



Paper 6

The Role of Evaluation in the 21st Century Foundation

Edward Pauly

SHORT VERSION

January 2005

| BertelsmannStiftung



ABOUT THE PROJECT

The foundation sector grows at a considerable pace and is about to entail significant changes in the three-way relationship of private wealth, public policy and non profit action. To cope with the challenges arising from these transformations the International Network on Strategic Philanthropy (INSP) was established in spring 2001. With the underlying assumption that strategic philanthropy is more effective philanthropy, the network has striven to professionalize foundation management, convene the excellent minds of the sector, clarify the guiding values behind foundation activities, and contribute to capacity building in the field. The 68 members of the INSP are representatives of foundations and support organizations, consultants and researchers from the US, Europe and other countries of the world that operate along the lines of strategic philanthropy.

The network now presents a number of high-quality papers on a range of important subjects regarding strategic philanthropy. These include topics such as the role of philanthropy in globalization, new innovative instruments for philanthropy, promoting philanthropy, the role of evaluation in foundations and effective board management. The papers are available for free download at the INSP's Web site at www.insp.efc.be.

INSP is an initiative of the Bertelsmann Stiftung in collaboration with Atlantic Philanthropies, the Charles Stewart Mott Foundation, the Compagnia di San Paolo, the Ford Foundation, the German Marshall Fund of the United States, and the King Baudouin Foundation. Along with the Bertelsmann Stiftung, three institutions – The Philanthropic Initiative, Inc., The Hauser Center for Nonprofit Organizations at Harvard University, and The Center for Civil Society at the University of California Los Angeles – help to coordinate the work of approximately 70 network members.

The INSP working groups are advised and coordinated by representatives of leading academic and consulting institutions:

Prof. Helmut K. Anheier, Director, Center for Civil Society, University of California

Joe C. K. Breiteneicher, President, The Philanthropic Initiative

Prof. Christine W. Letts, Associate Director, Hauser Center for Nonprofit Organizations, Harvard University

For more information, please contact:

Dirk Eilinghoff
Philanthropy and Foundations
Bertelsmann Foundation
Carl-Bertelsmann-Strasse 256
33311 Gütersloh
(0049) – (0) 5241–81 81391
insp@bertelsmann.de

ABOUT THE BERTELSMANN FOUNDATION

The Bertelsmann Foundation is Germany's largest foundation established by a private donor. In keeping with the longstanding social commitment of its founder, Reinhard Mohn, the Bertelsmann Stiftung is dedicated to serving the common good by encouraging social change and contributing to society's long-term viability. To achieve this, it maintains an ongoing dialog with all of society's stakeholders. The belief that competition and civic involvement form an essential basis for social progress is central to the foundation's work. In order to apply its expertise as effectively as possible, the Bertelsmann Stiftung is structured according to subject areas. The foundation's 280 employees focus on Education, Health, Economics and Social Affairs, International Relations, Corporate Culture and Promoting Philanthropy.

CONTENTS

INTRODUCTION 4

EVALUATION AS LEARNING ABOUT RESULTS 4

LEARNING THE IMPORTANT LESSONS – WHY EVALUATION IS POWERFUL, CREATIVE, AND OFTEN PROBLEMATIC 5

THE STATUS OF EVALUATION IN FOUNDATIONS 6

DIFFERENT KINDS OF EVALUATION PRODUCE DIFFERENT KINDS OF LEARNING 10

PROMISING RESULTS OF EVALUATION, AND NEW EVALUATION TOOLS FOR PHILANTHROPY 11

FIVE PITFALLS OF FOUNDATIONS’ EVALUATIONS AND HOW THEY CAN BE AVOIDED 14

STANDARDS FOR EVALUATION IN 21st CENTURY FOUNDATIONS..... 16

BIBLIOGRAPHY 17

ABOUT THE AUTHOR 18

INSP TITLES 19

INTRODUCTION

One of the most visible changes in philanthropy in the last twenty years has been the increased attention paid by foundations to the results of the activities they support. In many cases, foundations seek to improve their results by using new tools – including strategic planning, operational planning, and evaluation. This emphasis on results is part of broad “foundation effectiveness” and “strategic philanthropy” movements that are global in scope. Expanding the use of evaluation, and being more thoughtful about evaluation, are significant responses to the new emphasis on the results of philanthropy.

Yet evaluations are often problematic for foundations, for several reasons:

- Foundations often decide to use evaluation without making an informed assessment of the particular kind of evaluation - among the quite different evaluation approaches that are available - that will serve their purposes. Choosing an evaluation approach that is ill-suited for the foundation’s needs creates problems.
- When the results of an evaluation differ from the claims made by the foundation’s staff and grantees, or when staff fear that evaluation results will differ from their claims, the evaluation is often seen as the source of the problem.
- Many foundations place a higher priority on directly supporting high-quality social service organizations than on achieving specific results and using evidence of results to improve the organizations’ - and the foundations’ - performance.

The effective use of evaluation by foundations that seek to increase their focus on results, and the benefits and challenges of using evaluation effectively, are the topics of this paper. A foundation can shape its use of evaluation according to these guiding questions: What does the foundation most need to learn? What are the distinctive features and contributions of the activities the foundation is examining? What is the new understanding the foundation seeks to obtain from an evaluation? What kind of evaluation best fits both the foundation activities being examined and the foundation’s learning goals? When foundations carefully determine what they need to learn, it is straightforward for them to determine how to use evaluation effectively.

EVALUATION AS LEARNING ABOUT RESULTS

Discussions of evaluation often stumble because people attribute different meanings to the word “evaluation.” Evaluation in philanthropy can be defined as the use of systematic information-gathering and research activities to learn about the results of foundation-supported activities. Thus,

evaluation is both systematic learning about the results of a foundation's work and it is applied research on foundation-supported projects.

Defining evaluation as learning is a departure from current thinking in many foundations, which often see evaluations as efforts to find out whether or not a foundation-supported activity "worked." Yet because foundations have different missions and goals, and support many different kinds of activities, they need to learn different things, and different learning goals require very different evaluations. Learning what the foundation most needs to learn sometimes requires an evaluation that examines how a project was managed, how services were provided, and who received the services; sometimes it requires an evaluation that assesses the outcomes of the project; and sometimes it simply requires the documentation of how the foundation's resources were used. These differing forms of evaluation produce different kinds of learning.

LEARNING THE IMPORTANT LESSONS – WHY EVALUATION IS POWERFUL, CREATIVE, AND OFTEN PROBLEMATIC

In order to learn the right lesson from the foundation's work, the foundation (and often its grantees, as well) must first decide what lessons they most need to learn. Learning important and useful lessons from a foundation project is sometimes difficult – particularly if the project is complex or innovative. Yet foundations' activities cannot achieve their goals, receive recognition for their accomplishments, be improved, and become more widely used unless accurate and useful information about their operational experiences and their results is readily available. Used thoughtfully, evaluations gather, organize, and analyze the information needed to learn important lessons about the projects supported by foundations.

The uses of evaluation for foundations can be quickly summarized:

- Evaluations identify important and useful lessons reflecting the varied contexts and the diverse goals of foundations.
- Evaluations provide practical information on ways to improve program operations, thereby increasing opportunities for foundations to obtain good results.
- Evaluations counterbalance the all-too-human tendencies for foundation staff (like other leaders) to indulge in overly optimistic "delusions of success" (Lovallo, Kahneman 1999).
- Evaluations build reliable and useful evidence about the performance and effects of the innovations supported by foundations.

Despite the benefits for foundations of carefully-designed evaluations, many foundations have found that the evaluations they commissioned were not useful, timely, flexible when a project changed direction, or relevant to the foundation's decision-making. These failures occurred either because the evaluation was conducted poorly, or because the foundations did not select the most appropriate

evaluation approach and failed to gather the lessons that would have been of value. In either case, a careful effort to decide what the foundation really needed to learn, and what evaluation approaches would enable it to learn what it needed to learn, would likely have produced an evaluation that benefited the foundation.

THE STATUS OF EVALUATION IN FOUNDATIONS

Globally, most foundation activities are not evaluated, and only a minority of foundations uses evaluation. Most foundations do not systematically gather lessons based on the results of their activities. However, foundations' use of evaluation appears to be increasing.

The status of evaluation in US foundations

Evaluation is much discussed among US foundations, apparently reflecting two historical roots: the business sector's emphasis on using and releasing credible information on performance, growth, and profits, and foundation trustees' application of these management practices to the foundations they oversee; and the government sector's extensive use of evaluation to assess and improve government services in the absence of market feedback mechanisms. Overall, the status of evaluation in US foundations is mixed. There are some hopeful trends, but neither the business goal of measuring performance nor the government goal of improving services has yet been achieved in more than a modest number of foundations.

Frequency of evaluation use

In 2004, the Urban Institute and Grantmakers for Effective Organizations reported the results of a survey of the 3,000 US foundations that employ a professional staff (most with assets of at least \$5 million), providing the best information to date on the use of evaluation by US foundations. The survey found that 44% of US staffed foundations reported that they formally evaluate work they fund, and 75% of foundations with more than \$400 million in assets reported doing so. US foundations hire a wide variety of professionals to conduct evaluations, including independent non-profit research organizations, consultants, university faculty and graduate students, and others. More than half of staffed US foundations reported that they frequently require organizations they fund to report on the outcomes of their foundation-supported work; 56% said they did this "always" or "often." For many foundations, "evaluation" refers to reports received from grantees on their work. Of the foundations that reported using formal evaluations, 16% said they make evaluation results public "always" or "often"; 47% said they never do so (Ostrower 2004).

Nature of the evaluations

Most of the evaluation work done by US foundations falls into one of two categories, and examples representing other approaches can also be found:

1. Plan-versus-performance (or grant monitoring) evaluations document how foundation funds were actually used and whether the activities originally planned when the project began were actually carried out. This is the most common evaluation approach, and it is used by a wide range of U.S. foundations. The information is usually gathered and presented by grantee organizations' staff. Since many grantee organizations lack evaluation expertise, these evaluations vary in their quality and reliability.
2. Outcome evaluations assess whether various kinds of results sought by program operators were achieved. These evaluations include the whole range of program evaluations, implementation studies, participation studies, cost studies, and other studies that examine what happened as a result of the foundation-supported activity and why it happened. The outcomes evaluations commissioned by foundations generally go beyond grantees' plan-versus-performance reports to examine not just the activities that were conducted, but the effects of those activities and the reasons the activities were conducted productively or not so productively.

Program evaluations seek to determine whether the intended results were achieved by the foundation-supported activity. They begin by identifying the outcomes sought by program operators; they gather information on the outcomes achieved by the foundation-supported activities; and they draw conclusions about the extent to which the desired outcomes were achieved.

Implementation evaluations examine operational topics such as how a program was carried out, the proportion of targeted people who actually received services, and the management challenges encountered during the program. Careful implementation studies can provide highly practical lessons on how to improve the service delivery, the identification and recruitment of participants, and the management practices of both the foundation and its grantees. Good implementation evaluations have the capacity to provide lessons that improve organizations' effectiveness and benefit people well beyond those directly served by a foundation-supported project.

3. Evaluations aimed at providing lessons for policy and practice. A modest but growing number of U.S. foundations' evaluations go beyond the program evaluation approach to capture practical or policy-related lessons from innovative foundation-supported projects. Implementation studies are often a valuable resource for policymakers as they seek to fund well-designed services and articulate quality standards for programs that receive government funding. Reliable evidence on the effectiveness of new approaches can spread lessons about effective practices to all of the organizations in a sector. Since many fields

supported by foundations lack the research funding and the technical expertise needed to develop new methods and programs, the systematic evidence provided by good implementation studies is particularly valuable for them. These lessons for practice and policy, provided by evaluations, are often the building blocks of improved performance by government, non-profit organizations, and private-sector organizations.

Assessment of U.S. foundations' use of evaluation

While evaluation is used by some U.S. foundations, very few of them use evaluations as a central part of their efforts to achieve their core goals. Implementation evaluations are rarely used to improve the effectiveness of non-profit organizations or to improve the effectiveness of public policies. Foundations that support significant innovations rarely use evaluation to determine the effectiveness of the innovation, how it can be implemented more effectively, how it can target people who will benefit the most from the innovation, and what the innovation would cost if it were to be operated at large scale. Instead, the U.S. foundations that use evaluation mostly do so to find out whether the foundation's projects worked as they were expected to do.

The status of evaluation in European foundations

Despite of the modest number of European foundations that use evaluation, the growing use of evaluation in Europe is significant because it is often directly tied to the foundations' core goals. Some European foundations use evaluation as a tool for learning important lessons to make improvements in their work and the work of grantees, as well as to contribute to public policy discussions. Among the European foundations that support the development and use of innovations, evaluation is sometimes used to learn about the implementation and effectiveness of these innovations. Finally, the pressure to demonstrate accountability, and to respond to regulatory agencies, has led some European foundations to use evaluation as a tool for demonstrating the implementation and outcomes of their activities.

These evaluations contribute to the dialogue between foundations and the other leaders of European civil society and government who are working on solving social problems, as well as the media. This link between European foundations' evaluations and the surrounding context of the evolving governmental and civil society system is quite dissimilar from the stimulus for evaluation in the U.S., with its roots in business management and government program evaluations.

Probing the status of evaluation in foundations in Russia, Africa, South America, and Hong Kong

The use of evaluation in Russia, Africa, and Hong Kong reflects a history of foreign governments' aid donations, which have often required evaluations. These evaluations provided assurances of financial accountability and evidence of whether desired results were achieved, and the same evaluation goals have been adopted by foundations in these countries. Evaluations are less likely to

be used to improve foundations' decision-making or to provide information and lessons to other organizations on what works and why.

Numerous foundations use evaluation in Russia, and evaluation practices are well established among foreign donors, community foundations, and private foundations. This practice builds on the history of foreign donors' evaluation requirements. The International Program Evaluation Network (IPEN21) has led information exchanges, a publishing program, a website, conferences, training, and consulting programs on evaluation. U.S. governmental aid activities in Russia also supported training in evaluation in the mid-1990s. Currently, evaluations are conducted by foundation staff members, indigenous organizations, and contractors with foreign funding, including the NGO Support Center and Process Consulting. A small proportion of these evaluations are made public. The history of evaluation in Russia has created a climate in which programs use evaluations to strengthen their work.

Among foundations in Africa, foreign donors have a long history of requiring the use of evaluations, but in general evaluation results are not made public. A striking example of the role of evaluations can be found in the work of the AIDS Foundation of South Africa, which directly experienced the mortal consequences of ineffective service delivery methods, along with great uncertainty about the effectiveness of alternative service delivery methods – underscoring the need for evaluation, and causing the foundation to establish a Research and Evaluation Desk early in 2002. Overall, the recognized need for effective programs appears to have promoted the use of evaluation in many parts of Africa, although capacity limitations have slowed the spread of evaluation.

Evaluation is used by a modest number of the larger foundations in Hong Kong. Corporate foundations in Hong Kong do not generally use evaluation, because their efforts focus on building linkages to community groups rather than on achieving particular service outcomes. In contrast, some private foundations and NGO fund-raising foundations are established users of evaluation, particularly those working on development issues, where evaluation systems are seen as part of the development field's normal management and program improvement practices. Some of these foundations provide evaluation training to their partners. A sizable fraction of these foundations publish the results of their evaluations. The extent of evaluation by Hong Kong foundations is particularly notable because there are deep-seated challenges to evaluation there. These include "guanxi giving" (philanthropy that is highly responsive to social connections); the tradition of using glamorous events, such as dinners and galas, to finance civil society organizations; and a general avoidance of visibility by foundations.

In Uruguay, only a few foundations appear to use evaluations, and the few evaluation results are rarely made public. There has been considerably more use of evaluation by foundations in Argentina, where a sizable proportion of the larger foundations report that they support or commission evaluations. Many of these evaluations are made public, consistent with the considerable

transparency of foundation activities in Argentina. The history and traditions of the foundation sector play a major role both in Uruguayan foundations' limited use of evaluations, and in the greater and more transparent use of evaluations in Argentina.

DIFFERENT KINDS OF EVALUATION PRODUCE DIFFERENT KINDS OF LEARNING

Different foundations have very different needs for evaluations, because they have different missions, strategies, assets, and contexts. The experiences of many foundations show that there are well-developed evaluation approaches that can produce practical learning that is tailored to foundations' differing goals. Six broad kinds of evaluation approaches can contribute in different and distinct ways to foundations' work.

1. Monitoring

Evaluations for monitoring document how the foundation's funds were used, which planned activities and milestone accomplishments were achieved, and how much progress was made toward reaching the goals of the foundation-supported activity. Monitoring is typically regarded as an essential part of philanthropic due diligence (making sure that the foundation's resources were used for their intended purposes). A major use of monitoring studies is to increase projects' accountability to the foundation.

2. Implementation evaluations

These evaluations analyze the operational feasibility of a project, the organizational changes that are necessary to support the project, the incentives for organizations and individuals that are necessary so that the project can be carried out, the practical challenges faced by the project and its staff and how those challenges can be overcome, and the role of affected communities in enabling the project to succeed. A major use of implementation evaluations is to improve operational success, both for the foundation-supported program and, importantly, for other organizations doing similar work.

3. Impact evaluations

These evaluations seek to provide evidence on the effects of a foundation's project on its intended beneficiaries. This approach is useful for determining whether a particular kind of activity is capable of producing a specified result; when this is already known, it is often preferable to use implementation evaluations to make sure the program is being operated in accordance with the impact evaluation's findings.

4. Participation evaluations

Information about participation is important when service providers need to find out who benefits from the foundation's project: which people were affected, how many people were affected, and

what proportion of the intended beneficiaries were affected. This approach is useful for improving the targeting of services and deciding whom to recruit and how they can be recruited.

5. Cost evaluations

These studies gather and analyze cost information on the foundation's project, particularly information on the costs of expanding the project after the initial development costs have been amortized. This approach is useful for making decisions on financing and budgeting.

6. Evaluations of the logic and the state of knowledge and practice on a specified topic

There is often a great need for lessons about the broad state of knowledge on which programs and services are based, including knowledge about the causes of key outcomes and how best to produce these outcomes in specified contexts and situations. These studies are very valuable for strengthening the design of a foundation's programs, and they make a special contribution to the design of innovative programs.

As these six categories show, evaluation refers to many different kinds of learning. Consequently, the central question in designing a foundation's evaluation is, what do the foundation's leaders most need to learn so that the foundation and its grantees can achieve the foundation's goals? Foundations of many different sizes and types can use the tools offered by evaluation to learn about the results of their work. Even a small foundation can use evaluations effectively, for example by finding ways to strengthen a program's methods, or by assessing how the context of a project shapes its effects. An evaluation that uses just one of the approaches described here and does so effectively will accomplish more than a large, expensive evaluation that is broad and unfocused or uses a purely generic evaluation approach.

PROMISING RESULTS OF EVALUATION, AND NEW EVALUATION TOOLS FOR PHILANTHROPY

Evaluations have already produced a great deal of solid and practical knowledge drawn from the results of foundations' work. These evaluations have been put to use by practitioners, policymakers, and other leaders, as well as by foundation trustees and staff. These evaluations have replaced assumptions with evidence, have used new knowledge to stimulate discussion and debate, have solved practical problems, and have proved the feasibility of innovative programs. The following examples demonstrate the kinds of benefits foundations have obtained from evaluations and how their evaluations have contributed to society.

Evaluations improve performance and promote accountability through monitoring

Evaluations are one way, but not the only way, for foundations to find out whether their grantees are carrying out their work as planned and whether key activities have been completed. Monitoring is often used as the basis of "midcourse corrections" when a project encounters problems, and as

input for decisions on whether or not to renew funding. Monitoring is conducted in “real time,” that is, soon after the planned events were supposed to occur and long before the “final results” of a project can be assessed. Real-time monitoring makes it possible to intervene if a problem is identified. However, when the monitoring news is bad, conflicts can arise between those seeking accountability for results, and those counseling patience and flexibility in dealing with challenging problems.

The Wallace Foundation’s large project on improving the effectiveness of state and local school leaders in increasing students’ learning serves as a case in point. This project seeks to support innovation in the laws and practices of the states within the U.S. When the foundation’s program officers monitored the grants to 15 states, they found that little innovation appeared to be occurring in most of the states, partly because the states’ governors and other high officials – those with the authority to make change happen – were not involved. The states’ activities were limited in scope and importance, and lacked a sharp focus. Based on their monitoring, the foundation’s staff designed a second phase of the project, with more ambitious goals for the states; new kinds of technical assistance; and they added several states (chosen for their use of major innovations) to the project. These were major shifts, with larger roles for senior leaders, greater use of experts, and clearer plans for achieving results. As this example shows, it is through monitoring that foundations frequently take action to increase the likelihood that their work will produce significant benefits for society.

Evaluations provide important and action-oriented lessons about implementation

Foundations’ implementation evaluations have produced powerful evidence on the feasibility of innovative programs, how they can be operated more effectively, and how newly-designed programs can be refined to improve their performance. These evaluations provide powerful lessons on how to improve management, organization, staffing, incentives, and service design. In the 1960s, the widespread use in U.S. elementary schools of innovative mathematics curricula was not effective in helping students learn math. Implementation evaluations revealed that many teachers had not received adequate training in how to use the new curricula, resulting in poor-quality instruction. These evaluations showed how teachers could be trained to use the curricula well, and resulted in improved learning for students. These implementation evaluations triggered major improvements and expansions of effective training for teachers.

Evaluations provide powerful evidence about impacts

The significance of impact evaluations is that they provide credible, actionable evidence on whether or not a program or service should continue to be used. In 1995, the foundation-supported impact evaluation of adult mentoring for “at risk” children 10 to 16 years old found that the mentoring clearly reduced the young people’s use of drugs and alcohol, reduced their use of violence, and increased their school performance and attendance. The evaluation report, “Making a Difference: An

Impact Study of Big Brothers Big Sisters” by Public/Private Ventures, led to increased financial support for mentoring and encouraged more adults to become mentors. This is an example of how scarce social resources can be used more effectively when they are allocated using well-designed and well-implemented impact evaluations.

In general, impact evaluations are appropriate only for relatively established, stable programs. If an impact evaluation is conducted prematurely, it can lead to the misleading conclusion that the approach studied was a failure, when the truth is that it was simply not well implemented. The challenge for foundations is to understand when an impact evaluation is called for – and when another evaluation approach, such as an implementation evaluation, is more appropriate. Rushing to do an impact evaluation in the hope of discovering a “magic bullet” is frequently a recipe for disappointment.

Evaluations provide practical information about who participates and who benefits, thereby improving the targeting of services

Evaluations that examine participation issues have often been extremely valuable in determining whether the intended beneficiaries were participating in a foundation-supported program, how many of them (and what proportion of the intended target group) participated, and whether they received sufficient services to enable them to benefit from the services (the “dosage” question). For instance, the evaluation of Career Academies (secondary school programs that combine an academic emphasis with family-like support and career preparation) showed that the program attracted many students who did not need assistance in completing secondary school. The evaluation identified the importance of changing the eligibility and admission rules for Career Academies, so that scarce resources could be used for the students in greatest need of these programs.

Evaluations provide crucial information on costs

In The Wallace Foundation’s Pathways to Teaching Careers program, carefully-selected teacher aides received scholarships to complete college to become teachers for high-need schools. An evaluation of their teaching performance showed that they were slightly more effective than conventionally-prepared teachers, and they had greater job retention in hard-to-staff schools. The evaluation determined that after the development costs of the Pathways program were set aside, the operating cost was in the range of \$8,000 - \$13,000 per participant. This was considerably less than the foundation’s gross cost per participant (which included the development costs), demonstrating the importance of a careful cost analysis. The high job retention rate for Pathways teachers made these operating costs attractive to many school districts that have high turnover. Moreover, having clear and reliable evidence on costs was enormously valuable for policymakers and other potential users of Pathways. When evaluations provide reliable cost information, the resulting clarity about the amount of scarce resources needed to achieve a social benefit is a major contribution to public policy and large-scale social change.

Evaluations assess the logic and the knowledge underlying the design of new programs

After years of support for arts organizations' efforts to build greater public participation, The Wallace Foundation sought to understand the logic and effectiveness of the different approaches that were used. The evaluation took an innovative approach: it assessed the logic of the projects, by analyzing how people decide whether or not to participate in the arts. The study, "A New Framework for Building Participation in the Arts" by the Rand Corporation, showed how arts organizations can attract people who are initially disinclined to participate (by changing their perceptions through welcoming connections with their neighbors); people who are inclined to participate but are not doing so (by overcoming the practical barriers they face, such as parking, scheduled hours of service, or ticket prices); and people who are already participating (by deepening their experience of the artistry, through education and new programs). The evaluation broke new ground by logically connecting arts organizations' participation-building efforts with the needs of individuals as they decide whether to participate.

As these examples show, there are many ways that evaluation has already provided practical lessons that serve the strategic purposes and goals of foundations. The examples also make plain the fundamental value of evaluation for 21st century philanthropy: they create reliable information and lessons that provide significant social benefits and advance foundations' missions.

FIVE PITFALLS OF FOUNDATIONS' EVALUATIONS AND HOW THEY CAN BE AVOIDED

Foundations' evaluations are likely to fail when their context, design or execution reduces or undermines effective learning about the results of the foundation's work. This can happen in several ways:

Evaluations that do not seek significant learning

All foundation's evaluations are really valuable only if they address the most important learning needs of the foundation. All too often, foundations conduct evaluations that seek merely to provide a general impression of whether the foundation-supported activity did what it was supposed to do, or to ensure compliance with the foundation's expectations. In contrast, the foundations that benefit most from evaluations use them to find the answer to their burning questions, such as: What changes in the design of an innovation will provide the biggest boost to its effectiveness? Is the program serving the people who will benefit from it the most? If an innovation is used widely, what will it cost? What can be learned from the foundation-supported program that will show program operators how they can radically improve their work?

Low quality evaluations

All too many evaluations are untimely, inflexible, and unable to provide useful findings. These are frequent results of a “generic” program evaluation design (one that assesses whether a project’s original goals were achieved), rather than a more carefully designed implementation evaluation, participation study, cost evaluation, or an assessment of the logic and the knowledge base related to a foundation effort. A badly designed or badly executed evaluation is as lacking in value as an evaluation that fails to seek important learning.

The fear of evaluation and the need for sensitivity about future funding

Many foundation staff members and grantees fear that negative evaluation results will severely damage a project and the reputations of its supporters. Evaluators are often viewed as police officers, enforcing compliance with predetermined plans and punishing those who produce disappointing results. When evaluation is seen as policing, staff resistance and conflict with program operators is likely to follow. The most useful and productive role for evaluation is not the policing role – it is the learning role. Good evaluations are always sensitive to the concern that the evaluation’s findings will lead to the termination of funding or to a reputation for poor performance that could damage a grantee’s future fund-raising prospects. Instead of punishing innovative programs, evaluations that focus on implementation, costs, participation, or the logic and knowledge base typically identify valuable ways to improve the program’s performance. This approach addresses the concerns of program operators and foundation leaders by emphasizing the gradual and step-by-step processes of program improvement and capacity-building.

Excessive focus on specifying outcomes

Increasingly, foundations require grantees to specify their project’s desired outcomes before the project begins, as a management tool that is unconnected to learning but instead seeks compliance with the agreement between the foundation and the program operator. This approach has little to do with evaluation. If it has value, it would seem to be mostly when the foundation is paying for a well-defined product (such as might be purchased from a vendor) – for example, the construction of a new building, the hiring of specialized staff for a grantee, or the execution of a well-understood task (such as conducting an audit or drawing a new map of a specified area). Recording outcomes is no substitute for learning nuanced lessons about the results of a foundation’s activities.

Premature evaluation of innovations

If an innovation is evaluated before it has been fully implemented, important results and accomplishments will be neither observed nor appreciated. Innovations almost never unfold strictly according to a plan. Some foundation leaders fear that evaluations will result in the suppression of innovation and risk-taking, because they assume that the inevitable uncertainties of innovations will cause negative evaluation results. It is not evident that this has ever happened. But whether or not

this potential risk represents a real problem, it seems clear that the program evaluation model – in which the evaluation seeks to determine whether or not the original goals of a project were achieved – is a poor choice for evaluating most innovations, particularly early in their development. It is implementation evaluations, not conventional program evaluations, that are best suited to learn how an innovation unfolds and what results, if any, it may be able to achieve. Indeed, the lessons provided by high-quality implementation evaluations are likely to provide irreplaceable lessons that can accelerate the development and spread of an innovation, by systematically building up a rich store of evidence and understanding of innovative practices and outcomes. Innovations spread when their results have been carefully described and understood (Rogers 2003).

A consistent pattern emerges from an examination of these pitfalls. Foundations can avoid the major problems that have limited the usefulness of evaluations by determining what they most need to learn and using their evaluations to obtain those lessons.

STANDARDS FOR EVALUATION IN 21st CENTURY FOUNDATIONS

Foundations can provide valuable lessons for society, and they can provide evidence of the value to society of their work, by using evaluations. These benefits come directly from the harvesting by foundations of the most important learning that results from their work. The experiences of the foundations reviewed in this paper point to two critically important standards for evaluation in 21st century foundations:

- Foundations should have a learning plan, and should gather relevant information to capture the most important lessons from the results of their work.
- Foundations should make public significant information regarding what they have learned about the results of their activities.

These standards reflect foundations' compelling need for accurate information about the results of their work. It is a matter of great importance for foundation leaders to take the steps required to learn about and evaluate the results of foundations' work. This is the reason evaluation and the information it produces are arguably essential for responsible philanthropy. The same professionalism that requires careful stewardship of foundations' financial assets and systematic reviews of staff performance, requires foundation heads and trustees to use evaluation to gather information on the results and the lessons of foundations' activities.

BIBLIOGRAPHY

Rogers, Everett, *Diffusion of Innovations*, 5th edition, New York: Free Press, 2003.

Ostrower, F., *Attitudes and Practices Concerning Effective Philanthropy*, Washington, D.C.: Urban Institute, 2004.

Lovallo, D., and D. Kahneman. Delusions of Success. *Harvard Business Review*, 1999.

ABOUT THE AUTHOR

Edward Pauly

The Wallace Foundation

As The Wallace Foundation's Director of Evaluation since 1996, Edward Pauly leads the foundation's efforts to gain practical knowledge from the Foundation's work. The Wallace Foundation's evaluations – all of which are made public – seek to provide powerful findings that will be widely used by leading organizations and policymakers. Dr. Pauly and his staff have developed major evaluation projects on education, leadership effectiveness, youth development, after-school programs, adult literacy, and public participation in the arts and culture. Previously, he was the first coordinator of education research for the Manpower Demonstration Research Corporation (MDRC). His books include *The Classroom Crucible: What Really Works, What Doesn't and Why* (Basic Books), *Homegrown Lessons: Innovative Programs Linking School and Work* (Jossey-Bass), and, with Judith M. Gueron, *From Welfare to Work* (Russell Sage Foundation). Dr. Pauly received his Ph.D. in political science from Yale University. The Wallace Foundation is a private independent foundation based in New York City.

INSP TITLES

The Role of Philanthropy in Globalization
Helmut Anheier, Adele Simmons

Promoting Philanthropy: Global Challenges and Approaches
Paula D. Johnson, Stephen P. Johnson, Andrew Kingman

Effective Foundation Boards – The Importance of Roles
Christine W. Letts

Theories of Change
Theories of Social Change: Background Paper
Diana Leat
The INSP *Theory Of Change Tool* Manual

Philanthropy Program Design
Richard Mittenenthal

The Role of Evaluation in the 21st Century Foundation
Edward Pauly

Who Comes to the Table? Stakeholder Interactions in Philanthropy
Tomas E. Backer, Ira Barbell, Ralph Smith

Management Practices surrounding Program Professionals in US and European Foundations
Nadya K. Shmavonian

Innovations in Strategic Philanthropy – Lessons from Africa, Asia, Central and Eastern Europe, and Latin America
Helmut Anheier, David Winder

Tools for Good: A Guide to Vehicles for Philanthropy and Charitable Giving
Paula D. Johnson and Stephen P. Johnson