

StanfordSOCIAL INNOVATION^{REVIEW}

Sponsored Supplement to *SSIR*

Risky Business

By Paul Brest

Stanford Social Innovation Review
Summer 2012

Copyright © 2012 by Leland Stanford Jr. University
All Rights Reserved

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.