



CHAPTER ONE

USING EVALUATION TO ADVANCE A FOUNDATION'S MISSION

Laura C. Leviton and Marian E. Bass

Why should foundations and their grantees conduct evaluations? Evaluation should be done for a good reason; there are many good reasons and good uses for the results. In this chapter, we present several illustrations of the ways evaluation can be used in foundation-supported work. We want to capture the reader's imagination, express concretely what is possible to achieve, and point to some key conditions that we believe are necessary in order to conduct effective evaluations in the foundation context.

This chapter is intended for both the users and producers of evaluations. A precondition for useful foundation evaluations is that participants better understand their purpose. Participants include those who conduct the evaluations, the foundation staff members, as well as the grantees, who may themselves conduct evaluations or whose activities may be the subject of evaluation. Equally important, however, is for evaluation professionals to understand the distinctive context of foundation work. In the conclusion to the chapter, we summarize important features of the context, describe some potential uses of evaluation in the foundation context, and give specific recommendations for working within that context to improve evaluations.

The suggestions of Dana Phillips and Ernest House are gratefully acknowledged. Thomas Gilmore first suggested the material in Cases 8 and 9; William Bickel suggested the content in Case 2.

We make two basic points—emerging from two different perspectives—in the chapter:

1. *Looking inside a foundation*, useful evaluation depends on the organization's size, its culture, and the placement of responsibility and authority for evaluation.
2. *Looking outside a foundation*, evaluations will be better and more useful if they consider both the power and vulnerability of foundations and the predictable effects of foundations' power on grantees, evaluators, and the grantmakers themselves.

Two basic problems that must be confronted at the outset are that (1) most evaluations conducted for foundations are unsatisfactory (Patrizi, 2002), just as they are in the public sector, and (2) people bring preconceptions to the conduct and use of evaluations that are incorrect and often can harm the prospects for getting useful information. For example, grantees fear that they will be “evaluated”—and in poor practice, they sometimes are. Good practice is very different, as illustrated in this chapter. As a result of these problems, there is frequently tension and apprehension over the terms under which evaluations will be conducted, the ground rules of behavior, roles, and responsibilities, and the uses that might be made of the results. Foundation leadership, board members, staff, and grantees have frequently expressed disappointment over the value of evaluation results.

This situation is partly caused by confusion within philanthropy about the purposes for, and the likely consequences of, evaluation. People believe they understand what the word *evaluation* means, and often they do not. Even the trusted experts and veteran grant officers often get it wrong. Although the field of evaluation rests on decades of collective experience, in philanthropy both the users and producers of evaluation tend to be unfamiliar with some basic writings in the field. The result is that evaluations are often less useful than they might be, thus wasting time, resources, and valuable opportunities for learning. It is important to demystify the process and products of evaluation on behalf of the stakeholders (those with an interest in the program or its evaluation). It is important to identify the preconditions for getting useful information from evaluations.

An important product of evaluation's collective experience is summarized in the introductory chapter and helps to define what evaluation is really about. The Guiding Principles for Evaluators (see Exhibit I.1 in the Introduction), endorsed by the American Evaluation Association (AEA), is a key document that encompasses the wide-ranging approaches to be found in evaluation today. Both the users and producers of evaluation will benefit greatly from reading the principles in their entirety, either at the association's Web site (www.eval.org) or in any issue of *The American Journal of Evaluation*. The principles give stakeholders an idea

of what to expect from good practice (Leviton, 2001). By reading and sharing the principles, evaluators and those who commission evaluation can often alleviate tension and prevent disappointment.

It is also important to define what we mean by the “use” of evaluations. Leviton and Hughes (1981) describe three general kinds of use that still hold up under scrutiny today: (1) *persuasive use* that argues for or against a policy or program, (2) *instrumental use* in identifiable decisions, and (3) *conceptual use* (such as the serious consideration of results in decision-making, enlightenment about program reality, or testing assumptions about the underlying problem). Foundation work gives its own imprint to these general types, as seen throughout this volume. At the end of this chapter, we point to some distinctive purposes for evaluation in philanthropy.

Looking Inside Foundations

Foundations vary in many ways that are important to evaluation; many are covered elsewhere in this volume. This chapter focuses on just three internal drivers: (1) foundation size, (2) foundation culture, and (3) the location of responsibility and authority for evaluation.

Foundation Size

A foundation's size refers to its assets, payout (dollar amount of grants issued each year), and staffing. About 350 foundations pay out \$10 million per year or more, but only 29 of these gave more than \$100 million in grants in 2001 (Foundation Center, 2004). The amount of the payout often affects the number of staff a foundation can maintain to manage its grantmaking, as well as their degree of specialization within program areas and technical support functions such as evaluation. Other features also affect the degree of grant officers' involvement with evaluation: they may award and oversee all grants, or they may operate through a variety of intermediary, “re-granting” organizations (Szanton, 2002). “Operating foundations” have staff who actually manage and implement programs.

High-Quality Evaluation Is Within Reach. The size of the payout does not have to affect the quality or usefulness of evaluation, but smaller foundations may believe that high-quality evaluation is simply beyond their reach. In truth, cost is an abiding issue among the stakeholders for evaluation of all kinds, whether it is conducted for government, for individual agencies and organizations, or for foundations. Decision makers often have no idea what an adequate study should cost.

However, cost is not the central issue for foundations in allocating evaluation resources. A more important issue is to select the most cost-effective approach, given the purpose of the evaluation. In other words, what are the real information needs of the intended audience, how important is the result, and what standard of evidence is required? High-quality evaluation does not necessarily mean high cost, given the many legitimate approaches available, and cost should be driven by decisions about what is a “good enough” standard of evidence for the purpose at hand (Rossi, Lipsey, and Freeman, 2004).

Chapter Seven outlines the challenges of resource allocation in more detail. Sometimes low-cost options may be highly appropriate to the purpose, as seen in Chapter Ten. At other times, however, it is essential for strategic purposes to invest more in the evaluation than in the program. As described in Case 1, the strategic aim of a foundation initiative can be lost, unless there is an appropriate evaluation focus. And evaluation should never be conducted *pro forma*, that is, just because someone believes it is supposed to be conducted. If evaluation is likely to be of poor quality, or if it has no purpose (for example, when foundation staff and program managers do not expect to use evaluation results in their decisions or thinking), then any evaluation is too expensive.

Case 1: Strategic Evaluation in Field Building

Program managers and advocates tend to view evaluation as taking away needed resources from programs. However, foundation officers sometimes value it because they recognize ways in which evaluation can help them build a field in line with the foundation’s mission (Hirschhorn, Kelly, and Gilmore, 2002). Examples of *field building* include:

- Encouraging participation in the arts (The Wallace Foundation)
- Supporting research in health policy (The Robert Wood Johnson Foundation)
- Strengthening community leadership (The W. K. Kellogg Foundation)
- Creating the field of clinical epidemiology in less-developed countries (The Rockefeller Foundation)
- Building the capacity of youth development programs (The Edna McConnell Clark Foundation)

Individual programs may work together to build a field. For example, The Robert Wood Johnson Foundation (RWJF) funds several programs designed to improve the effectiveness of public health agencies, including leadership training, information infrastructure improvement, and advocacy and public awareness ini-

tiatives. Evaluation as part of a field-building strategy can work either at the level of individual programs or across various programs that are associated with building the field.

In 2003, funding for the evaluation of two national programs came into question at RWJF. In both cases, the grant officers said there was no point in funding a program if there was no funding for evaluation of the program. Grant officers of RWJF often believe that the evaluation of individual programs will help them build a field. They see strategic advantages for the national programs for which they have oversight. Evaluations aid the strategy by signaling the importance of the field being built, by informing funders of progress over time, and by assessing the value of the field to society.

The dynamics certainly are not perfect, but, often enough, both RWJF staff and grantees believe that evaluations can leverage important changes in a field of interest. This is the case whether the findings are positive, negative, or mixed. If the evaluation is positive, that is good publicity for the field. Because the evaluations are independent and generally of high quality, more confidence can be placed in the findings. If the evaluation is negative or mixed, then that becomes a basis for further learning in a field. It may not be pleasant, but it is the basis of a productive dialogue (for an example, see the description of the SUPPORT study in Chapter Six).

What Is Meant by the "Standard of Evidence"? A wide range of legitimate approaches is employed in evaluation today. Foundations have different program aims, and therefore they prefer different evaluation methods to give them the information they need. In the range of evaluation approaches, there is no single preferred method; there are simply appropriate and inappropriate methods to answer the question that is being posed. However, the legitimate evaluation approaches all use methods that have recognized standards of evidence—agreed-upon rules and criteria for the work—to which reputable evaluation practitioners adhere. The methods employed range from the quantitative analysis of so-called "hard outcomes" at one extreme (such as changes in the prevalence of a disease) to the qualitative, social-constructivist interpretation of a program's "story" at the other (such as stakeholder perceptions of participation in the arts). Foundation officers and grantees should be familiar with, and insist on, the standard of evidence for the methods that are to be used.

Much of the confusion regarding the choice of an evaluation approach centers on the use of cause-probing methods such as randomized experiments or quasi-experiments. These methods have had an illustrious, but uneven, history in demonstrating the worth of public policy and program approaches. Unfortunately, opinions

about their use are often ill-informed. Some believe that no evaluation question is worth asking unless it involves a rigorous scientific assessment of outcomes. Yet in many cases a randomized experiment is premature or simply inappropriate to the situation, while other methods can help reduce uncertainty or keep an investment on course. Others believe that social experimentation is unethical and often overly negative about the true worth of social programs. Yet without social experimentation, foundations may help to advance programs and policies that remain unproven or eventually become discredited. These views of randomized experiments are all erroneous (Dennis, 1994). Field experimentation is one tool for the tool chest, to be used judiciously (see Chapter Twelve).

Size Can Affect Foundation Goals. Because of foundations' size, evaluations that address their goals will differ greatly in content, focus, structure, and consequences. The largest foundations often aspire to achieve societal change far beyond their actual payout. Evaluations for these foundations are often linked to effecting social change or to building a field. Staff of these foundations may be national or international experts in a field of work; their prestige derives, in part, from the funds they provide and their track record for achieving social change. Evaluation could either bolster or threaten their status in the field, with predictable consequences.

Other foundations have a local or regional mission, including most community and conversion foundations. Like the largest foundations, they fund worthy causes and may contribute to field building, but they are also important civic leaders. They often have local influence behind the scenes, far beyond the actual dollars they can provide. For these philanthropies, longstanding relationships with other organizations play an important role. Evaluation could affect those relationships by offering constructive or harmful criticism and by affecting nonprofits' bottom line, with predictable consequences.

The foundation community of Pittsburgh, Pennsylvania, offers a noteworthy illustration of philanthropy's civic role. Until the mergers of the 1980s, Pittsburgh had the third-largest concentration of corporate headquarters in the United States. Dating from the time of "robber barons" such as Andrew Carnegie and Henry Clay Frick, the wealth of the corporate leaders gave rise to significant private and corporate foundations, still among the largest in the country. Operating through the Allegheny Conference on Community Development, their civic leadership is celebrated for creating a far better quality of life in the region (for history and accomplishments visit www.accdpel.org). Evaluation played an important role in some of these improvements. In Case 2, the story of public education in Pittsburgh illustrates the power of foundations' civic leadership and the ways that evaluation can assist that leadership, but also people's ambivalence about foundations' power.

Case 2: Evaluation, Civic Leadership, and Philanthropic Power

Pittsburgh in the 1980s developed one of the finest urban school systems in the nation, thanks, in part, to the work of David Bergholz, associate director of the Allegheny Conference. Evaluation guided those improvements by determining the specific needs and providing information about “what works” in school improvement (Bickel and Cooley, 1986). However, philanthropy's role in this process was distinctly behind the scenes. More recently, the foundations' civic leadership and foundation-supported evaluation have been used in a very different way.

By 2002, the Pittsburgh school board had become notoriously dysfunctional, making the district almost impossible to administer. Three foundations had made grants to the school district to implement Literacy Plus, a program to increase the number of third-graders reading at grade level. However, foundation staff were concerned about implementation, so they required that the district meet specific performance measures. When these conditions went unmet, the foundations withdrew funding for the program. They did so publicly, as a rebuke to the school board.

This decision was controversial, because philanthropy so often chooses a discreet and private role. Grant Oliphant, associate director of The Heinz Endowments, explained that the decision “was influenced by our growing emphasis on evaluation and accountability. It also reflected our ongoing commitment to the strategic use of communications as a tool for effective grantmaking” (personal communication, December 2002).

In *H: The Magazine of the Heinz Endowments*, King and Oliphant note:

Accountability has become almost a cliché in the philanthropic field. But there is a squeamishness that often sets in with foundations when accountability reaches this point, the painful moment at which funding to a valued grantee or cherished program may actually have to be withheld.

“We must be that small, though significant, resource that can stay the course, for better or for worse . . . it is the requirement for being effectively at the table” David Bergholz, [now the] executive director of the George Gund Foundation in Cleveland, argued in a scathing critique of our action published in the *Post-Gazette*. . . . The problem with this sort of reasoning is that it makes a mockery of accountability. It suggests that foundations must continue to provide funding even when, despite their best efforts, a grantee has demonstrated an inability or unwillingness to abide by mutually agreed upon terms of behavior and performance. It's nothing more than a blank check for underperformance. Accountability is not accountability unless it has teeth. . . .

There were some immediate positives. Overnight, the district's governance and leadership problems became headline news and gained the crisis status they so richly deserved. . . . Within days, Mayor Tom Murphy, who previously had kept a low profile on issues pertaining to the schools, announced the creation of a community-wide task force to examine the district's problems and recommend

solutions. . . . This was precisely the sort of community action we had hoped to spark through our decision, which is why all three foundations are funding the commission's work. The commission has no authority to implement its findings on its own. But it has a high-profile charter, membership and degree of public support that will make its work difficult to ignore (King and Oliphant, 2002, pp. 20 and 24).

Foundation Culture

Culture determines whether foundations attach importance to evaluative information and what kind of evaluation is preferred. Some believe that the standard of evidence should be as close to rigorous scientific method as possible; others rely on less formal, though systematic, methods of assessment. The culture also largely drives the kinds of uses to which evaluation will be put. (More discussion of the effects of foundation culture can be seen in Chapter Eight.)

Values and Expectations. A foundation's culture is defined by what its stakeholders value and expect. A foundation's internal stakeholders include the board of directors, senior management, grant officers, and, to a surprising extent, its grantees. For grantees, the boundary is often permeable and demonstrates the power of personal relationships. Foundations may ask trusted grantees or former grantees to sit on advisory committees or to author a white paper guiding grant initiatives. Grantees may be hired as grant officers; they may even be former grant officers. And sometimes they become foundation board members. In Case 3, we present an oral history of the emergence of an evaluator as a trusted advisor to a major foundation.

Case 3: The Rewards of Integrity

This is an oral history about how a researcher became an influential advisor to a foundation by sticking to the integrity of an evaluation. The incident must be anonymous. Most details are substantiated, but that is secondary to the role this incident plays in the foundation's culture. An anthropologist would say that it is part of the organization's "folklore" (or even a creation myth), to describe how and why the foundation supports evaluation. It also illustrates how "kinship" relations are formed in this type of elite organization (Weatherford, 1985).

In the 1980s, a large foundation created a national initiative that, over time, proved to be fatally flawed. A university-based evaluation team was hired at the outset of the program. Their report concluded that the program had no discernible effects and pointed to some very cogent and constructive reasons for the lack of effect.

The foundation had sought a highly placed, politically powerful advisory committee for the initiative. According to several parties, the chair of the committee threatened the evaluation team with political consequences if they released their conclusions—and one member of the evaluation team was coming up for tenure. The evaluation team defied the national advisory committee, presented the foundation with their original conclusions, and subsequently published their findings. There is some reason to believe, however, that the evaluation team may have lost a book contract due to political pressure on reviewers.

From the foundation's perspective, however, the committee's threat boomeranged. Far from having its intended effect, the threat against the evaluation team enhanced the team's credit in the foundation's eyes. The foundation officers, already unhappy with the way the program was going, knew they could trust these evaluators to tell the truth, courageously. Over the next decade, the foundation entrusted several other major evaluations to this team. The untenured evaluator received tenure, became a trusted foundation consultant, and eventually directed a prestigious fellowship program funded by the foundation.

Mission and Vision. Philanthropy is still an intensely personal activity, so the relationships among stakeholders matter a great deal in defining what is valued. The personal factor helps make the process more flexible and responsive, in contrast to the public sector. But at the same time, there are very few uniform procedures and rules for distributing funds, providing feedback to applicants, or wielding influence—a situation radically different from that of government funders (Downs, 1994). As seen in Chapter Six, the relative absence of procedures requires that individual grant officers develop their own ways to cope with the demands of stakeholders, with effects on evaluation.

Foundation staff are legally accountable only to the board of directors for the use of foundation funds, and the foundation is accountable only to the IRS to use the funds within legal requirements. Based on the work of certain commissions and national associations, there also exists some consensus about how foundation staff and boards of directors should comport themselves to avoid conflict of interest or abuse of power. Beyond these fundamentals, individual foundations have had to develop their own rules and traditions about fiduciary responsibility and duties.

The staff and board hold expectations about accountability and stewardship of both the financial resources and the foundation's reputation. Foundation staff make grants in line with what they conceive to be the foundation's mission, as well as a vision of the public interest. (Of course, visions of the public interest vary greatly across the political spectrum, and this spectrum is also represented

in philanthropy.) The foundation's mission and the staff member's vision become central to conducting effective evaluation.

The foundation can often value both positive and negative findings, as long as they are constructive and in line with a shared definition of the mission and vision. For example, an evaluation that offers new knowledge to use in building a field would be deemed highly useful by some foundations (Hirschhorn, Kelly, and Gilmore, 2002). As seen in Cases 4 and 5, other foundations prize evaluation that permits grantees to improve their services or to make the case for funding to other grantmakers and public agencies.

Case 4: Effects of Evaluation on Subsequent Funding

The Local Initiatives Funding Partners program is especially revealing about how grantmakers respond to evaluation. Under this RWJF program, local organizations initiate health-related programs with support from a variety of philanthropic funders. RWJF then matches local support dollar-for-dollar. The funding partners range from the very largest community foundations in the country, such as The Cleveland Foundation, to small charities and individual donors that band together to seek an award. The median number of funding partners is five, and RWJF grants range in size from \$50,000 to \$500,000. Projects range from substance abuse prevention and treatment to health promotion and improved access to medical care. Mathematica Policy Research recently conducted a survey that followed the fate of projects after RWJF funding ended (Stevens, Peikes, and Patel, 2003). These included 120 projects that were funded from inception of the Local Initiatives Funding Partners program in 1986 through July 2001. Seventy-five percent of these projects survived and are in operation today. Of these, 61 percent obtained subsequent funding from other foundations, 59 percent from public revenues, 50 percent from individual donors, 42 percent from corporate donors, 35 percent from client fees or services, and 24 percent from United Way.

Among the surviving projects, evaluation may have been critical to subsequent funding; 92 percent had evaluated their activities, and 50 percent of surviving projects believed that a positive evaluation had been important to promoting their sustainability. These percentages may seem low to evaluation professionals, but they seem high indeed to many foundation staff and nonprofit organizations. These comments illustrate the organizations' view of evaluation:

Whether you are recruiting volunteers or asking for financial support, you need to sell benefits, not features.

The RWJF grant allowed us to prove that the . . . model works. We then took the finished product back to the government. Then it was a "no brainer" for them: they could choose between jailing drug offenders for \$36,000 per year, or treating them more effectively in the community for less than \$6,000.

Case 5: Evaluation in Social Enterprise Philanthropy

Nonprofit organizations can sometimes develop earned income streams, in addition to receiving private donations and grants from foundations and government sources. Earned income is helpful to the nonprofit mission as a way to (1) defray the cost of program operation, (2) promote more flexibility and fiscal health, and (3) help produce a "soft landing" once a grant ends. Social enterprise philanthropy (also termed *venture philanthropy*), aims to develop earned income strategies for nonprofit organizations in areas such as

- Helping them market and charge for their intellectual property
- Helping them create or expand a fee schedule for services
- Providing a low-cost loan or a program-related investment, or underwriting a loan obtained elsewhere

Even though a loan might carry risk, across a portfolio the foundation may see a "blended return on investment," consisting of debt repayment and interest, as well as the expected social benefits (Emerson, 2003).

Social enterprise philanthropy is more prevalent in the western United States, where foundations that were recently founded on high-technology fortunes find it an appealing model. Foundations of all sizes engage in social enterprise philanthropy, including The Pew Memorial Trusts, The Time-Warner Foundation, The UPS Foundation, and The Illinois Facilities Fund. More information is available from the Social Enterprise Alliance or from Grantmakers for Effective Organizations (www.geofunders.org). Descriptions of nonprofits that use a social enterprise model can be obtained from The Roberts Enterprise Development Fund (www.redf.org) and from Independent Sector (www.independentsector.org).

Several evaluation questions emerge about the social enterprise model. These include the following:

What is the "blended rate of return" that can be expected? Can the foundation's contribution go farther, dollar-for-dollar, in social enterprise organizations?

Will earned income streams defray the cost of a service, or is there increased overhead for social enterprise organizations, compared to other nonprofits?

What indicators of progress are reasonable in order to evaluate a social enterprise investment?

Are the chances for sustainability greater when this model is employed?

If social enterprise organizations and their missions and services are more likely to survive, then both the earned income stream and the sustained programs may be due to good management. However, if these organizations are more likely to disappear, then that presents a problem. Answers may determine the future of this idea.

Responsiveness. Foundation culture and perceived duty may also dictate the amount and style of responsiveness to grantees and grant applicants. Responsiveness means different things to different foundations, especially in evaluation, where there are legitimate differences in purpose and approach. Some large foundations take on a totally participatory or community-driven approach. Others have historically operated in a top-down approach, prizing independence of evaluation “to keep the process honest.” These styles are contrasted in Case 6. At a minimum, though, responsiveness for foundation staff means that they must clearly convey what their mission and vision are, what they will fund, and what the “rules of engagement” are for grantees and for the fields that foundations support. Evaluation is important in conveying this information.

Case 6: Rules of Engagement

Evaluators and grantmakers will do well to clarify exactly what the rules of engagement are when an evaluation is conducted. One very important issue is whether the evaluation questions and methods will be participatory, or will be determined by outsiders on behalf of the funder. Contrast The Robert Wood Johnson Foundation (RWJF) with The W. K. Kellogg Foundation (WKKF). In 2001, WKKF was the eighth-largest and RWJF was the fourth-largest foundation in the United States by asset size (Foundation Center, 2003). RWJF espouses a view of evaluation that often (but not exclusively) relies on quantitative methods of social science. Moreover, the foundation values independent evaluation as a mechanism to ensure an honest answer. The process is seldom participatory, however, and the focus is less often on grantees’ own program improvement than on building a field of health-related work, such as access to medical care, tobacco control, or addressing the nursing shortage (Hirschhorn, Kelly, and Gilmore, 2002). WKKF takes a radically different approach to evaluation: methods are more often qualitative; research questions are often set through a participatory process, and the products are more often conducive to “lessons learned” than to definitive statements about the causal effects of grantmaking (see www.wkkf.org). Doing evaluation that is even more participatory in some respects is the Group Health Community Foundation, of the Group Health Cooperative of Puget Sound at www.ghcfoundation.org. Other foundations may view evaluation as an audit or accountability mechanism, or purely as a performance-monitoring tool for managing a portfolio of grants.

Both the RWJF and WKKF approaches have advantages and disadvantages. A participatory model does not guarantee a useful or high-quality product, although it can and regularly does produce them. It takes an important set of skills to navigate a seemingly cumbersome process of consultation with community and program stakeholders and still produce a high-quality product. In the same way, an independent evaluation mandated from outside the programs themselves can often be threatening to those programs. It takes skill to gain the buy-in and cooperation of apprehensive

program stakeholders. It also takes skill to show the client how consulting the stakeholders will help to craft a more meaningful question and gain more useful results. Evaluators' independence from the program does not mean ignorance of the program, does not mean a "black box" evaluation (one that only examines results, to the exclusion of attempting to understand the program), and does not mean oppression of program staff and clients. In truth, good-quality evaluations must somehow transcend these distinctions of "independent" and "participatory."

Responsibility and Authority for Evaluation

An effective evaluation function requires both authority from foundation leadership and a shared sense of responsibility, or buy-in, among foundation staff. We believe that some models of funding and organizing a large foundation's evaluation are superior to others. Eleanor Chelimsky (2001) agrees. She created what was arguably the most effective evaluation agency in the federal government for the U.S. General Accounting Office. In describing how foundation evaluation might be organized, she notes, "The point that needs to be made here, with some vigor, is that evaluation offices need independence, resources, strong staff, an agreed agenda for work, focused procedures and clear reporting channels among other things, if their products are to be objective, technically competent, relevant to foundation needs, and useful" (p. 26).

The Budget Process Dictates Authority over Evaluation. Independence of evaluation is strongly influenced by where the evaluation sits in the organization and, in particular, who controls the money. One model was in force at The Rockefeller Foundation beginning in 1990, when the foundation created a formal evaluation unit. The Central Evaluation Unit was established, reporting to a vice president. However, funding for evaluation was provided through the annual budget of each program division. The practical result was that each year, before each division submitted its budget for the coming year, the evaluation director would meet with each division director to determine which, if any, programs were ripe for evaluation for the coming year. They would then negotiate the selection of an appropriate evaluator. Final approval of programs subject to evaluation rested with the president of the foundation.

In this model, the evaluation director was weak relative to the division directors. She could encourage but could not compel the selection of a particular program for evaluation. Similarly, she could encourage or discourage the choice of evaluator, but with budgetary authority vested in the program rather than in the evaluation unit, she did not have the authority to make the final

choice. Relationships were cordial and professional, but there was always a risk that the foundation's grant officers would strongly endorse an evaluator with inappropriate links to, or interest in, the program.

Typically, The Rockefeller Foundation's program director and the evaluation director were able to negotiate good choices of both programs and evaluators. On at least one occasion, however, the credibility of evaluation was placed in jeopardy. Evaluation staff were required to act with great tact and diplomacy in interactions with grant officers. This need was reinforced by the natural tendency of at least some grant officers to fear that an evaluation might encourage senior managers to terminate, or drastically alter, programs to which they had devoted many years and much determined effort. The evaluation director needed to establish that she could be trusted to work in the best interest, not only of the foundation but of its individual programs and staff members. With no budgetary authority, the evaluation director had instead to establish authority indirectly, via trust and influence.

The Robert Wood Johnson Foundation exemplifies a different organizational model for evaluation within large philanthropies. At RWJF, the Research and Evaluation Unit has its own vice president, who reports directly to the foundation president. Evaluation competes for funding along with other proposals and has the authority to select the programs to be evaluated, as well as the evaluators. Grant officers are consulted, but the final choice lies with the vice president for research and evaluation and with the president of the foundation.

Recommendation: Independent Authority but Shared Responsibility. In both models, the evaluative enterprise only becomes meaningful when the foundation's grant officers accept that evaluations are conducted for the right reasons:

- To discern the strengths of program design and implementation
- To discover weaknesses amenable to redress
- To improve both grantmaking and the implementation of program plans

Learning cannot be optimal, and sometimes it is not even possible, if grant officers suspect that the evaluators are "out to get them" or to kill their programs. Evaluation works best within foundations when both evaluation staff and foundation grant officers come to the relationship in a spirit of collaboration and mutual support. This collaboration sets an important tone for the programs being evaluated and the evaluators who carry out the work. We have observed that useful evaluations in philanthropy are neither independent nor subordinate to program; instead, they strive for the best characteristics of both models by an arrangement that shares responsibility for producing a useful product.

Looking Outside Foundations

Foundations respond to their environment as well. The important stakeholders outside the organization include grantees, government, and the public. Their responses dictate the way evaluation can be conducted, as well as the ways evaluation can demonstrate its value to foundations. Some of the most important external forces include foundations' power, their vulnerability to criticism, and grantee perceptions of evaluation in light of foundations' power.

Foundations' Power

Foundation staff often caution that their relationship with grantees is unequal; many do what they can to equalize it. One reason they do so is to make sure they are apprised of issues that are important to the success of a grant. As in other settings, information gets selectively filtered to those in power, yet conscientious stewardship requires the negative as well as the positive news. Effective program officers often realize that their status as a funder works to insulate them from negative program reality, thus necessitating the outside perspective that evaluation can offer. An illustration of the need for outside perspectives is seen in Case 7.

Case 7: Telling 'Em What They Want to Hear

Recently, an evaluator of our acquaintance who works in the arts attended a meeting of grantmakers and grantees on the East Coast. At that meeting, one of the grantees asked the grantmakers, "Do you people realize how often we are just putting stuff on paper that you want to hear so that we can get funding to do what we wanted to do in the first place?" The arts funders were astonished by this statement and insisted it never happened to them. The alternative explanation is that they were insulated from program reality.

Several kinds of philanthropic practice could redress this problem. Certainly, exercising due diligence before the fact could go a long way toward establishing that grantees will do the work in line with the proposal. Consistent monitoring by the grant officer might also keep the grantee on track—or at least keep the foundation officers in touch with what could and could not be accomplished. However, evaluation might also help put foundation officers in touch with this reality and perhaps ensure that funding was used as intended. Monitoring, making site visits, and observing products could all provide for accountability.

Program Side Effects. Foundations' power can produce both intended and unintended consequences. Evaluation can play a useful role by documenting these consequences, especially the side effects of philanthropy's power that might be anticipated or controlled. Two examples illustrate how the power of philanthropy can produce these side effects, as seen in Cases 8 and 9.

Case 8: Evaluation of the Green Revolution

Rockefeller Foundation president Gordon Conway has written extensively about the Green Revolution, both its positive and negative effects (Conway and Ruttan, 1999). In 2003, Conway articulated what the foundation has learned from this experience and how it would incorporate that learning into a new approach to combat hunger in Africa. An important principle going forward is to involve communities, small farmers, and nongovernmental organizations in better formation of a "pro-poor" food policy.

The first Green Revolution offered farmers new crop varieties that allowed them to improve their agricultural yields—to grow more wheat and rice per acre. . . . The call to action was the prediction of massive famine for India in the 1960s. . . .

The Green Revolution has been criticized both rightly and wrongly, thoughtfully and not. . . . It lowered food prices for everyone, including both urban and rural poor. This was important because the poor spend the highest proportion of their income on food. Prior to the Green Revolution, half of the population in the less developed countries did not get enough to eat; today that percentage has fallen to one-fifth. The Green Revolution also greatly increased employment in the rural non-farm economy and, while patterns varied, agricultural wages generally rose. Across Asia, the annual income per person rose several fold, and the number of poor people dropped rapidly. In 1975, six out of ten Asians lived in poverty; by 1993, it was only two out of ten East Asians and four out of ten South Asians.

However, the Revolution had real failings. These failings—and I am using a very broad brush in this historical sketch—were both environmental and social. [The new crop varieties required more water, fertilizer, and pesticides and avoided the need for labor-intensive soil aeration, crop rotation, and working organic matter into the soil.] But these steps, we understand today, are key to long-term sustainability. The results of omitting them were soil erosion, nutrient depletion, falling water tables and salinization. . . . The seeds for the new varieties and the inputs required . . . were expensive, affordable mainly by the richer farmers (who by Western standards were still quite poor). The "rich" got richer and in some regions bought out the poor—to quote the critics, they "drove them off the land." . . . by the early 1970s many of the problems associated with the Green Revolution had been well documented and better understood. During the next two decades, governments in many parts of the world . . . put in place policies . . . to ensure that the Green Revolution might ben-

efit more of the poor. Village-level studies of the second-generation effects of the new technology in South India showed that small farmers had gained proportionally more income than large farmers, and the distribution of income improved. Nor, in most regions of India, did the disproportionate land distribution worsen. . . .

The world and the Rockefeller Foundation have changed considerably since the early Green Revolution days. We now understand the need to balance social and environmental concerns with getting new technologies right. . . . We are also trying to seize and develop the opportunities of globalization: *more transparency of process, more reporting of adverse effects in remote places, more informed criticism, and the possibility for faster and more precise course correction when the debate so indicates* (emphasis added; Conway, 2003, pp. 1–6).

Case 9: The Ford Foundation's Effort to Decentralize Governance of New York City Schools

This initiative worsened a growing rift between African Americans and Jews—groups whose historic solidarity promoted many civil rights achievements.

[In the late 1960s, the Ford Foundation was funding] experiments in structural change in public education, specifically the decentralization of the New York City school system. In 1967, McGeorge Bundy [president of The Ford Foundation] headed a panel appointed by Mayor Lindsay to make recommendations for school reform. Once the panel's report was published, the foundation gave support for two "demonstration districts" set up by the Board of Education in largely black areas, which in effect delegated authority over the schools in each district to newly formed community councils. The ensuing conflicts between the local councils, the Board of Education, and the teachers' union over questions of curricula and the right to hire and fire teachers led directly to the school strikes of 1968 and greatly exacerbated tensions between the black and Jewish communities. The role which the foundation, and Bundy in particular, had played in the whole affair drew great criticism. Albert Shanker, the head of the New York City teachers' union, even went on a nationwide speaking tour to warn other teachers' organizations of the threat to their position and power represented by "irresponsible intervention" of philanthropic organizations (Nielsen, 1972, p. 357).

One wonders whether prior consultation with the affected parties or formative evaluation early in the process would have apprised the foundation of the danger of side effects. From what we know, The Ford Foundation might not have changed its strategy. Yet the foundation's primary commitment was to the betterment of minorities. At least it would have known the potential tradeoffs for the civil rights movement, not to mention the public criticism this episode would bring down.

Revealing Flawed Assumptions. Evaluation is often highly useful when it reveals flawed assumptions (Leviton, 2003). A rather humble example is seen in Case 10. Testing assumptions can often help to preserve grant initiatives, make sure that money is well used, and preserve foundation reputations.

Why are such tests of assumptions necessary? Foundations often rely on a circle of elite experts they trust, and those elites often share assumptions about the world. These assumptions are occasionally wrong, or at least out of step with program stakeholders outside the circle. Philanthropy is not alone in this; erroneous but shared assumptions regularly cause both business and public policy mistakes (Allison and Zelikow, 1999; Mitroff, 1998). (This issue relates to the central point of Chapter Five: the benefits for philanthropy of being inclusive and listening to diverse voices when planning and making grants and when evaluating the consequences.)

Case 10: Engaging a Community's Leadership

In the early 1990s, the first author conducted evaluation for local funders, focused on a planning process to develop substance abuse services in three low-income, minority neighborhoods. The planning process was intended to be inclusive of community leadership in those neighborhoods. One of the neighborhoods had long been ignored by the city's power structure; it was the first time in a long time that the funders had sought this community's input on any topic whatsoever. However, the funders assumed they knew the community's leadership, and they clearly did not. An initial meeting erupted with resentment, as people shouted, "You don't even know who our real leadership is. Where are they? And why do you keep inviting people from [agency]—they don't represent us; they colonize us!"

The funders were greatly dismayed by this reaction. However, the evaluation put it in perspective for them. It pointed out that many good meetings in low-income communities can begin with residents' need to vent their frustration and hurt and to test whether the invitation to participate is made in good faith. The report also offered methods to better identify community leadership in the future. The process continued to be tense from time to time, but progress was made in planning for services, which are in place to this day. The evaluation report gave the funders some confidence to stick with the process in this neighborhood.

Philanthropy's Vulnerability

Foundations are vulnerable to criticism, whether deserved or not, both as a group and individually. Throughout the past century, both insiders and outsiders have called foundations elitist, overly privileged, unduly powerful, or sometimes simply wrong-headed (Brilliant, 2001; Cohen, 2002; Council on Foundations, 2003; Nielsen, 1972).

Consequences of Criticism. The justice of these criticisms is less important to our purposes than is the fact of foundations' vulnerability to damage to their reputations. The United States subsidizes foundations through favorable treatment under the tax law (Porter and Kramer, 1999). When foundations are criticized, this arrangement can be endangered. For example, in the early 1970s Congress created the Commission on Private Philanthropy and Public Needs "to clear the air after a string of scandals and reports of fiduciary misdeeds befouled the previously pristine reputation of American foundations in the late 1960s" (Cohen, 2002, p. ii; see also Brilliant, 2001). In 2003, the House of Representatives cited excessive executive compensation at one major foundation in proposing to exclude administration costs from the 5 percent IRS payout requirement. The bill did not pass but would have been a disaster for foundations that rely on in-house expertise and contract research to carry out their strategies (Foundation Center, 2004).

Foundation culture places great importance on the foundation's good name. It is difficult for outsiders to comprehend the importance of this consideration. Foundations rely on their "reputational" capital, as well as their financial resources, to shape programs and influence whole fields of endeavor. The good name of individual charities is difficult to come by and easy to lose. Foundation staff frequently allude to the fate of the United Way after its scandal of the early 1990s (Glaser, 1994; Sinclair, 2002). Charitable contributions to United Way increased only 6.9 percent between 1992 and 1996, although total private giving increased 24 percent during the same period (National Committee for Responsive Philanthropy, 1998).

Demonstration of Worth. Evaluation can provide important evidence of the value of philanthropy, helping to balance any criticism that a foundation might receive. Evaluation can demonstrate the worth of philanthropic activity in several ways, such as showing the value of a program design or intervention. However (and crucial to this discussion), evaluation is also valuable if it helps foundations avoid negative side effects and subsequent criticism through a midcourse correction (Chelimsky, 2001).

Grantees' View of the Evaluation Process

Although the focus of this chapter is on evaluation as perceived and planned within foundations, grantees will also benefit when evaluation has a sharper focus and a clearer purpose. Yet they will have understandable worries about a process that puts them in a position of relative weakness and potential loss of control.

Predictable Dysfunctions. Foundations' power translates into apprehension about evaluation among grantees. The need for trust and collaboration therefore carries over to the foundation's grantees and the evaluators who are commissioned to do

the work. All too often, we see the same dysfunctional reactions to evaluation that are so common in other settings. We have seen defensive grantees damage themselves through misbehavior; we have also seen evaluators harm their ability to generate useful information by rigidity in thinking and by a disregard for the grantee's legitimate anxiety. These situations convey the impression that a foundation is being high-handed with grantees.

Interdependence of Evaluation and Program. Realistically, foundation evaluation often has to depend on the cooperation of the program. This can work to restore some of the power balance in the relationship and establish a business-like exchange agreement between the grantees and the evaluator. After all, grantee staff must be interviewed to get a sense of process. Program participants may need to be interviewed to assess effects of the program. Even when data collection does not have to rely on grantee cooperation (as in the case of existing datasets), fairness dictates that the evaluation should reflect an understanding of the program logic model or theory of change, as well as reported barriers and facilitators for program implementation and achievement of outcomes. Therefore, the grantees need to be kept involved over the course of the evaluation. Besides, people have the right to be kept apprised of something so important to their fate. Fairness (and the paper trail demonstrating fairness) often provides the best protection from future attacks on the process or the results.

Conclusion

Evaluations that consider both the internal and external context of foundations have a greater chance to be of good quality, relevant, and useful. Expectations about evaluation's purpose and the ways it might be used also affect usefulness. Examples throughout this chapter illustrate both the potential and the reality of useful evaluation.

The Foundation Context

Inside the foundation, we have argued for the importance of three features that drive evaluation choices:

1. Foundation *size* (assets, payout, and staffing), which
 - *Does not* necessarily affect the quality or usefulness of evaluation (see Chapters Eight and Ten)
 - *Does* influence evaluation's purpose, because large foundations often have different program goals from the smaller ones (Cases 1, 2)

- *Does* affect foundation staff and grantees' sensitivities about evaluation (Cases 2, 3, 10)
- 2. Foundation *culture* (shared values and expectations), which
 - Dictates whether evaluation is valued (see Chapters Three and Eight)
 - Values some approaches to evaluation over others (Cases 1, 3, 4, 6)
 - Prizes evaluation findings (positive or negative) that address the foundation's shared mission and the foundation staff member's vision of the public interest (Cases 1–5)
- 3. Location of *authority and responsibility* for the evaluation function, which
 - Affects the independence of evaluation from program (Chelimsky, 2001)
 - Requires trust and collegiality between foundation program and evaluation staff to ensure a meaningful evaluation (Cases 1, 3, 10)

Outside the foundation, we believe three important evaluation drivers include

1. Foundations' *power* in society, which
 - May insulate staff from on-the-ground reality so that evaluations can offer information about problems and opportunities (Cases 7–10)
 - May lead to unintended side effects of grantmaking—side effects that evaluation might detect at an early stage for midcourse correction (Cases 8, 9)
 - May lead to untested assumptions for which a countervailing evaluation perspective is helpful (Cases 7, 9, 10)
2. Foundations' vulnerability to criticism, which
 - Drives the need to demonstrate the value of grantmaking
 - May lead foundations to rely on evaluation to avoid public mistakes (Cases 8–10)
3. Grantee *perceptions of the evaluation process* (fear, extent of buy-in), which
 - Can assist or detract from the work (Cases 3, 6, 10)
 - Can lead to acceptance or attacks on results (Cases 3, 4, 6)
 - Can be managed through a fair and transparent process (Cases 6, 10)

Some Constructive Uses for Foundation-Sponsored Evaluation

Inside the foundation, evaluations might be used in any of the following ways:

- Exercising due diligence by foundation staff to gain information in advance about the potential effectiveness or worthiness of a proposal (Cases 4, 5, 9)
- Reducing uncertainty about the investment in a program or strategy (Cases 4, 5, 8, 10)
- Correcting funders' assumptions about a societal need, a program strategy, or the political climate for social change (Cases 7–10)

- Keeping the foundation staff accountable to the board for their choices of grantmaking strategies or program funding (Cases 1, 3, 5, 8)
- Providing a rationale for renewal of funding to the foundation's senior leadership and board (see Chapter Seven; Cases 1, 4, 10)

Outside the foundation, evaluations might be used in these ways:

- Keeping grantees accountable for performance (Cases 2, 3, 5, 7)
- Portraying the rationale for foundation choices to the rest of the world (Cases 1, 2, 8)
- Building a field of work (see Chapter Nine; Cases 1, 8)
- Affirming or disproving the value of a program intervention to encourage outsiders' attention and support (Cases 1, 2, 4, 8)
- "Fixing" the wayward investment (Cases 1, 2, 7, 8)
- Providing the rationale for renewal of funding to the foundation's staff (Case 4)

The Road to Better Foundation-Sponsored Evaluation

Most important, perhaps, is the potential that evaluation presents to elevate the debate about program investments and grant awards. Because philanthropy is still so intensely personal, evaluation properly conducted can depersonalize the process. This can be done through an appeal to altruism and duty—the higher motivations behind many foundation officers' choice to work in philanthropy. It can be especially effective when staff view program grants as portfolios of work for which they have fiduciary responsibility. Viewed in that light, both positive and negative evaluations become important information to those who manage those portfolios. The information can be used to advance a field through knowledge and program improvement. Just as a positive evaluation can validate a foundation's strategy for improving a field, a negative evaluation can steer its future activities toward newer and more productive activities.

References

- Allison, G. T., & Zelikow, P. (1999). *Essence of decision: Explaining the Cuban missile crisis* (2nd ed.). Upper Saddle River, NJ: Pearson PTP.
- Bickel, W., & Cooley, W. W. (1986). *Decision-oriented educational research* (Evaluation in Education and Human Services, no. 11). New York: Kluwer Academic Publishers.
- Brilliant, E. L. (2001). *Private charity and public inquiry: A history of the Filer and Peterson Commissions*. Bloomington, IN: Indiana University Press.

- Chelimsky, E. (2001). What evaluation could do to support foundations: A framework with nine component parts. *American Journal of Evaluation*, 22, 13–28.
- Cohen, R. (2002). Introduction. In *The state of philanthropy, 2002*. Washington, DC: National Committee for Responsive Philanthropy (pp. iii–vii). Available at www.ncrp.org. Accessed June 17, 2003.
- Conway, G. (2003, March 12). *From the green revolution to the biotechnology revolution: Food for poor people in the 21st century*. Woodrow Wilson International Center for Scholars Director's Forum. Available at www.rockfound.org/documents/566/Conway.pdf. Accessed July 22, 2003.
- Conway, G., & Ruttan, V. W. (1999). *The doubly green revolution: Food for all in the twenty-first century*. Ithaca, NY: Comstock Publishing Association.
- Council on Foundations. (2003). *Council on Foundations, Inc. records, 1949–1981*. New York: Rockefeller University. Available at www.rockefeller.edu/archive.ctr/council.html. Accessed June 18, 2003.
- Dennis, M. L. (1994). Ethical and practical randomized field experiments. In J. S. Wholey, H. P. Hatry, and K. E. Newcomer (Eds.), *Handbook of practical program evaluation*. San Francisco: Jossey-Bass.
- Downs, A. (1994). *Inside bureaucracy* (reissue ed.). Prospect Heights, IL: Waveland Press.
- Emerson, J. (2003). *Total foundation asset management: Exploring elements of engagement within philanthropic practice*. Available at www.redf.org. Accessed May 20, 2003.
- Foundation Center. (2004). *Top U.S. foundations by total giving*. Available at fdncenter.org/research/trends_analysis/top100giving.html. Accessed May 27, 2004.
- Glaser, J. S. (1994). *An insider's account of the United Way scandal: What went wrong and why*. New York: Wiley.
- Hirschhorn, L., Kelly, M., & Gilmore, T. N. (2002). "Field" work in philanthropic strategy. Contribution to the project, Improving the Practices of Philanthropy, supported by The Robert Wood Johnson Foundation, The Ewing Marion Kauffman Foundation, The John S. and James L. Knight Foundation, and The David and Lucile Packard Foundation. Philadelphia: Center for Applied Research.
- King, M., & Oliphant, G. (2002). Tough love. *H: The Magazine of the Heinz Endowments*, 2, 10–25. Available at http://www.heinz.org. Accessed July 9, 2003.
- Leviton, L. C. (2001). Building evaluation's collective capacity: American Evaluation Association presidential address. *American Journal of Evaluation*, 22, 1–12.
- Leviton, L. C. (2003). Evaluation use: Advances, challenges, and applications. *American Journal of Evaluation*, 24, 525–535.
- Leviton, L. C., & Hughes, E.F.X. (1981). Research on the utilization of evaluations: A review and synthesis. *Evaluation Review*, 1981, 5, 525–547.
- Mitroff, I. I. (1998). *Smart thinking for crazy times: The art of solving the right problems*. San Francisco: Berrett-Koehler.
- National Committee for Responsive Philanthropy. (1998). *Charity in the workplace, 1997*. Washington, DC: Author. Available at www.ncrp.org/reports/charity97.htm. Accessed March 28, 2004.
- Nielsen, W. A. (1972). *The big foundations*. New York: Columbia University Press.
- Patrizi, P. (2002). *Briefing notes to the Evaluation II Roundtable*. Presented at the meeting of the Grant-makers Evaluation Network, Council on Foundations, Washington, DC, April 4, 2002.
- Porter, M. E., & Kramer, M. R. (1999). Philanthropy's new agenda: Creating value. *Harvard Business Review*, 77, 121–130.

- Rossi, P. H., Lipsey, M. W., & Freeman, H. L. (2004). *Evaluation: A systematic approach* (7th ed.). Thousand Oaks, CA: Sage.
- Sinclair, M. (2002, March 1). William Aramony is back on the streets. Parsippany, NJ: *The NonProfit Times*. Available at www.nptimes.com/Mar02/npt2.html. Accessed April 15, 2003.
- Stevens, B., Peikes, D., & Patel, S. (2003). Preliminary findings from the survey "When the Funding Stops." Princeton, NJ: Mathematica Policy Research. Manuscript submitted for publication.
- Szanton, P. (2002). *Toward more effective use of intermediaries*. A contribution to the project, Improving the Practices of Philanthropy, supported by The Robert Wood Johnson Foundation, The Ewing Marion Kauffman Foundation, The John S. and James L. Knight Foundation, and The David and Lucile Packard Foundation. Philadelphia: Center for Applied Research.
- Weatherford, J. M. (1985). *Tribes on the Hill: The U.S. Congress—Rituals and realities* (Rev. ed.). New York: Bergin and Garvey.