Perception Reports administered by the Center for Effective Philanthropy, website user surveys, and other constituent feedback activities. We have found that the self-imposed requirement for external feedback to include in our annual report encourages us to identify more feedback opportunities than we might otherwise—a good development.

6. How are we performing on measures of financial health and organizational effectiveness? In this section we track a number of indicators related to the foundation’s investment performance, operating ratios and costs, board and staff diversity, and institutional developments. In contrast to other parts of the report, for this section we are able to draw on ready sources of financial benchmarks for comparable institutions in the field and for the foundation’s past performance.

It is important to note that when answering these six questions, we operate with the assumption that the measurement needs to fit the subject matter. As much as possible, we rely on quantitative data for maximum

precision and clarity. For some topics in our framework, such as exercising leadership or gathering constituent feedback, a quantitative approach may not be as useful, so we try to balance the quantitative and qualitative.

EVOLUTION, REFINEMENT, AND IMPROVEMENT

Our Performance Assessment Framework has evolved using feedback from the board and our experience creating and using the annual performance reports. The most important evolution has been to change the sequence and emphasis of topics. The framework initially led with a review of new grantmaking during the reporting year, but we have reorganized the report to focus attention on the progress and results of our previously awarded grants. We include a grantmaking summary as an appendix to the annual performance report, but we rely on quarterly dashboards to keep the board members up to date on recent approvals.

An important goal in creating the Performance Assessment Framework was to create a view of the foundation’s work as a whole rather than as a series of parts. It was the quest for a holistic view that motivated us to organize the report by topics—the six

questions—rather than by programs. That said, as a multipurpose foundation, we have three very distinct bodies of work in our three core programs, so although we take stock of the foundation as a whole, we do not attempt to aggregate measures into a single index for the foundation. Rather, we weigh progress and challenges to learn from each.

AUDIENCES FOR ASSESSMENT

Over time we have broadened our thinking about the audiences for our performance assessment work. The main audience for our annual performance report is the foundation's board of directors. The report is one of the primary deliverables for our annual retreat at which we conduct in-depth conversations with the board about our work. As part of the evolution of our framework and with our board's feedback, we came to define three additional audiences for our performance assessment work.

Our staff, especially the program staff, are a second important audience that can derive benefit from both the assessment process and results. The annual performance report represents a regular check-in on our ongoing process of strategy development and refinement. Our work in evaluation and performance assessment helps focus program staff on the goals and outcomes for their grantmaking. We have found that challenges in performance assessment often help uncover areas where our strategy needs refinement and elaboration. The process of reviewing progress indicators and other material to develop the annual performance report helps us reflect on how the progress we've made should inform the work ahead.

In addition to the two internal audiences, we believe that the analysis in our annual performance report can provide grantees and other funders—our third target audience—with a better understanding of how we define success in our work. Over time, we believe this understanding can facilitate collaboration toward shared goals. We reconceived the foundation's public annual report (which is different from the annual report we provide to the board) to integrate performance assessment content so that our grantee partners have easier access to the information.

The general public is a fourth, although lower priority, audience. By including more

performance assessment in our public annual report and making it generally available, we seek to contribute to a broader understanding of philanthropy's role in society beyond the grantmaking transaction.

LESSONS LEARNED

Literature on measurement and evaluation reminds us that consciously building in opportunities for learning will help us use our results in actionable ways. At the Irvine Foundation, that wisdom has helped to guide our approach to performance assessment across the foundation. From the beginning, we knew we were experimenting; building on the work of pioneers in foundation-wide assessment, such as the Robert Wood Johnson Foundation. We also

leading up to board decisions about grant awards at designated intervals throughout the year. When one cycle concludes, another begins—or more often, cycles overlap and docket deadlines are constantly looming for foundation staff. We have concluded that this structure tends to discourage careful reflection, thoughtful analysis, and distillation of lessons learned. Often, these activities are considered luxuries at their best and distractions at their worst.

For a foundation to embrace fully a commitment to institution-wide performance assessment, careful consideration must be given to addressing this structural barrier, which tends to focus us primarily on the next deadline at the expense of what we are learning and how we are using that learn-

A commitment to foundation performance assessment can force us both to consider context and to build in opportunities for learning.

knew we did not have the answers, and now, after six years of creating and evolving our framework, we know that we still have more to learn. At the same time, we are equally persuaded that our commitment to foundation-wide performance assessment has made us a better foundation—one in which a culture of reflection, learning, and refinement has become even more pronounced, and in which we actively find ways to use what we learn from performance assessments to improve our strategic thinking and, by extension, to deepen our impact.

Recognizing that our assessment framework remains a work in progress, we offer the following three lessons in the hope that they may inform others who are interested in foundation-wide performance assessment.

Lesson 1: The traditional structure of philanthropic activity can conflict with a commitment to performance assessment, so it is important to address related barriers and incentives.

Incentives within foundations are typically organized around the core activity of awarding grants. The IRS mandate to spend grant dollars annually helps to drive the requisite work of grantmaking—proposal review, site visits, docket preparation—all

ing. We have discovered a related temporal challenge. Because of regulations related to payout, grant budgets are often organized around an annual calendar. Our program goals and aspirations rarely follow annual timelines, nor should they, if they are sufficiently ambitious in scope. So how do we reconcile these different timelines?

At the Irvine Foundation, the discipline of producing an annual performance report to our board (and then sharing it publicly) has oriented us toward a sharper focus on reporting progress, not necessarily final results. To do so, we have had to orient ourselves toward identifying shorter- and medium-term indicators and measures of progress that we can track and report on in annual increments.

The need to clarify these indicators and measures of progress has been a valuable contribution to our strategy development and refinement as well, because it has forced us to articulate more clearly both the logic and sequence of outcomes we seek. We believe that any foundation committed to assessing its performance must determine how best to keep the grantmaking work moving forward while creating the space for consideration of broader progress assessment.

Lesson 2: Foundation performance as-

assessment both requires and fosters a culture of reflection and learning that can lead to ongoing refinement and program improvement.

Much of today's business literature speaks to the importance of building adaptive organizations that stay closely attuned to their external environments and retool strategies in ways that align with that ever-shifting context. Although foundations do not necessarily think in terms of "competitive advantage" or worry about going out of business, foundations do have an obligation to remain attentive to the context of their work.

Similarly, we need to find ways to create feedback loops that permit us to use what we learn to improve our strategies. A promising development over the past decade has been a more intentional focus on evaluation for learning and improvement rather than simply for auditing purposes or declaring success or failure. Although we certainly need to use evaluation to guide our understanding of whether we are succeeding or failing, we stand to benefit significantly if we can use what we learn to improve our strategies and their execution.

In this respect, performance assessment is inextricably linked to program strategy, and it has been our experience that a focus on assessment has improved the rigor and logic of our program strategy. As previously noted, a key contribution of this process has been defining progress indicators that allow us to determine whether we are making the kind of progress we seek in the short term, and if not, to understand what that might imply about our strategies. We have discovered that even the process of identifying these indicators can help us to surface possible instances of underdeveloped strategic thinking or unrealistic expected outcomes.

A commitment to foundation performance assessment can force us both to consider context and to build in opportunities for learning. From the beginning, our Performance Assessment Framework contained a section on what we call "program context." It was designed to encourage us to collect data about the contextual issues related to our programs. We also annually collect data on other philanthropic investments in the particular areas we fund. Examining both of these contextual data sets helps us understand better how we view the foundation's

contribution and track other philanthropy and the broader trends in our focus areas. Interestingly, this contextual data has been one of the areas that our board has found the most interesting. It has helped them to position our work in a broader context and has encouraged us to find other ways to expose the board to the environment and context for the foundation's efforts.

Understanding the context and reflecting upon our learning, however, can be useful only if it can then be translated into action. This is where we have tried to demonstrate our commitment to adaptability without falling prey to a tendency by foundations to embrace the issue of the day. For example, in our California democracy program we have remained focused on the goal of advancing more representative public policy decision making. As we have approached this priority from different angles, such as voter mobilization and civic engagement, the program team has analyzed the products of its grantmaking and adjusted its strategy as appropriate.

Lesson 3: Successful foundation performance assessment requires both buy-in and engagement from stakeholders at all levels.

The old adage that "leadership starts at the top" holds true for foundation performance assessment. Without demonstrated commitment and visible leadership from the board, the CEO, and other senior officers, foundation performance assessment simply cannot succeed. We put it this starkly because the effort required for foundation performance assessment mandates full institutional commitment and cannot be the province of just the evaluation director.

Obtaining the buy-in of the board from the very beginning was essential. We created a board task force to help us design our approach to foundation performance assessment, and from the outset we viewed the board as the lead initial stakeholder. We did so because we believed it would advance board members' understanding of the foundation's progress and impact—an important outcome by itself—but we also knew it would signal both internally and externally just how important this work was to the foundation.

For the process to take hold, it also requires broad-based buy-in from the staff. We were perhaps well-served by the fact

that our foray into foundation performance assessment coincided with the adoption of a new strategic plan and the staff change that often accompanies a shift in new directions. The foundation's staff embraced this approach from the outset, and it became part of the way we did business. Nonetheless, we recognize the need to ensure that staff view performance assessment not as an additional burden, but rather as a tool to inform better decisions.

CONCLUSION

This article explains how we have structured and evolved our foundation's performance assessment. We are eager to share our approach, not just because we are persuaded that adopting an institutional view of performance assessment can improve a foundation's work, but also because we want to create a broader community of foundations committed to learning from each other about this important dimension of our work. Fortunately, we have moved beyond the question of whether measurement and evaluation are useful, but there remains much to be explored about how the tools of measurement and evaluation can be applied across a foundation's work. That's where we have been experimenting for the past six years and where we know we still have a great deal to learn.

Among the questions that we have yet to explore is how an approach to foundation performance assessment can work in organizations of different sizes and scopes. Although the Irvine Foundation is a relatively large foundation, we are also regional in scope and focus on only a few program areas. How this approach can be applied in a foundation with a broader mandate and more diverse portfolio of grantmaking remains an open question.

We hope this contribution to the body of knowledge on performance assessment can support robust exploration of questions such as these. In the end, our institutions exist to provide a public benefit, and we must therefore embrace opportunities that might enable us to deliver on that promise more effectively. We believe that a commitment to foundation performance assessment offers one such opportunity, and we know there are more experimenting to be done, more learning to be accumulated, and many stories to be shared. ♦



The Aspen Institute's mission is twofold: to foster values-based leadership, encouraging individuals to reflect on the ideals and ideas that define a good society, and to provide a neutral and balanced venue for discussing and acting on critical issues. The Aspen Institute does this primarily in four ways: seminars, young-leader fellowships around the globe, policy programs, and public conferences and events. The Aspen Institute Program on Philanthropy and Social Innovation seeks to maximize the impact of social-sector leaders in contributing to the good society at home and abroad. It pursues its mission through convenings, leadership development initiatives, communications, and strategic partnerships. It hosts the Aspen Philanthropy Group, an agenda-setting body of foundation leaders at the cutting edge of change, and it spurs dialogue among leaders from the private, public, and social sectors in working groups on specific issues of concern.

The Aspen Institute is based in Washington, D.C.; Aspen, Colo.; and on the Wye River on Maryland's Eastern Shore, and has an international network of partners.

The Aspen Institute
One Dupont Circle, NW
Suite 700
Washington, D.C. 20036
www.aspeninstitute.org

This sponsored supplement was produced by *Stanford Social Innovation Review* for the Aspen Institute Program on Philanthropy and Social Innovation. The supplement was underwritten by a grant from the Ford Foundation.

Illustrations: Adam McCauley | Design by David Herbick
