i n s p

Paper 6

The Role of Evaluation in the 21st Century Foundation

Edward Pauly

FULL 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

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. This is because:

- Evaluation is not a single, homogeneous commodity; instead, there are many different kinds of evaluation, each one suited for different purposes, and foundations often decide to use evaluation without making an informed assessment of the kind of evaluation that will serve their purposes;

- When foundations use an evaluation approach that is inappropriate for their needs, the evaluation often creates problems and disappointment;

- When an evaluation's results differ from the claims made by the foundation's staff and grantees, the evaluation (rather than the unexpected performance result) is often seen as the source of the problem;

- The widespread anxiety of many foundation leaders and staff members regarding evaluation problems, and the possibility of negative evaluation findings, are significant barriers to the effective use of evaluation; and

- Many foundations place a higher priority on supporting high-quality social service organizations through contributions of money and other resources than on achieving specific results and basing their decisions on evidence of results.

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. Using evaluation effectively to illuminate the results of a foundation's activities requires the foundation's leadership to answer to the question, What does the foundation most need to learn? Based on the answer a particular foundation gives to this question – and different foundations will give very different answers to it – the foundation can shape its use of evaluation according to three guiding themes: *What are the distinctive features and contributions of the foundation activities being examined?*, *What is the new understanding the foundation seeks to obtain from an evaluation?*, and *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, and when they understand

the distinctive features of the activities about which they seek to learn, it is straightforward for them to determine how to use evaluation effectively.

Evaluations can strengthen foundations' accountability, management, understanding of the results of their work, and their credibility in disseminating powerful lessons about those results. Evaluations, used with the judgment foundations expect of all their work, are an essential and powerful tool for making foundations' work effective – that is, for enhancing the results of foundations' work.

Indeed, without the effective use of evaluation by foundations, it is difficult to imagine that foundations will achieve credibility for their claims of significant accomplishments worthy of the attention and use of other parts of society.

EVALUATION AS LEARNING ABOUT RESULTS

Nearly everyone thinks they know what "evaluation" means – that is, nearly everyone has a particular kind of evaluation in mind when they think of the word "evaluation" – but discussions of evaluation often stumble because the word evaluation has many different meanings. Some brief comments outlining a broad definition of evaluation may avoid problems later.

Latin roots suggest that evaluation means "taking out the value" (that is, the worth), or "finding the value." In order to pull out or assess the value of their work, foundations use evaluation to learn what happened – what resulted – through their activities (both grantmaking and operations).

- *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.[1]

Thus, evaluation is both *systematic learning about the results of a foundation's work* and it is *applied research* on foundation-supported projects. In contrast with other kinds of research, evaluation seeks to be useful; it is not "pure research" conducted for the sake of research. Peter Senge, author of *The Fifth Discipline*, has said that evaluation requires researchers who are sincerely interested in the world of practice and are highly respectful of that world, and practitioners who are sincerely interested in research and seek to use it to better their work.

Defining evaluation as learning is a departure from the ways many foundation leaders think about evaluation. Many people in foundations think evaluations are efforts to find out whether or not a foundation-supported activity was carried out as it was originally planned. This definition is far too narrow to be appropriate for the many different kinds of foundations, the varied goals of foundations, and the enormous diversity in the projects supported and conducted by foundations. Experience demonstrates that it is not useful for foundations to conduct a formal evaluation of every

---

[1] Based on the work of Michael Quinn Patton, Peter Rossi, and others.

project according to an expectation that the project should correspond to the plans outlined in the original proposal; this approach makes sense for some foundation projects, but not for all. And even more importantly, this approach misses many opportunities to obtain valuable lessons, insights, tools, experiences, and discoveries by learning from foundation-supported work.

Because foundations have different missions and goals, and consequently support and participate in many different kinds of activities, they need to learn different things – and different learning goals require very different evaluations. There is no single variety of evaluation that will enable a foundation to achieve its goals. It follows that different kinds of evaluation are needed by different foundations. In addition, whatever the kind of learning that is sought by a foundation, an evaluation that is not thoughtfully designed and carefully implemented is unlikely to produce the well-grounded, reliable findings that are the basis of learning. In other words, foundations' evaluations must combine effective execution with a carefully-designed focus on the particular kinds of learning that are most relevant to the foundation's goals.

Foundations use evaluation in different ways, and the differences matter. For example:

- Grant monitoring, or plan-versus-performance, evaluations document how grant funds were used.
- Outcome evaluations yield lessons about the results achieved through the use of a particular service.
- Other kinds of evaluation studies provide information on topics such as innovative ways to deliver services, overcoming people's reluctance to participate in services, effective ways of obtaining financing for a new service, and lessons for policymakers and program operators.

These differing forms of evaluation produce *different kinds of learning,* with major consequences for foundations' work, as this paper will show.

LEARNING THE IMPORTANT LESSON – WHY EVALUATION IS POWERFUL; CREATIVE; AND OFTEN PROBLEMATIC

It is not difficult to learn about the results of many foundation projects; foundation staff members need only ask their grantees (or other partners), What happened? The answers they hear are often straightforward ones: The money was spent as planned, or it was not; the proposed tasks were carried out, or they encountered problems and were changed; services were provided, or were not provided; the services were given to people in need of them, or to others, or to no one.

If they wish, foundations can witness the unfolding of the projects they support. By witnessing a project, or by gathering information from witnesses, they can find out what happened in considerable detail.

Often, a foundation needs to answer a different question than "How did the foundation-supported project unfold, and how was the foundation's money spent?" For example, the foundation may need to find out how the management of the foundation-supported activity can be improved; or how to attract more participants (or more participants with particular needs); or how to build more community support for the activity; or how to reduce the activity's cost (or increase its revenues). These examples point to the differences in the contexts of foundations and the organizations they support, as well as the different challenges that must be overcome in different environments. In order to learn the right lesson from the foundation's work, the foundation and its partners must first decide what lesson they most need to learn – and that lesson will only rarely be about how the foundation's money was spent.

Learning important and useful lessons from a foundation project is sometimes difficult – particularly if the project is innovative, or complex, or complicated. Anheier and Leat (2002) have pointed to foundations' support for innovations as a distinctively valuable niche for foundations, one that they are uniquely capable of filling.[2] Yet innovations cannot spread and undergo additional development and improvement unless accurate and useful information about their results is readily available, and evaluations can fill this role. They provide reliable information on important topics, combined with analysis and assessments of the significance of the findings. Used thoughtfully, evaluations can gather, organize, and make sense of the important lessons about the results of projects supported by foundations.

Without evaluations, foundation staff can only gather anecdotal accounts of activities, reports of expenditures, and the documentation of a few events. There are times when this kind of information is useful and even important (such as for assuring that foundation funds are used for appropriate philanthropic purposes and with responsible, conscientious management), but it rarely meets foundations' need for lessons that can be used to improve the foundation-supported project, other projects, other organizations' practices, or policymaking.

Despite these benefits of evaluation for foundations, it is not widely used. In their widely cited article in *Harvard Business Review*, "Philanthropy's New Agenda: Creating Value," Porter and Kramer (1999) wrote, "The overall failure to evaluate the results of foundation grants is the most telling danger signal of all. Almost no money is set aside for program evaluation. Many foundations are ambivalent about whether funds should be spent on evaluation and whether assessing the performance of past grants can improve future grant making. This ambivalence about evaluation is reinforced by the performance criteria used to judge foundation staff. These tend to emphasize the paper trail of pre-grant analysis and recommendations and give little credit for achieving the real-world results that motivated the grant in the first place. Program evaluation, therefore, has only a downside: failure risks censure, but success adds no reward. … Without evaluation, a foundation

---

[2] Anheier, H., and D. Leat, *From Charity to Creativity: Philanthropic Foundations in the 21st Century*, Stroud, England: Comedia, 2002.

will never know whether or not it has been successful. The most basic premise of strategy – striving for superior performance – is violated if performance is not measured."[3]

In other words, Porter and Kramer argue that if foundations fail to conduct evaluations, their ability to obtain high-quality results is deeply jeopardized.

One very significant reason for foundations' limited use of evaluation is suggested by Lovallo and Kahneman's study of the tendency in businesses to avoid skepticism. Their analysis is entitled "Delusions of Success" (2003).[4] They take note of the fact that three quarters of business initiatives (start-ups, new manufacturing plants, mergers and acquisitions) flounder, and they attribute the predominance of failed vision to "delusional optimism": business leaders over-emphasize projects' potential benefits and under-emphasize likely problems, due to an all-too-human (and organizational) bias toward highlighting the positive. In the for-profit economy, economic results are a partial counterweight to delusions of success; good results encourage the continuation of effective problem-solving, and bad results discourage the "delusional" over-optimism that ignores problems rather than solving them. In the philanthropic sector, the evidence provided by evaluations may be a potentially significant counterweight to the powerful psychological mechanisms that produce undue and inaccurate fantasies of success, by providing information and lessons on what works and why. However, foundations do not use evaluations enough to make this counterweight an effective one, and this is one of the reasons that over-optimism and inadequate attention to results continues in the foundation world.

Another critically important use of evaluation is to gather rigorous and persuasive evidence of successful results from effective foundation-supported innovations. The importance of this role is underscored in Anheier and Leat's book *From Charity to Creativity: Philanthropic Foundations in the 21st Century* (2002). They argue that the unique contribution that foundations can make to society is to support the development and introduction of "truly new approaches based on creative thinking and working that go beyond existing professional, departmental, and organizational boundaries and mind sets." For these innovations to benefit society, *learning* about their implementation and results is needed, and that is where evaluation comes in. Evaluation is the critically important vehicle for "setting out systematically to analyse and learn from failures and successes. Instead of treating evaluation and analysis of current and past grants as an expensive luxury, foundation cultures need to define these as necessities. Instead of worrying about spending too much on learning and evaluation, foundation cultures need to encourage anxiety about the costs of spending too little." Without evaluation, the opportunity to share the benefits of important innovations widely is, for practical purposes, lost – as if they were isolated and easily forgotten experiments whose results were never shared with the community of potential users.

---

[3] Porter, R., and M. Kramer, "Philanthropy's New Agenda: Creating Value," Harvard Business Review, 1999.
[4] Lovallo, D., and D. Kahneman, "Delusions of Success," Harvard Business Review, November 2003.

This brief review of 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 for foundations, their grantees, and others to obtain improved results.
- Evaluations counterbalance the all-too-human tendencies for foundation staff (like other leaders) to indulge in delusions of success.
- Evaluations build reliable and useful evidence about the performance and the effects of the innovations supported by foundations.

It is worth noting that conventional discussions of evaluation in philanthropy place relatively little emphasis on these benefits (with the partial exception of the benefit of identifying ways to obtain good results). Instead, they generally see evaluation as an elaborate way to "close the loop" on foundation-supported projects by gathering information on whether the original plan was carried out.[5] No wonder most foundation leaders place a low priority on evaluation!

Despite the benefits for foundations of carefully-designed evaluations, it is important and sobering to recognize that many foundations have found that the evaluations they commission have not been useful, timely, flexible when a project changed direction, or relevant to the foundation's decision-making, as the evaluation directors of large U.S. foundations reported to Patrizi and McMullan.[6] In other words, the promise and the benefits of evaluation have often eluded foundations that have used evaluation. 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 garner the lessons that would have been of value. In either case, a careful effort to determine what the foundation really needed to learn, and what evaluation methods would enable it to learn what it needed to learn, would likely have produced and evaluation that benefited the foundation.

If evaluation is to strengthen the work of foundations, this pattern of ineffective evaluation design and use must change.

THE STATUS OF EVALUATION IN FOUNDATIONS

Most foundation activities are not evaluated. Even more important, most foundations do not seek to learn important lessons from their activities; rather, they focus on initiating the next round of new foundation-supported projects. Foundations place a high priority on developing new projects and

---

[5] For example, Council on Foundations, *Evaluation for Foundations: Concepts, Cases, Guidelines, and Resources*, San Francisco: Jossey-Bass Publishers, 1993.
[6] Patrizi, P., and B. McMullan, "Evaluation in Foundations: The Unrealized Potential," W. K. Kellogg Foundation, 1998 (www.wkkf.org).

new grants, often leaving little staff time for follow-up and evaluation. The benefits of evaluation for gathering the lessons needed to achieve the foundation's goals, improving the foundation's performance, counterbalancing over-optimism, and determining the results of significant innovations are thus lost. In an overview of the challenges of outcome measurement and evaluation in U.S. foundations, Walker and Grossman wrote, "The overwhelming majority of projects, organizations, initiatives and programs supported by philanthropy [in the U.S.] are not formally assessed for outcome achievement. Many do not even generate basic descriptive information about the content, quantity, and quality of what they do, much less assess what they accomplish."[7]

In the U.S., Europe, and other nations, the available information suggests that relatively few foundations use evaluation. At the same time, there is also considerable evidence that the use of evaluations is growing and that evaluations are becoming more closely adapted to the needs of foundations.

Within these broad outlines, the use of evaluation varies greatly among foundations and regionally, as the following overview shows.

A. The status of evaluation in U.S. foundations

Most of the published reports on foundations' use of evaluation are from the United States. Evaluation is much discussed among U.S. foundations, apparently reflecting two historical roots: the business sector's emphasis on using and publicly 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 improve government services in the absence of market feedback mechanisms. Overall, the status of evaluation in U.S. foundations is mixed. There are some positive 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 U.S. foundations.

Frequency of evaluation use.

A substantial proportion of U.S. foundation leaders believe that evaluating the activities supported by their foundation is an obligation of good professional practice – although the meaning of this belief is unclear, in the absence of an agreed-upon meaning of the term "evaluation." In 2003, the Urban Institute and Grantmakers for Effective Organizations reported the results of a survey of the 3,000 U.S. 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 staffed U.S. foundations.[8] Many of these foundations reported that they gather some sort of evidence on the outcomes of

---

[7] Walker, G., and J. Grossman, "Philanthropy and Outcomes: Dilemmas in the Quest for Accountability," Public/Private Ventures, 1999; www.ppv.org.
[8] Ostrower, F., "Attitudes and Practices Concerning Effective Philanthropy," Washington, D.C.: Urban Institute, 2004; www.urban.org.

foundation-supported activities; 44% of U.S. staffed foundations reported that they formally evaluate work they fund, and 75% of foundations with more than $400 million in assets reported doing so. U.S. 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 U.S. 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 the information on outcomes that is reported to the foundation when the grantee organization completes its foundation-supported work. Of the foundations that reported using formal evaluations, 16% said they make the evaluation results public "always" or "often"; 47% said they never do so.[9]

Nature of the evaluations.

Most of the evaluation work done by U.S. foundations falls into one of two categories, and examples representing other approaches can also be found. The broad types of evaluation approaches are: *Grant monitoring, or plan-versus-performance evaluations* that document how foundation funds were actually used and whether the activities originally planned when the project began were actually carried out; various kinds of *outcome evaluations* that assess whether various kinds of results intended by the program operators were achieved; and other *evaluations seeking to provide lessons for policy and practice*.

Plan-versus-performance (or grant monitoring) evaluations are the most common evaluation approach, and are used by a wide range of U.S. foundations, both large and small. They are often but not always carried out by grantee organizations' staff at the request of foundations, and since many grantee organizations lack evaluation expertise (they were awarded grants for very different reasons than having evaluation expertise!), they are varied in their quality and reliability. Examples of U.S. foundations that have developed and used careful monitoring methods include the W. K. Kellogg Foundation, the Lilly Endowment, and the Doris Duke Charitable Foundation. When monitoring information is provided by the organization that has done the work being monitored, the information is sometimes self-serving, because organizations that receive money from foundations do not wish to damage their prospects for future funding. Monitoring reports compare the activities described in the original grant proposal to the activities that were actually completed. They typically list the activities paid for by a foundation grant, and report whether these activities were actually carried out: the meetings that were held, staff salaries that were paid, the services and food that were distributed, the immunizations that were performed, and the buildings that were built. It is beyond the scope of a monitoring report to provide reliable evidence on whether larger outcomes

---

[9] Ibid.

(such as improvements in health or learning, or changes in risky or problematic behavior) were achieved, or on any lessons that were learned on how the foundation-supported activities could be strengthened.

Outcome evaluations include the whole range of impact 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 outcome evaluations commissioned by foundations generally go beyond grantee 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. Many outcome evaluations seek to gather information on whether the intended results were achieved by the foundation-supported activities. This approach is widely known as *program evaluation*. Program evaluations begin by identifying the outcomes sought by program operators (for example, a reduction in smoking by teenagers enrolled in a high school health class); they gather information on whether the outcomes were achieved (such as by mailing a survey to the enrolled students, and sometimes to a comparison group of students as well, asking how often they smoke, or whether they smoke more or less than before the class); and they draw conclusions about the extent to which the desired goals were achieved (such as the extent to which the proportion of surveys returned permits reliable judgments about the behavior reported in the survey). By gathering and analyzing these data, program evaluations systematically assess whether the foundation-supported activities result in the changes in people's behavior that were sought. Examples of U.S. foundations that have frequently used this evaluation approach include the Robert Wood Johnson Foundation, the James Irvine Foundation and the Pew Charitable Trusts. Program evaluations are often intended to enable program operators, foundation staff and trustees to judge whether the foundation-supported activity achieved its goals. As such, they are not designed or intended to go beyond quite narrow conclusions, such as whether the activity should continue to receive funding from the foundation. Some outcome evaluations take a different approach than the program evaluation model; they examine the *implementation* of the program – how well the program was carried out, the proportion of targeted people who actually received services, and the management challenges encountered during the program. These implementation studies analyze the likely effect of the program's implementation on its results. Careful implementation studies 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, as well as other organizations that do not receive support from the foundation. Indeed, 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.

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 learn practical and

policy-relevant lessons from innovative foundation-supported projects. With innovation comes surprises – the discovery of unforeseen barriers and challenges; greater than expected benefits for some groups of participants, and smaller than expected gains for other groups; and the gradual accumulation of practical experience as new approaches receive their first test in the "real world" of management foul-ups, intractable problems, and the effects of human nature in service providers and service recipients. Implementation studies are often a valuable resource for policymakers as they seek to manage well-designed services and articulate quality standards for programs that receive government funding. For example, government policy-makers are often reluctant to undertake new actions if they cannot demonstrate to their overseers that specific desirable outcomes are very likely to result from their use of actions suggested by a foundation project. Solid evaluation evidence is the critically important knowledge base they need in order to move forward. Similarly, the leaders of non-profit organizations may be reluctant to adopt new approaches without reliable evidence on their effectiveness – and with such evidence, they can spread lessons about effective practices to all of the organizations in their 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. Evaluations that provide useful lessons for practice and policy hold great value for society, because their lessons are the building blocks of improved performance by government, non-profit organizations, and private-sector organizations. Among the U.S. foundations that have used this approach are the Robert Wood Johnson Foundation, The Wallace Foundation, the John S. and James L. Knight Foundation, the W. T. Grant Foundation, and the Commonwealth Fund.

Assessment of U.S. foundations' use of evaluation.

While evaluation is used by some U.S. foundations, it is noteworthy that very few of tem use evaluations as a central part of their efforts to achieve their core goals. Foundations that seek to improve the effectiveness of non-profit organizations rarely use implementation evaluations to serve this goal. Foundations that seek to contribute to effective policy in a particular field only occasionally use evaluations to do so. And foundations that support the development of significant innovations rarely use evaluation to determine the effectiveness of the innovation, how it can be implemented more effectively, how it can identify and target the people who will benefit the most from the innovation, and what the innovation would cost if it were to be operated at large scale. Overall, U.S. foundations have not yet made much use of evaluation to strengthen their grantees, the fields they serve, the policymaking process, or organizations other than their grantees. Instead, the U.S. foundations that use evaluation mostly do so to find out whether the foundation's project worked as it was expected to do. This very limited use of evaluation leaves a great deal of room for achieving better results by increasing the use of carefully-designed evaluations.

The status of evaluation in European foundations

New developments in the use of evaluation by European foundations reflect the distinctive context and opportunities they face. In a November 2002 meeting convened by INSP, leaders of seven European foundations discussed a wide range of innovative evaluation and learning activities they have integrated into their work. The evolution of evaluation appears to be already well advanced in a sizable number of European foundations. Interestingly, this change has, as yet, attracted relatively little notice; only a single notable study has examined the use of evaluation by European foundations.[10] Consequently, European foundations' discussions of their experiences with evaluation are likely to be a productive part of ongoing efforts to enhance the effectiveness of their work. Overall, the number of European foundations that use evaluation appears to be modest, but the growing use of evaluation is significant because evaluations appear to be used in ways that directly contribute to achieving the foundations' core purposes and goals, unlike the situation in the U.S. where evaluations are mostly used in ways that are ancillary and complementary to foundations' core purposes and goals.[11]

The European context appears to be a major stimulant for foundations' use of evaluation, for several reasons:

- 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 significant information to the policymaking process.

- Some European foundations engage heavily in supporting the development and use of innovations, and use evaluation to learn about the implementation and effectiveness of these innovations.

- The need to demonstrate accountability, and to respond to pressure from regulatory agencies, has led some European foundations to use evaluation as a tool for demonstrating the implementation and outcomes of their activities.

Many European foundations explicitly structure their work to support independently developed, risk-taking innovations that go beyond existing governmental and civil society activities. In these efforts, evaluations provide reliable, credible learning and evidence that enables key audiences to understand and interpret the results of the innovations. In turn, these lessons contribute to a dialogue between foundations and the other leaders of civil society and government who are working on solving social problems, a dialogue that may also include the media.

---

[10] Compagnia di San Paolo, "European Foundations and Corporate Funders: Evaluation Methods," Turin: Compagnia di San Paolo, 1998; www.compagnia.torino.it.

[11] The information on European foundations reported in this section was gathered by the author, from three sources: discussions at meetings in 2002 and 2003 of the International Network on Strategic Philanthropy, whose members include twenty leading European foundations; presentations by staff from seven European foundations at a special INSP meeting on evaluation in November, 2002; and information gathered through interviews conducted for INSP in 2002 by the Center on Philanthropy and Civil Society at the Graduate Center of The City University of New York, under the direction of Professor Kathleen D. McCarthy. See also Compagnia di San Paolo, op. cit.

This close connection between evaluation and the surrounding European context of the evolving civil society system is quite dissimilar from the context for foundations' use of evaluations in the U.S., with its roots in business management and government program evaluations. European foundations increasingly use evaluation to inform their efforts to add significant innovative improvements to the activities of civil society and the state. Evaluation provides a strong basis for a dialogue with civil society and government that goes well beyond criticism of existing programs, and promotes true innovation.

- Examples include the Bernard Van Leer Foundation's evaluations of early childhood programs, the Bertelsmann Foundation's evaluations of public libraries' programs for young people, and the Joseph Rowntree Foundation's evaluations of innovations in a wide range of social policy fields.

In a departure from the U.S. foundation practice of using evaluation narrowly to monitor foundation-supported projects and to asses the extent to which they achieve their stated goals, some European foundations have developed ways to use evaluations to identify "next steps" for the foundation and the partners it supports, as well as ways to improve the foundation's execution of its strategies.

- Examples include the Zeit Foundation's evaluations of its branded projects (including the Bucerius Law School and Summer School), the Charities Aid Foundation's evaluations of the impacts of its grantmaking, the Compagnia di San Paolo's use of ex-ante evaluation to assess the prospects for future work, the King Baudouin Foundation's use of varied evaluation methodologies to analyze completed projects, and the Freudenberg Foundation's development of evaluation tools that are specifically appropriate for its anti-violence programs. Other European foundations are using the "Theory of Change" approach (also known as the Theory of Action, or logic model, approach) to carefully plan and evaluate the implementation and effectiveness of their activities.[12]

European foundations are also communicating the lessons from their evaluations more broadly than previously through their Internet websites. This is a notable part of the evaluation work of the Joseph Rowntree Foundation, the TSB Lloyds Foundation and the Bernard Van Leer Foundation, among others. By making evaluation evidence and findings public, they provide valuable information to leaders, organizations that provide services, and the public. [13]

The work of these foundations is significant because they are using evaluation as a core activity contributing to the foundations' particular strategic purposes and goals. They avoid the use of

---

[12] Carol H. Weiss, "Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families," in *New Approaches to Evaluating Community Initiatives, Volume 1: Concepts, Methods, and Contexts*, edited by J. Connell, A. Kubisch, L. Schorr and C. Weiss; Washington, D.C.: The Aspen Institute (Aspen Roundtable on Comprehensive Community Initiatives for Children and Families), 1995. See also "W. K. Kellogg Foundation Logic Model Development Guide," www.wkkf.org.
[13] On making evaluation results public, see also Luc Tayart de Borms and Emmanuelle Faure, "Transparency and Accountability," in Andreas Schluter, Volker Then, and Peter Walkenhorst, editors, <u>Foundations in Europe: Society, Management and Law</u>, London: The Directory of Social Change and Bertelsmann Foundation, 2001.

standardized or generic program evaluation approaches that is common in the U.S. This tailoring of evaluations to the core strategies of European foundations holds great promise for the development of valuable and powerful lessons for use by foundations, their partner organizations, other civil society organizations, and government.

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

Some initial indications of the use of evaluation by foundations outside Europe and the U.S. can be gleaned from a report on foundation practices in selected nations that was prepared in 2002 for INSP by the Center on Philanthropy and Civil Society (CPCS) at the Graduate Center of The City University of New York.[14]

In several of these nations, a single theme was found. Donors' need for assurances of financial accountability, and for evidence on whether desired results were achieved, have contributed to the use of evaluation in Russia, Africa, and Hong Kong, reflecting a history of foreign governments' requirements for evaluations as a condition of aid donations.

In Russia, the INSP/CPCS report found that numerous foundations use evaluation, and evaluation practices are well established among foreign donors, community foundations, and private foundations. For many years, foreign donors have required evaluations to assure financial accountability, to monitor the use of funds to determine whether they were used according to the original plan, and to check on whether the desired results were achieved. As a result, many civil society organizations are very familiar with evaluation practices. The International Program Evaluation Network (IPEN21) has conducted information exchanges, a publishing program, a website, conferences, training, and consulting programs on evaluation in Russia. U.S. governmental aid activities in Russia also supported evaluation training during 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. The results of these evaluations are mostly kept private, although there are examples of the publication of selected foundation evaluations. The history of evaluation use in Russia has created a climate in which programs use evaluations to strengthen their work.

Among foundations in Africa, according to the INSP/CPCS report, foreign donors have a long history of requiring the use of evaluations, but in general do not publicize evaluation results, leaving to their partner organizations the decision about whether and how to make evaluation findings 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. This foundation's emphasis on learning is increasing,

---

[14] See footnote 11.

and draws on the many lessons from its partner organizations' experiences; its evaluation topics have included the work of traditional healers and the effectiveness of the foundation's technical support activities. Some foundation leaders in southern Africa report that capacity problems limit their use of evaluation. For others, the reporting of results and the use of evaluation are essential tools of their work. Overall, the recognized need for effective programs appears to have promoted the use of evaluation in many parts of Africa.

There appears to be a moderate amount of evaluation use among the larger foundations in Hong Kong, according to the INSP/CPCS report. Roughly half of the dozen foundations surveyed for the report said they support or commission evaluations. Hong Kong's corporate foundations do not generally use evaluation, because their efforts focus on using philanthropy to build linkages to community groups rather than on achieving particular service outcomes. For them, outcomes other than a positive relationship with the organizations they fund are not of primary importance, so evaluation is not seen as adding value. In contrast, private foundations and NGO fund-raising foundations often use evaluation. This is particularly true for foundations working on development issues, because evaluation systems are an established part of the development field's normal management and program improvement practices. Some foundations that work on development provide evaluation training to their partners. A sizable fraction of these foundations publish the results of their evaluations. The largest Hong Kong foundation, which draws its enormous revenues from the operation of horseracing and wagering businesses, occasionally funds outcome 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. Overall, evaluation appears to be increasing in Hong Kong foundations.

The INSP/CPCS report also assessed the use of evaluation by foundations in Uruguay and Argentina, and produced findings that differ considerably from those in the other nations surveyed. In Uruguay, only a few foundations appear to support or commission evaluations, and the practice of making evaluation results public has received little, if any, support, according to the INSP/CPCS report. There has been considerably more use of evaluation by foundations in Argentina, with as many as half of the larger foundations reporting 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 their respective foundation sectors play a major role both in Uruguayan foundations' limited use of evaluations, and in the greater and more transparent use of evaluations in Argentina.

In nations whose foundations' practices have been shaped by a history of international development aid, the use of evaluation by government donors has created a legacy of expertise in program evaluation. This, in turn, has affected the ways foundations use evaluation: Evaluations are used to

assure financial accountability and to determine whether desired results were achieved. 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.

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. There are many different kinds of evaluations, and foundations need to tailor their evaluations to fit their particular circumstances. Types of evaluations should be thought of as distinctive tools that pertain to different needs, and produce qualitatively different kinds of findings. They offer foundations an array of options from which to choose the evaluation approaches best suited to each particular foundation's agenda.

The experiences of a number of foundations show that there are well-developed evaluation approaches that can produce practical learning tailored to foundations' goals. Six broad types of evaluation approaches can be found in foundations' work. The evaluation approaches described in this section and the next section roughly correspond to the three categories of evaluations discussed above: Grant monitoring evaluations; outcome evaluations (here divided into studies of implementation, impacts, participation, and costs); and other kinds of evaluations that seek to provide logic, knowledge and lessons to inform policy and practice.

1. *Monitoring evaluations*: 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 proposed goals of the foundation-supported activity. Monitoring is typically regarded as an essential part of philanthropic due diligence, that is, taking care that the foundation's resources were used for their intended purposes. A major goal of monitoring evaluations 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 shaping the project. A major use of implementation evaluations is to *strengthen program operations* and practices, 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 program activity is capable of producing a specified result*; when this is already known, it is generally preferable to use implementation evaluations to make

sure the program is being operated in accordance with the findings of previous impact evaluations.

4. *Participation evaluations:* Information about participation is valuable 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 best be recruited to use the program's services.

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, budgeting, and program expansion.*

6. *Evaluations of the logic and the state of knowledge and practice for 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 program*, and they make special contributions to the design of innovative programs.

Evaluation, then, is a suite of different types of learning. To use evaluation, a foundation should begin by determining what it most needs to learn so that the foundation, its grantees, and those it seeks to inform and influence, can achieve their goals. The foundation's answer this question will enable it to decide how to use evaluation well.

Some foundations need to learn about the implementation of an innovation it is supporting; others need to learn whether a particular activity produces the impacts on people it is seeking; still others need to answer operational questions that grow out of the situation in a particular region, a community's history, or a group of people in need; and for others, cost information is what is most needed. The central question for designing a foundation's evaluations is, What do the foundation's leaders most need to learn so that the foundation and its partners can achieve the foundation's goals?

Foundations of many different sizes and types can use the suite of tools offered by evaluation to learn about the results of their work. A foundation need not be large to make use of the types of evaluation discussed here. Large staffs and large budgets are not required to identify what the foundation most needs to learn and to gather relevant information; what is required is a sustained focus on the lessons that are most valuable for the foundation. A small foundation that focuses on developing and refining a single program approach so that it can be used effectively throughout the foundation's region or home community can create and spread practical knowledge about the outcomes of that approach; the foundation can use that single evaluation's findings to strengthen many similar activities in their region or home community. If a small foundation focuses as

consistently on obtaining the lessons that it needs most as it does on working with a grantee to launch a new project, it can learn things that will be very important both for the foundation and its key constituencies.   Deciding what the foundation most needs to learn and tailoring a modest evaluation to gather the needed information is much more productive for a foundation than spending a large sum of money on an unfocused or generic evaluation.   The way is open for foundations with modest assets to make very effective use of evaluation.

PROMISING RESULTS OF EVALUATION, AND NEW EVALUATION TOOLS FOR PHILANTHROPY

As foundations have gained experience in using evaluation, the list of examples of significant benefits from evaluation has grown rapidly. Useful evaluation approaches and tools have also been developed, and are being shared among foundations.

An expanding repertoire of tools now supports foundations' evaluation efforts. The W. K. Kellogg Foundation's evaluation handbook, one of the most frequently downloaded Internet resources in the foundation community, is widely known. Other foundations and organizations that provide services for foundations, including The Wallace Foundation, the Robert Wood Johnson Foundation, the Bertelsmann Foundation, the James Irvine Foundation, the John S. and James L. Knight Foundation, the W. K. Kellogg Foundation, the Pew Charitable Trusts, Grantmakers for Effective Organizations, the Innovation Network, United Way of America's Outcome Measurement Resource Network, the Harvard Family Research Project, and the TCC Group, have created valuable tools and have made them widely available.

Foundation-commissioned evaluations have already produced a solid and practical knowledge base about 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 foundations' evaluations have contributed to society.

1.   Evaluations improve performance and promote accountability through monitoring.

 Grant-making foundations often need information on whether their grantees are carrying out their work as planned. Operating foundations need to manage their work, and this requires accurate progress reports from their staff and partners. Monitoring information on foundation-supported projects is used to make "midcourse corrections" and to decide whether or not to renew funding.

Monitoring is now widely viewed as an essential part of philanthropic due diligence, and it is typically part of the work of program staff.

Monitoring typically focuses on how project funds have been used and whether key project activities have been completed. If the funds have not been used according to the original plan, or if important activities have not been completed on time, monitoring assesses whether the project has become problematic.[15] 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 to make changes in a project while it is still underway, if a problem is identified. In many cases, a project's problems may arise due to shifting conditions or unexpected difficulties that force a change in the original plan. Because part of monitoring is assessing the apparent causes of problems and crafting appropriate interventions, the person doing the monitoring needs to possess a deep understanding of the context and circumstances of the foundation's grantee (or other partner organization).

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 foundation's trustees may focus on accountability; the program officer may seek to support and improve the planned work; and the grantee may seek flexibility in addressing tough challenges. Evaluators sometimes feel pulled between the differing views of these stakeholders. When a project's problems are severe, the possibility that the foundation may lose confidence in the grantee, or may reduce the project's funding, can make cooperation difficult.

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 the addition of several states (chosen for their use of innovative approaches) to the project. These were major changes, with larger roles for senior leaders, greater use of assistance from experts, and clearer plans for achieving desired 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.

"Learning from experience" is central to effective monitoring. This example will sound familiar to many foundation leaders, because many foundations make significant revisions in their projects'

---

[15] From Martha Campbell, "Grants Monitoring," James Irvine Foundation, n.d.

design, activities, capacity-building, and information-gathering, all based on monitoring by the foundation's staff. No one should take lightly the importance and value of monitoring that enables foundations to learn from their experience about how to shape their activities in ways that result in greater benefits for society.

While monitoring information is intended to be used solely by a foundation and the organizations it directly supports, other kinds of evaluation seek to benefit policymakers, leading practitioners, other leaders, and the public, as well as the foundation and its grantees. These evaluation users sometimes have needs and concerns that differ from the concerns of the foundation and its grantees.

2. 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.

In the 1960s, the widespread use in U.S. elementary schools of innovative mathematics curricula was not effective in improving students' understanding of mathematics and their ability to solve math problems. Implementation evaluations of these curricula revealed that many teachers had not received adequate training on how to use the new curricula, resulting in poor quality instruction. The curricula were well-designed, but many teachers simply did not learn how to use them well. The evaluations showed that the *implemented* curriculum was different from the *intended* curriculum. These implementation evaluations also showed how teachers could be trained to use the curricula well, and thereby led to improvements in students' learning. These early implementation evaluations triggered a great expansion of training for teachers.

Implementation evaluations are under-used. They provide invaluable evidence and practical lessons on management problems, organizational incentives, staff quality, and needed improvements in services.

3. 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 achieves desired effects and should continue to be used. Impact evaluations are thought by many people to be "what evaluation is." To the contrary, impact evaluations are simply one approach from the rather lengthy menu of evaluation approaches described in this paper and elsewhere. An impact evaluation is a tool that is appropriate for some situations, and inappropriate for many others. In general, impact evaluations are appropriate only for relatively established, stable programs. In the words of the experienced evaluators Walker and Grossman (1999), "funders should probably focus outcome and impact assessment resources only on stable programs and initiatives that have a track record and have refined their substantive

strategies based on years of operational experience. It is in examining a mature program that one is likely to provide an accurate assessment of the impact and outcome potential of a particular substantive strategy."[16] If an impact evaluation is conducted prematurely, it can lead to the misleading conclusion that the approach studied is a failure, when the truth is that it was simply not well implemented. The challenge for foundation leaders 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, and a missed opportunity to understand the challenges of effective program implementation.

Many foundation staff members fear that an impact evaluation will have negative results, and that such results will make it difficult to continue supporting the project that was evaluated, and related foundation-supported projects. When foundation staff members have a strong emotional commitment to a project, the possibility of negative evaluation results can seem threatening. If a foundation-supported program has been carefully developed and effectively implemented, and field leaders need to determine whether the program is effective, then an impact evaluation may be appropriate – and negative findings are just as valuable as positive findings, because they show whether or not the program is adequate or sufficient to achieve its purpose, and whether a new program design is needed. Negative results are valuable and necessary to avoid wasting the time, and in some cases wasting the life chances, of people who need some other program or service if they are to benefit. It does them no service to keep a program's limitations buried in uncertainty or ignorance. Negative findings from impact evaluations of well-implemented programs are a crucial element in efforts to make progress toward real social benefits.

There are numerous examples showing the benefits of impact evaluations when they are used appropriately. In 1995, the foundation-supported impact evaluation of adult mentoring for "at risk" children 10 to 16 years old found that the mentoring clearly reduces young people's use of drugs and alcohol, reduces their use of violence, and increases their school performance and attendance. The evaluation report, "Making a Difference: An Impact Study of Big Brothers Big Sisters" by Public/Private Ventures,[17] led to increased financial support for mentoring and encouraged more adults to become mentors. Scarce social resources can be used more effectively when they are allocated based on well-designed and well-implemented impact evaluations.

---

[16] Walker and Grossman, op. cit.

[17] Tierney, J., and J. Grossman, "Making a Difference: An Impact Study of Big Brothers Big Sisters," Philadelphia: Public/Private Ventures, 1995 (reissued 2000, www.ppv.org). The foundations that supported this evaluation were the Lilly Endowment, Inc., The Commonwealth Fund, The Pew Charitable Trusts, and an anonymous donor.

4. 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 participated 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).

The evaluation of the Career Academies showed the benefits of participation evaluations. Career Academies are secondary school programs that seek to increase the likelihood that students will graduate from secondary school. They combine courses that prepare the students for university study (that is, Career Academy's curriculum is academic and challenging rather than the undemanding courses that fail to meet universities' entrance requirements) with summer on-the-job learning and strong interpersonal support from a small community of teachers and classmates. The evaluation showed that a very high proportion of students identical to those in the Career Academy (identified through a random lottery held because there were more applicants than could be accepted), but who received only the normal high school program without the features of Career Academies, graduated from secondary school. This showed that the Career Academies attracted and served many students who did not need assistance in completing secondary school. The evaluation showed the importance of changing the eligibility and admission rules for Career Academies, and helped the program's leaders better target the scarce resources available for special school programs on the students who can benefit most from them.[18]

5. 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 enable them to become teachers in high-need schools. The evaluation of their teaching performance showed that they were slightly more effective than conventionally-prepared teachers, and they had greater job retention than other teachers. The evaluation determined that after the development costs of the Pathways program were set aside, the program's operating cost per participant was in the range of $8,000 - $13,000.[19] This is considerably less than the foundation's gross cost per participant (including development costs, such as planning and foundation-sponsored conferences), demonstrating the importance of a careful cost analysis. The high job retention rate for Pathways teachers made these operating costs acceptable, and even attractive, to many school districts that have low retention for teachers. Moreover, having clear and reliable evidence on costs was enormously valuable for policymakers and other potential users of the Pathways approach. When evaluations provide useful information on the costs of a

---

[18] Kemple, J., "Career Academies: Impacts on Students' Initial Transitions to Post-Secondary Education and Employment," Manpower Demonstration Research Corporation, 1999 (www.mdrc.org).

particular program approach, 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.[20]

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

After a decade of support for leading arts organizations' efforts to build greater public participation, with more than four hundred grants supporting the broadening, deepening, and diversification of participation in the arts and culture, The Wallace Foundation sought to understand the logic and effectiveness of the different approaches used by the arts organizations they supported. The resulting evaluation used an innovative approach: it assessed the logic of the projects, and instead of cataloguing the different approaches used to increase participation, it presented (in the words of the report's title) "A New Framework for Building Participation in the Arts."[21] The evaluators carefully examined how people decide whether or not to participate in the arts. The evaluation then identified the strategies used by arts organizations to attract people who are initially disinclined to participate (changing their *perceptions* by welcoming them into the organization and its programs); people inclined to participate but not doing so (by overcoming the *practical* barriers they face, such as parking, scheduled hours of service, or ticket prices); and people already participating (deepening their *experience* of the artistry, and their connection with the organization that hosts the artistic experiences). The evaluation broke new ground by logically connecting arts organizations' participation-building efforts with the needs of individuals as they decide whether to participate. The result has been a new level of discussion among arts organizations as they apply this new knowledge and logic to their participation-building efforts.

As the preceding examples make clear, there are many ways that evaluation – "the use of systematic information-gathering and research activities to learn about the outcomes of organized efforts to bring about social change" – has already provided benefits and practical lessons that have served the strategic purposes and goals of foundations. Looking ahead, it is clear that evaluation can and should be used in ways that reach far beyond monitoring and impact evaluations, valuable as these may be in some situations. The only barriers to using evaluations effectively are the limits on our imagination.

The examples also make plain the fundamental value of evaluation for 21st century philanthropy: Evaluations create useful lessons that provide significant social benefits and advance the strategic

---

[19] Clewell, B., and A. Villegas, "Evaluation of the DeWitt Wallace-Reader's Digest Fund's Pathways to Teaching Careers Program," Urban Institute, 2001. See also B. Clewell and A. Villegas, "Absence Unexcused: Ending Teacher Shortages in High-Need Areas," Urban Institute, 2001; www.urban.org.

[20] Cost studies are not the same thing as cost-benefit studies (and the closely-related Return On Investment – ROI – studies), which require considerable more information and more elaborate methods than cost studies. This section argues that there is considerable benefit in the learning that can result from a relatively simple and straightforward cost study, without the added difficulty that is required to calculate the cost-benefit ratio.

[21] Kevin McCarthy and Kimberly Jinnett, "A New Framework for Building Participation in the Arts," Santa Monica, Ca.: Rand Corporation, 2001; www.rand.org.

purposes and goals of foundations. That is the reason foundations should use evaluation as a core tool enabling them to achieve their missions.

FIVE PITFALLS OF FOUNDATIONS' EVALUATIONS AND HOW THEY CAN BE AVOIDED: LESSONS AND APPROPRIATE CRITICISMS BASED ON FOUNDATIONS' EXPERIENCES WITH EVALUATION

Because foundations' evaluations are essentially tools for systematic learning – rather than tools to ensure compliance, a common role for evaluations conducted by government agencies – the threats they face are threats to learning. The pitfalls that undermine foundations' evaluations are issues that block the gathering of important and practical lessons about the results of foundations' work. This can happen in several ways:

1. Evaluations that do not seek significant learning

The most important lesson from foundations' experiences with evaluation is this: *A foundation's evaluations are really valuable only if they address the most important learning needs of the foundation.* All too often, foundations conduct "general purpose" evaluations that address issues of little interest to the foundation's leaders and staff – evaluations that seek to provide a general impression of whether the foundation-supported activity "did what it was supposed to do," despite the fact that the foundation staff already knew what happened, and thus did not need the evaluation and were unlikely to learn anything from it. In stark contrast, the foundations that benefit most from evaluations use them to find answers to their burning questions – 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 would benefit from it the most? If an innovation were spread widely, what would it cost? What can be learned from the foundation's program that will teach other program operators how they can radically improve their work? Evaluations that address a foundation's most important learning needs can be of enormous value. Evaluations that ignore the questions that are most important to a foundation are of little or no value to the foundation (and probably to anyone else).

2. Low quality evaluations

Even if an evaluation addresses a critically important learning need for a foundation, there is no guarantee that it will create valuable learning. Foundation leaders and staff are well aware that all too many evaluations are untimely, inflexible, and unable to provide useful findings.[22] These failures frequently arise when a generic program evaluation design – one that assesses whether a project's original goals were achieved – is used when it would have been more appropriate to use an implementation evaluation, a participation study, an evaluation of the relevant existing knowledge

---

[22] Patrizi and McMullan, op. cit.

base, or real-time monitoring, or when a good evaluation design is executed badly. The obvious lesson is that a badly designed or badly executed evaluation is as useless as an evaluation that fails to seek important learning or no evaluation at all. Evaluations must be designed and conducted thoughtfully and well, and must be carefully tailored to the needs of a particular project and its supporting foundation, or disappointment is sure to result.

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

Another kind of evaluation-related distress occurs with some frequency in the philanthropic sector: distress due to findings that are negative or are critical of the activity that was evaluated, or findings that are not as positive as desired. Many foundation leaders, staff members, and grantees fear that negative evaluation results will severely damage a project and the reputations of its supporters. Given the great difficulties and uncertainties faced by many foundation-supported projects, and the optimism of their supporters, this anxiety should not be surprising. Most progress is slow and most advances come one step at a time, but the deadlines for decisions on additional funding arrive quickly, and such decisions are more difficult if the early evaluation findings are discouraging. Foundation staff members and program operators are understandably nervous about the effects of an evaluation on decisions about additional funding, whether or not an evaluation produces valuable practical lessons.

Interestingly, it was disappointing evaluation findings on several U.S. foundations' very large, projects aimed at improving community support for children that led evaluation experts to create the "Theory of Change" evaluation approach.[23] In this approach, the developers of a new project carefully specify the steps in a complex project – beginning with initial planning and start-up, to recruiting participants, to the implementation of services, to interim results, and finally to the desired outcomes – and the evaluators then collect data on the progress made toward completing each step in the plan, in order to document the project's learning and initial progress as well as to assess the barriers to success that were encountered and how they might be overcome in the future. "Theory of Change" evaluations seek to produce practical lessons from the early stages of a project, and have been embraced by some foundations for this reason. The lessons that led to the creation of this planning approach came from "disappointing" evaluations.

For many foundation leaders and staff members, evaluators are often viewed as policemen, enforcing compliance with predetermined plans and punishing those who produce disappointing results. Few, if any, foundations use evaluation in this way, but the threatening image of "the evaluator as policeman" can easily replace the image of "the evaluator as gatherer of lessons and as learner." When evaluation is seen as policing, staff resistance and conflict with program operators

---

[23] Connell, Kubisch, Schorr and Weiss, op. cit.

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. Contrary to this worry, sometimes long-term funding is the result of an evaluation – for example, when the evaluation requires gathering information on results by tracking the results of a very lengthy program or treatment. More commonly, evaluations that include an implementation study provide concrete, practical suggestions for ways to improve the program's performance and overcome performance problems. Implementation evaluations do not conclude with a "thumbs up or thumbs down" verdict that labels a project as a success or a failure. Exaggerated language is rare in evaluation reports, which are more likely to describe complex outcomes than either grandiose accomplishments or complete program collapses.

4. Excessive focus on specifying outcomes

"Grantees report that never before have grant negotiations with foundation staffs been so focused on specifying outcomes," according to evaluators Walker and Grossman.[24]   Sometimes this requirement is tied to evaluation, and often it is not. Some foundations require that a project's desired outcomes be specified in advance, and be documented by the program operators, 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. Is it has any value, it would seem to be mostly when the foundation is paying for a well-defined product, such as the construction of a new building, the hiring of specialized staff for a grantee, or the execution of a well-understood task (such as an audit or drawing a new map of a specified area). Documenting outcomes is no substitute for learning nuanced lessons about the results of a foundation's activities.

5. Premature evaluation of innovations?

If an innovation is evaluated before it has been fully implemented – like a flower that is observed before it blossoms – important parts of its performance will be neither observed nor appreciated. Innovations, whether supported by a foundation or not, almost never unfold strictly according to a plan. Some foundation leaders fear that if evaluations are used to assess innovations, the result will be a suppression of risk-taking and a reluctance to support innovations – because they assume that the inevitable uncertainties and the uneven progress that are typical of innovations will produce disappointing evaluation results. In this vein, Anheier and Leat have argued against the premature

---

[24] Walker and Grossman, op. cit.

use of evaluation because it may have the unintended effect of reducing foundations' support for innovation.[25]

Whether or not the premature use of evaluation has harmed worthy innovations (and documented reports of this are difficult or impossible to find), 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 at learning how an innovation unfolds and what results, if any, it may be capable of achieving. Indeed, the lessons provided by high-quality implementation evaluations are likely to provide irreplaceable lessons that can accelerate the development and spread of effective innovations, by systematically building up a rich store of evidence and understanding about innovative practices and outcomes. Innovations spread when their results have been carefully described and understood.[26]

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 need to learn and using their evaluations to obtain those lessons.

STANDARDS FOR EVALUATION IN 21ST CENTURY FOUNDATIONS

In an increasingly skeptical world, it is not very likely that foundations' contributions will be widely seen as benefiting society without reliable and highly credible evidence of their effects. Foundations can provide much-needed evidence of the value of their work, along with many practical lessons of considerable value for society, by using evaluations well. The benefits of evaluation are the benefits that come from harvesting the most important learning that has been seeded by foundations' work. The experiences 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 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 the foundation's work. Thus, evaluation and the information

---

[25] Anheier and Leat, op. cit.
[26] Rogers, E., *Diffusion of Innovations*, 5th edition, New York: Free Press, 2003.

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 directors to use evaluation to gather information on the results and the lessons of foundations' activities.

BIBLIOGRAPHY

Anheier, H., and D. Leat, *From Charity to Creativity: Philanthropic Foundations in the 21st Century*, Stroud, England: Comedia, 2002.

Campbell, Martha, *Grants Monitoring*, James Irvine Foundation, n.d.

Clewell, B., and A. Villegas, *Evaluation of the DeWitt Wallace-Reader's Digest Fund's Pathways to Teaching Careers Program*, Urban Institute, 2001.

Clewell, B. and A. Villegas, *Absence Unexcused: Ending Teacher Shortages in High-Need Areas*, Urban Institute, 2001; www.urban.org.

Compagnia di San Paolo, *European Foundations and Corporate Funders: Evaluation Methods,* Turin: Compagnia di San Paolo, 1998; www.compagnia.torino.it.

Council on Foundations, *Evaluation for Foundations: Concepts, Cases, Guidelines, and Resources*, San Francisco: Jossey-Bass Publishers, 1993.

Kemple, J., *Career Academies: Impacts on Students' Initial Transitions to Post-Secondary Education and Employment*, Manpower Demonstration Research Corporation, 1999 (www.mdrc.org).

Lovallo, D., and D. Kahneman, *Delusions of Success*, Harvard Business Review, November 2003.

McCarthy, Kevin and Kimberly Jinnett, *A New Framework for Building Participation in the Arts,* Santa Monica, Ca.: Rand Corporation, 2001; www.rand.org.

Patrizi, P., and B. McMullan, *Evaluation in Foundations: The Unrealized Potential*, W. K. Kellogg Foundation, 1998 (www.wkkf.org).

Porter, R., and M. Kramer, *Philanthropy's New Agenda: Creating Value,* Harvard Business Review, 1999.

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

Tayart de Borms, Luc and Emmanuelle Faure, *Transparency and Accountability*, in Andreas Schluter, Volker Then, and Peter Walkenhorst, editors, *Foundations in Europe: Society, Management and Law,* London: The Directory of Social Change and Bertelsmann Foundation, 2001.

Tierney, J., and J. Grossman, *Making a Difference: An Impact Study of Big Brothers Big Sisters*, Philadelphia: Public/Private Ventures, 1995 (reissued 2000, www.ppv.org).

Walker, G., and J. Grossman, *Philanthropy and Outcomes: Dilemmas in the Quest for Accountability*, Public/Private Ventures, 1999; www.ppv.org.
Ostrower, F., *Attitudes and Practices Concerning Effective Philanthropy,* Washington, D.C.: Urban Institute, 2004; www.urban.org.

Weiss, Carol H., *Nothing as Practical as Good Theory: Exploring Theory-Based Evaluation for Comprehensive Community Initiatives for Children and Families*, in: New Approaches to Evaluating Community Initiatives, Volume 1: Concepts, Methods, and Contexts, edited by J. Connell, A. Kubisch, L. Schorr and C. Weiss; Washington, D.C.: The Aspen Institute (Aspen Roundtable on Comprehensive Community Initiatives for Children and Families), 1995. See also "W. K. Kellogg Foundation Logic Model Development Guide," www.wkkf.org.

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 Mittenthal

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