The Foundation Review

Volume 5 | Issue 2 Article 5

1-1-2013

Benchmarking Evaluation in Foundations: Do We Know What We Are Doing?

Julia Coffman
Center for Evaluation Innovation

Tanya Beer
Center for Evaluation Innovation

Patricia Patrizi
Patrizi Associates

Elizabeth Heid Thompson
Patrizi Associates

Recommended Citation

Coffman, Julia; Beer, Tanya; Patrizi, Patricia; and Heid Thompson, Elizabeth (2013) "Benchmarking Evaluation in Foundations: Do We Know What We Are Doing?," *The Foundation Review*: Vol. 5: Iss. 2, Article 5.
DOI: 10.4087/FOUNDATIONREVIEW-D-13-00009.1
Available at: http://scholarworks.gvsu.edu/tfr/vol5/iss2/5

doi: 10.4087/FOUNDATIONREVIEW-D-13-00009.1

Benchmarking Evaluation in Foundations: Do We Know What We Are Doing?

Julia Coffman, M.S., and Tanya Beer, M.A., Center for Evaluation Innovation; and Patricia Patrizi, M.S., and Elizabeth Heid Thompson, B.A., Patrizi Associates

Keywords: Evaluation, philanthropy, foundations, evaluation in philanthropy, evaluation in foundations, benchmarking, benchmarking research, Evaluation Roundtable, foundation evaluation directors, learning, strategic learning, performance management, organizational learning, knowledge management, Center for Evaluation Innovation

Key Points

- Evaluation in philanthropy – with staff assigned to evaluation-related responsibilities – began in the 1970s and has evolved, along with philanthropy, in the four decades since. What has not changed, however, is a regular questioning of what foundations are doing on evaluation, especially since the world of philanthropy regularly shifts, and changes in evaluation resourcing and positioning tend to soon follow.
- This article presents new findings about what foundations are doing on evaluation and discusses their implications. It is based on 2012 research that benchmarks the positioning, resourcing, and function of evaluation in foundations, and follows up on a 2009 study that used a similar design.
- The participating foundations were surveyed and interviewed. They were asked about the range of activities they used to produce evaluative information about grantmaking, perceptions about the adequacy staff time and money for evaluation, and how and how well they use evaluative information throughout the life cycle of grantmaking programs and strategies.
- Benchmarking research was conducted for the Evaluation Roundtable, a network of foundations seeking to improve how they learn about the results of their grantmaking and enhance the difference they make.

Introduction

Much like the discipline of evaluation itself, the evaluation function in philanthropy –with staff assigned to evaluation-related responsibilities – is a fairly recent phenomenon in the United States. Its roots trace back only to the 1970s, when pioneers like the Robert Wood Johnson, Ford, and Russell Sage foundations began making serious commitments to evaluation (Hall, 2004).

Forty years ago, evaluation in philanthropy looked much different than it does today. At the time, less than two percent of foundations had professional staff, making evaluation mostly a larger foundation concern. For those few foundations making concerted evaluation investments, the focus was on assessing individual grants, often as grants were closing.

As philanthropy has evolved in the decades since, so has evaluation in philanthropy. When the 1990s brought a huge increase in the number of foundations, interest in evaluation surged (Patrizi & McMullan, 1999). Factors credited include more donors and trustees coming to philanthropy with a results orientation, the professionalization of nonprofit management and incorporation of a business orientation, an increase in the diversity of methods and tools available for evaluating different types of grants, and high-profile and generous foundation champions for evaluation (Hall, 2004). Similarly, evaluation received a boost in the last decade with the rise of strategic philanthropy, in which foundations seek to achieve their own clearly defined goals, pursue those goals in collaboration with grantees, and then track their success in achieving them (Brest, 2012; Patrizi & Thompson, 2011). Evaluation in foundations has again expanded as new methods and tools have been introduced for evaluating increasingly long-term and adaptive foundation strategies where traditional program evaluation approaches are not a good fit (Britt & Coffman, 2012; Preskill & Beer, 2012).

The one thing that has not changed over the last four decades, however, is a regular questioning of what foundations are doing on evaluation, especially since the world of philanthropy regularly shifts, and changes in evaluation resourcing and positioning tend to follow. In addition to questions about what foundations are doing, questions arise regularly about whether foundation evaluation investments – where they exist – are as useful as they can be, and if not, how to improve them. This has been true since the president of the Russell Sage Foundation wrote a 1973 essay on this topic, "Do We Know What We Are Doing?" (Heimann, 1973). As the title of this article suggests, that overarching question still dominates discussions about this topic.

This article presents new findings about what foundations are doing on evaluation and discusses their implications. It is based on 2012 research that benchmarks the positioning, resourcing, and function of evaluation in foundations. This study was conducted for the Evaluation Roundtable, a network of foundations seeking to improve how they learn about the results of their grantmaking and enhance the difference they make. The Center for Evaluation Innovation conducted the research and also leads the Evaluation Roundtable.

This is not a study about evaluation in all of philanthropy. As was the case decades ago, the hiring of professional staff dedicated primarily to evaluation-related activities is still something that mostly larger foundations do. As such, this is a study of primarily larger foundations and other foundations known for their commitment to

The one thing that has not changed over the last four decades ... is a regular questioning of what foundations are doing on evaluation, especially since the world of philanthropy regularly shifts, and changes in evaluation resourcing and positioning tend to follow.

evaluation. Most of the foundations that participated are among the 100 largest U.S. grantmaking foundations as ranked by asset size (Foundation Center, 2013).[1]

Methodology

Research questions for the 2012 benchmarking study focused on evaluation staffing, investments, practices, and use. Specifically, they asked foundation participants about the range of activities used to produce evaluative information about grantmaking, perceptions about the adequacy of resources (staff time and money) for evaluation, and how well foundations use evaluative information throughout the life cycle of grantmaking programs and strategies.

Data-collection methods included both a web-based survey and telephone interviews with foundation representatives. In most cases, the person who had primary responsibility for or led evaluation activities completed the survey and interview. This research is a follow up to similar benchmarking research conducted by Patrizi As-

[1] There are more than 50,000 foundations in the U.S. The vast majority have much smaller assets than the foundations in this study and function with no paid staff. While a study of how they and other funders like community foundations experience and invest in evaluation would be a fascinating endeavor, their inclusion fell ouside of the scope of this research.

TABLE 1 2012 Benchmarking Study Participants

	Annual Grantmaking Budget		
Evaluation Reports to	Under $50 Million (13)	Between $50 Million and $200 Million (12)	More Than $200 Million (6)
CEO (15)	• Arcus Foundation • Barr Foundation • The Colorado Trust • New York State Health Foundation • Otto Bremer Foundation (Canada) • Sierra Health Foundation	• The California Endowment • John S. and James L. Knight Foundation • Margaret A. Cargill Philanthropies • MasterCard Foundation (Canada) • Ontario Trillium Foundation (Canada) • Wallace Foundation	• John D. and Catherine T. MacArthur Foundation • Robert Wood Johnson Foundation • William and Flora Hewlett Foundation
Administrator, such as CFO, COO (8)	• Skoll Foundation • Lumina Foundation • Skillman Foundation	• The Colorado Health Foundation • Foundation Lucie et Andre Chagnon (Canada) • Rockefeller Foundation	• The Pew Charitable Trusts • Bill and Melinda Gates Foundation
Program VP or director (8)	• California HealthCare Foundation • Edna McConnell Clark Foundation • J.W. McConnell Family Foundation (Canada) • Rothchild Caesarea Foundation (Israel)	• Annie E. Casey Foundation • The Duke Endowment • James Irvine Foundation	• David and Lucile Packard Foundation

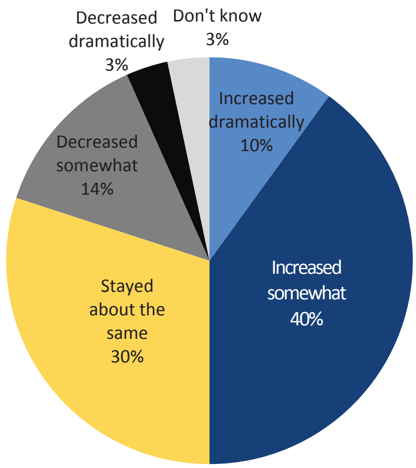
sociates in 2009 (Patrizi Associates, 2010). Many of the same survey questions were used in 2009 and 2012 to allow for an examination of patterns over time.

Thirty-one foundations participated: 26 based in the U.S., four based in Canada, and one based in Israel. (See Table 1.) In terms of foundation size as measured by annual grantmaking budgets, respondents were 13 smaller foundations (under $50 million), 12 mid-size foundations (between $50 million and $200 million), and six large foundations (more than $200 million). Participants also differed on their evaluation reporting structures (evaluation unit reports to the chief executive officer, an administrative leader such as a chief operating officer, or a program leader).

As this article discusses, evaluation in foundations now encompasses a broad range of activities. Because the shape of the evaluation function in philanthropy has expanded in recent years, both the 2009 and 2012 benchmarking studies focused more broadly on the use of and demand for "evaluative information" rather than solely on "evaluation." For shorthand, the term evaluation is used here to represent the suite of foundations' evaluation-related activities.

As in the 2009 benchmarking study, the survey analysis segmented responses by annual grantmaking budget and evaluation reporting. (See Table 1.) Only interesting patterns by size and reporting structure are reported here using this segmentation. Interview data were coded and

FIGURE 1 Funding Levels for Evaluation Relative to Size of Overall Budget



analyzed, and select interview findings were included to add depth to survey results. Because of the limited sample size and different foundation participants across years, comparisons from the 2009 to the 2012 benchmarking data are limited. Findings here provide data from the 2009 research as reference points and identify potential patterns where appropriate.

Findings

Eight main findings emerged from the benchmarking research. A more detailed account of benchmarking data and findings also is available (Center for Evaluation Innovation, 2013).

Foundation Commitment to Evaluation Is Increasing

Using investment levels and staffing for evaluation as indicators, 2012 benchmarking data suggest that foundation commitment to evaluation is rising. Fifty percent of the 31 foundations surveyed said that evaluation investments relative to grantmaking had increased during the last two years. (See Figure 1.) Thirty percent said their investments had stayed about the same.[2] A similar but even stronger result was found in 2009, when 62 percent perceived a recent increase in evaluation investments relative to grantmaking, even amid an economic crisis that was affecting the size of foundation endowments.

Similarly, the average number of staff dedicated to evaluation has increased since 2009, except in smaller foundations, where it decreased slightly. The average number of full-time employees across all foundations increased from 3.0 in 2009 to 4.2 in 2012. Three-fourths of foundation respondents had at least one full-time employee dedicated to evaluation-related activities, and the larger foundations had an average of 10. Of those with less than one full-time employee at the director level, all but one were smaller foundations. Whether current staffing levels are adequate to address evaluation demands within foundations is another issue, as discussed below.

[2] Actual expenditures related to evaluation are not tracked in a consistent manner, if at all, by most foundations in our research. Interviews revealed that tracking has become increasingly difficult as the range of evaluative activities has expanded and as responsibility for evaluation has been distributed among staff outside of the evaluation unit.

Evaluation-unit responsibilities now go well beyond managing evaluation contracts or assessing individual grantee results. Many are now leading a range of evaluation practices that include performance management, knowledge management, organizational learning, and strategic learning.

The Role and Scope of Evaluation in Foundations Is Expanding

Since the beginning of foundation evaluation 40 years ago, there has been a quest to generate "meaningful or reliable findings that could provide any real guidance to foundation decision makers" (Hall, 2004, p. 34). That quest remains a core concern for foundation evaluation, even as the types of decisions foundation staff make have expanded well beyond the basic, "Should we fund it?" to include questions about how staff can ensure grantmaking is as effective as it can be.

As foundations grapple with generating meaningful data for difficult decisions associated with increasingly complex strategies, they are experimenting with the role and scope of their evaluation functions. For the most part, this has meant expansion in the types of evaluation work in foundations. While the number of full-time employees dedicated to evaluation has increased on average for the foundations in the 2012 study, so have evaluation responsibilities. In fact, the expansion of responsibilities may be occurring at a faster pace than the growth in evaluation staff.

Reflecting a trend also found in 2009, of the 26 foundations with a unit or department dedicated to evaluation-related activities, more than one-third (37 percent) had either created a new unit or changed its name in the past two years. Changes were motivated by either an expansion of responsibilities or a shift in emphasis on types of evaluative responsibility. Evaluation-unit responsibilities now go well beyond managing evaluation contracts or assessing individual grantee results. Many are now leading a range of evaluation practices that include performance management, knowledge management, organizational learning, and strategic learning.

On one hand, this expansion reflects a seemingly positive trend. It suggests that foundations are recognizing that different kinds of decisions require different types of data and that there is much to be learned from grantmaking across programs and across time. Indeed, more foundations are attempting to adopt measurement, evaluation, and learning systems that attempt to match data and information to specific needs. At the same time, interviews with evaluation staff paint a more complicated picture of the day-to-day realities that these shifts have created.

With an expansion and diversification of responsibilities, the evaluation function can be complex and challenging to manage. While these activities are evaluation-related, they are not synonymous. (See Table 2.) This can create internal confusion about what these activities can and cannot deliver. It also can be challenging to do all of these things well, especially for foundations that have only one full-time position or less devoted to this function.

Compounding this is the reality that many of these activities require skills that are not often part of traditional training for program evaluators or social science researchers. Benchmarking interviews revealed that foundations are being challenged to find individuals qualified for increasingly multifaceted evaluation positions.

Some evaluation units in larger foundations are responding to an expansion in responsibilities by restructuring and assigning staff to distinct roles, such as a learning officer, a performance management officer, and an evaluation officer. This, however, can have the opposite effect, creating duplication and role confusion. Many

TABLE 2 Types of Evaluation Responsibilities

Evaluation	The systematic collection and analysis of data to draw conclusions about the conceptualization, design, implementation, effectiveness, and/or utility of grants or programs.
Performance management	The ongoing collection of information for monitoring how a grantee, program, strategy, or the foundation as a whole is doing, usually in relation to a set of predetermined goals.
Knowledge management	Practices to identify and disseminate lessons and experiences.
Organizational learning	Practices to ensure an organization applies lessons and adapts.
Strategic learning	Gathering of data and insights from a variety of information-gathering approaches – including evaluation – to inform decision-making specifically about strategy.

others hire consultants to fill gaps, at the risk of having parallel efforts that are not integrated or complementary or that even may conflict, such as performance-measurement systems that create incentives for staff to adhere to a particular course of action even while separate evaluation activities conclude that a change in direction is needed.

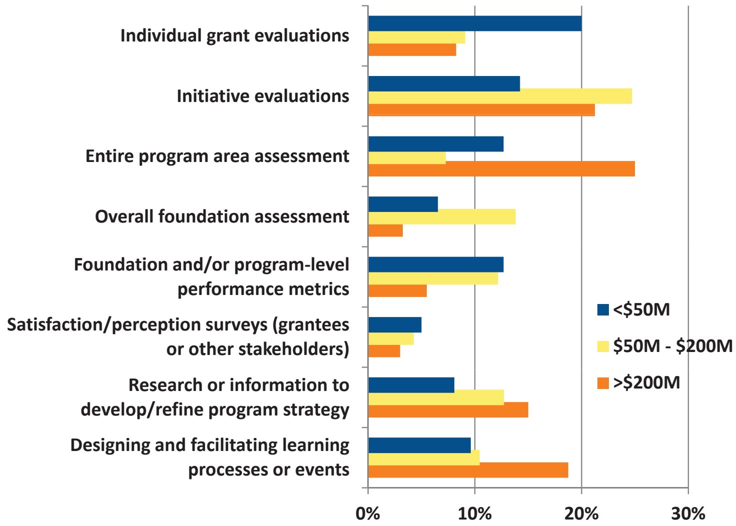
Finally, evaluation staff indicated that the addition of new responsibilities can be motivated by factors that do not include a clear need for that function. Changes, for example, might be a reaction to a perception that evaluation has a poor track record of informing foundation decisions. Rather than identifying how to improve what already exists, there is a tendency to add new activities, thinking that the "next big thing" in evaluation might be the answer. This constant churn in evaluation scope and responsibilities can leave evaluation staff feeling "punch drunk," and the activities they manage may end up performing below expectations in terms of their ability to deliver meaningful information. In some cases, interviews with evaluation staff suggested, going deeper on fewer evaluation activities and identifying where they can be applied most appropriately might deliver more usable evaluation information than multiplying the types of evaluative activities foundations perform.

Evaluation Staff Are Juggling More Than in the Past

Foundations were asked to estimate the percentage of total staff time over the course of a year that was dedicated to evaluation activities. Eight activities were assessed based on previous benchmarking research that indicated how evaluation staff spend their time: individual grant evaluations, initiative evaluations, entire program-area assessment, overall foundation assessment, program or foundation-level performance metrics, grantee/stakeholder satisfaction or perception surveys, research to inform strategy, and learning facilitation.
In sync with the finding about the expanding scope of evaluation units, the findings show that evaluation staff in most foundations spend their time on the majority of these activities. On average, they divide their time among 6.5 activities, many of which require varied skills and expertise.

The distribution of how much time staff spent on these activities differed, however, for smaller foundations compared to mid-size and large. The top two activities for smaller foundations were individual grant evaluations and initiative evaluations. (See Figure 2.) For mid-size foundations, they were initiative evaluations and overall foundation assessment; for large foundations

FIGURE 2 Average Percentage of Evaluation Staff Time Dedicated to Activities, by Foundation Size



they were entire program-area assessment and initiative evaluations. Evaluation staff at smaller foundations spend more time on smaller-scope activities (individual grants), while larger foundation staff devote more time to broader-scope activities (program areas, overall foundation).

Foundations of all sizes were similar, however, in that evaluation staff tend to focus much more on the production of data than on the facilitation of learning from it. Only large foundations said more than 10 percent of evaluation staff time was devoted to learning activities. This finding suggests that the growth in the types of evaluation-unit activities – as well as the multiple units of analysis on which staff are working – is contributing to a general proliferation of evaluative data and information in foundations. Benchmarking interviews indicated this proliferation has had some unintended consequences.

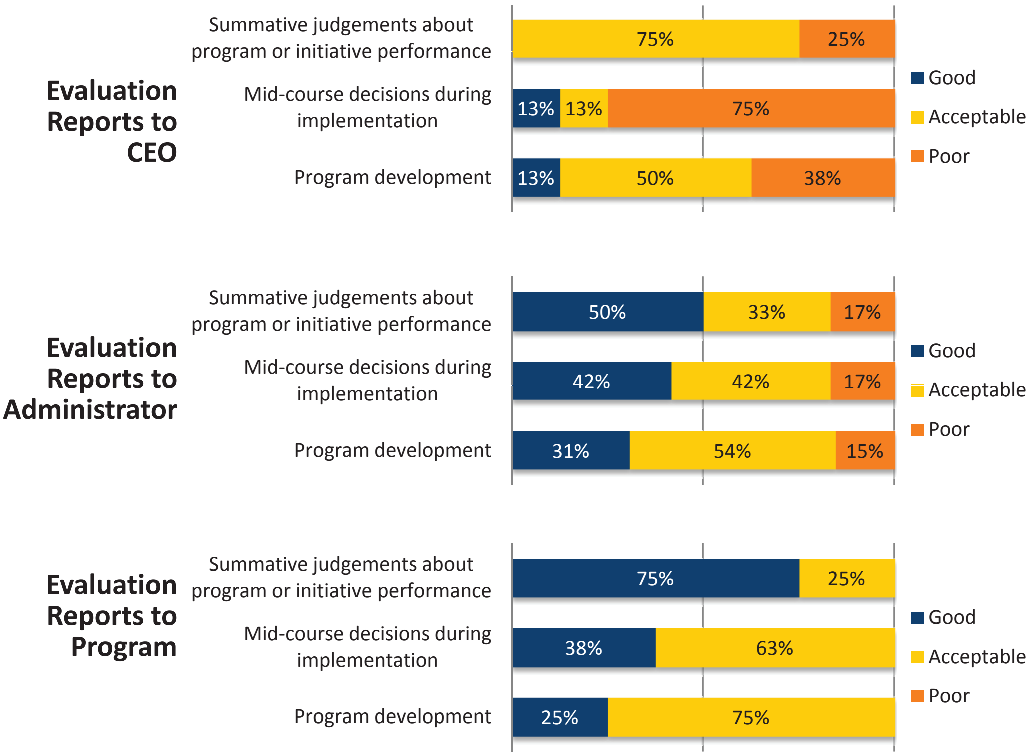
While producing usable information is the goal of most evaluation staff and adding more evaluation activities would seemingly help to achieve that goal, the supply of information might be growing faster than the need for it. Instead of leading to smarter decisions, more and more data can overwhelm people and leave them feeling like they are "drinking from a fire hose" of information (Frank & Magnone, 2011). As evaluation staff said during benchmarking interviews, better balancing their activities between producing high-quality information and working with foundation staff to use it (e.g., by helping them to ask better questions at the right time) might be a more productive use of their time.

Structure Matters

Most foundations that participated in the 2012 research have evaluation directors. Three-fourths had a full-time staff person at the director or manager level, charged with managing the expanding range of evaluation-related activities. That person also may lead evaluation capacity-building efforts among foundation staff, coordinate with communications and strategy leaders to develop integrated and aligned foundation practices and terminology, and help create a learning- or performance-oriented culture.

There has been much debate about where the evaluation director should sit in the organization to ensure that evaluation has the greatest possible impact. Three main options exist – reporting to the president or CEO, to an administrator like a COO, or to program leaders. The argument for

FIGURE 3 Assessment of Program Staff's Use of Evaluation to Inform Grantmaking Decisions



reporting to the CEO is that leadership buy-in and support is critical to securing evaluation resources. In addition, it enables the evaluation director to speak directly to the person who most profoundly shapes foundation culture and can help ensure staff take-up of evaluation. This executive-level positioning also gives the evaluation director access to top-level foundation discussions where evaluation can be championed. The argument for reporting to a high-level administrator is that those individuals are closest to the management of day-to-day practice in foundations. While they tend to be closer to program staff than the CEO and therefore influential in integrating evaluation into the grantmaking process, they also maintain enough distance from programs to ensure that an evaluation director is less likely to be pressured to suppress data that could negatively reflect on program staff performance. Finally, the argument for reporting to program leaders is that close alignment with program staff will increase the chances that evaluation is integrated and used because the evaluation director is closer to actual grantmaking and asking the questions in which program staff are most interested.

As was the case in 2009, across small, mid-size, and large foundations it is most common for the evaluation director to report directly to the CEO. (See Table 1.) Almost twice as many evaluation directors reported at that level than to an administrator or program leader. But 2012 data also indicate that there might be tradeoffs to this positioning and that there are no clear-cut answers to where the evaluation director should sit to maximize evaluation's utility.

CEO reports said that they perceived higher levels of management support for evaluation and were more satisfied with the level of investment in evaluation. However, evaluation directors who reported to program leaders were more satisfied with evaluation's use by program staff. (See

Because strategies are "owned" by program staff, ... evaluation and program staff must work together at all stages of the strategy life cycle. This is a considerable change from 20 years ago, when evaluation staff, like the process itself, sat much farther from program decision making.

Figure 3.) This finding is different from the 2009 benchmarking study, which found a clearer case for positioning at the CEO level (Thompson & Patrizi, 2010). In 2009, CEO reports were more satisfied with all three aspects: management support, investment levels, and program use. That there might be no best or consistent answer to this question and instead a set of tradeoffs appears to be further supported by the 2012 data. Foundations are experimenting with different configurations on this issue. One-third of the 18 foundations that responded to both the 2009 and 2012 surveys had switched their evaluation director's positioning in the last two years. Three switched from a CEO report to a program report, and three switched from program report to CEO report.

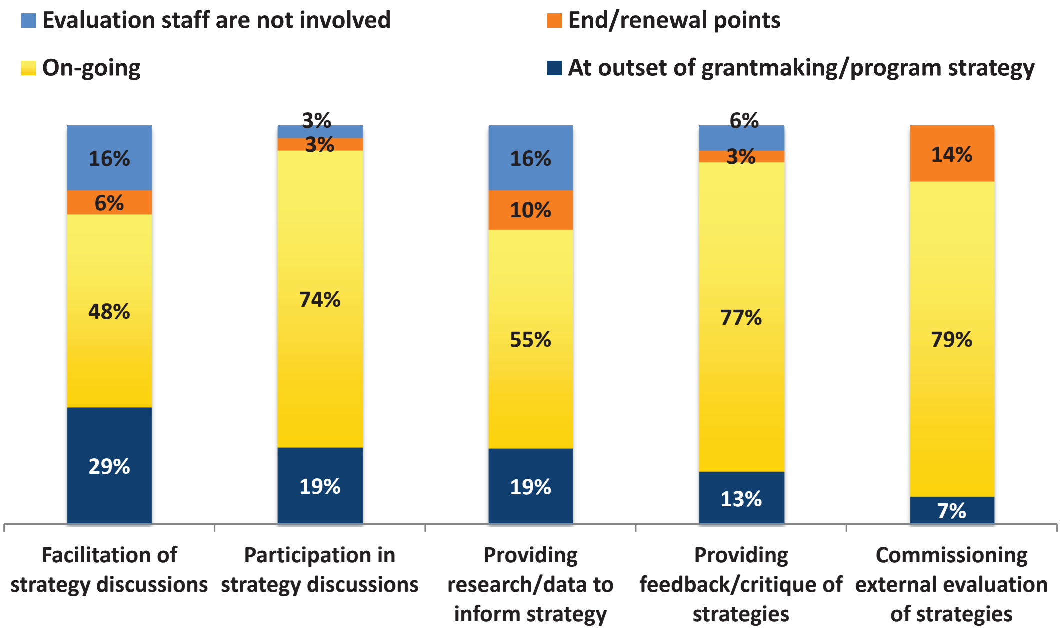
Finally, while this research did not collect data on this issue, there is some indication that foundations are looking for different backgrounds, skills, and expertise in their evaluation directors than they used to. Twenty years ago, evaluation directors commonly came from academia and had a strong background in applied social science research. They were methodology experts and practicing evaluators. A review of position descriptions for evaluation directors circulated recently suggests qualifications might be shifting as expectations for those positions diversify and expand. Because the director must work closely with program staff and help guide the learning process, backgrounds as practicing evaluators in a foundation's focus program areas are being weighed alongside substantive knowledge about philanthropy and strategy as well as soft skills such as communication and facilitation.

Evaluation Is Being Integrated Throughout the Strategy Life Cycle

In a *New Directions for Evaluation* issue devoted to this topic, Patton and Patrizi (2010) concluded, "Strategy is a new unit of analysis for evaluation" (p. 5). Particularly with the shift toward strategic philanthropy, foundations have been challenged to reframe evaluation from an older model of "post hoc" assessment of grantee projects and programs for accountability purposes to one that is more focused on examining foundation strategy and informing learning about it from start to finish. (Most often those strategies are focused on foundation initiatives, or multigrantee long-term efforts designed to achieve specific goals.) In response, many foundations have shifted from the use of evaluation for proof ("Did it work?") to a focus on also using evaluation for improvement ("What did we learn that can help us make it better?"). This means that evaluation staff and evaluation activities are aiming to be more integrated throughout the strategy development and implementation process. Because strategies are "owned" by program staff, it also means that evaluation and program staff must work together at all stages of the strategy life cycle. This is a considerable change from 20 years ago, when evaluation staff, like the process itself, sat much farther from program decision making.

Both 2009 and 2012 benchmarking data show that this evaluation-to-program integration is occurring. Program staff are heavily involved in evaluation activities and, for many, evaluation and program staff are sharing responsibility for them. This is especially true for activities that include program or foundation performance assessment, initiative evaluations, and program-area assessments.

FIGURE 4 Evaluation Staff Involvement in Activities to Inform Grantmaking/Program Strategy



Benchmarking data also show that evaluation staff are helping to inform program strategy, and are focusing more on informing it throughout its lifecycle. For the majority of foundation respondents, evaluation staff are participating in, critiquing, and facilitating strategy discussions as well as providing data to inform strategies and commissioning external evaluations of them. (See Figure 4.) In addition, for the most part, evaluation staff are doing these things in an ongoing way as strategies are developed and executed. This is a notable change from 2009, when respondents said that evaluation staff participation in strategy dropped off considerably during the ongoing stages of strategy evolution. For example, in 2009, only 27 percent reported that evaluation staff were heavily involved in "providing feedback or critique" in an ongoing way. That percentage jumped to 77 percent in 2012. These data suggest some improvement from what Patrizi and Thompson (2011) concluded based on their 2009 benchmarking research: "We've found that many foundations make the mistake of approaching strategy development as an upfront, analytic exercise that ends when implementation begins" (p. 53).

The purpose of integrating evaluation staff into the strategy life cycle is to help ensure that evaluation data and information is relevant, timely, and useful to strategy decisions. Benchmarking data indicate that this is occurring. Most foundation respondents in 2012 said that the use of evaluation to inform all stages of the strategy life cycle was at least acceptable, although about one-fourth of respondents said its use was still poor in informing midcourse decisions and in making summative judgments about program or initiative performance. As the next finding suggests, there is still considerable room for improvement in ensuring data actually are used and that learning takes place as grantmaking strategies are implemented.

Evaluation Use and Learning Continues to Be a Challenge

In spite of a positive assessment of program staff's use of evaluation throughout the grantmaking life cycle, benchmarking interviews revealed deeper frustration among evaluation staff about the extent to which foundations use evaluation data to understand how complex strategies unfold "on the ground" and then make appropriate adjustments. The list of barriers to the effective use of evaluation was familiar. More than two thirds of foundations cited staff time and workload as the top barrier; almost half cited the timeliness of data, and almost one-third mentioned organizational culture and negative attitudes about evaluation.

Many foundations are testing process improvements to address some of these barriers, such as regularly carving out reflection time, commissioning evaluation approaches that return data more rapidly, instituting planning processes that force greater clarity about strategy, and building evaluation capacity among staff. Yet many interviewees still expressed doubt about whether these process adjustments – or the increasing volume of data available – are leading to changes in strategy that observably improve grantmaking results. Few could provide concrete examples of meaningful midcourse strategy changes that resulted from program staff learning from data, even when data were available.

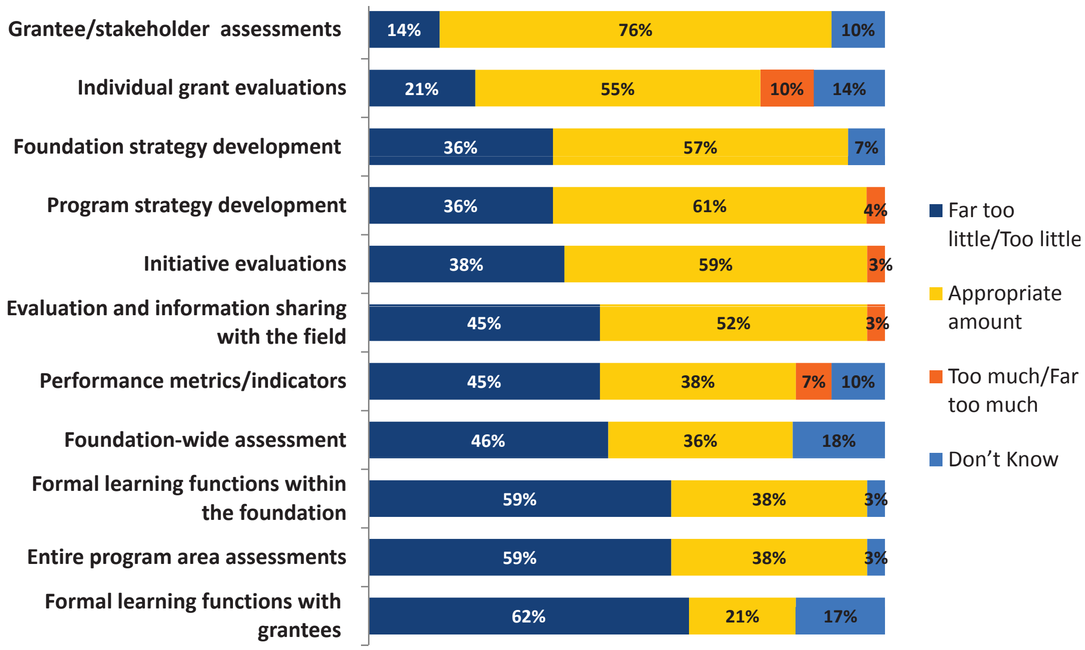
If foundations are addressing process barriers to learning and producing more timely evaluative information without seeing clear evidence of improved evaluation use, it is worth investigating what other factors are blocking evaluation use, particularly as a strategy is unfolding. Benchmarking research participants raised two possible explanations.

First, a growing emphasis on particular kinds of measurement is detracting attention from other types of evaluative work that could help to guide strategy implementation. Several foundations reported committing an increasing portion of evaluation resources to collecting and aggregating data that answer oversimplified questions, such as, "Did we move the needle?" and "Did we achieve our objectives?" These are not questions that help navigate the complex problems that many foundations are addressing. Often the foundation approach to strategy development superficially simplifies the situation and obscures the need for knowledge that will arise during the strategy process. Many foundations complete their strategy process and then fail to stay in touch with the inevitable changes that could hamper strategic success. This oversimplification has a way of becoming encrusted in everything that follows – evaluation questions, monitoring plans, conversations with board and management (Patrizi, 2012).

Additionally, too much focus on reporting performance indicators rather than investing both time and dollars in deeper evaluative work can thwart learning, and staff and boards might mistakenly assume that positive movement on indicators alone means strategies are effective, or the converse. Foundation choices about the scope of evaluative activities and the questions that drive data collection can have a profound impact on real learning.

The second factor affecting evaluation use relates to how institutions, groups, and individuals process information and make choices. Foundation interviewees raised concerns that despite more formalized or institutionalized learning and reflection time, it is not clear that foundation staff or grantees behave differently as a result of the findings or lessons even when they participate in generating them. It matters not just that decision-makers meet to learn; it also matters how they meet, digest information, and apply it. To ensure that evaluation is used, the structures and processes that support learning in foundations need to do a better job of connecting learning to actual foundation questions and decisions (while also making sure the right questions are being asked). In addition, learning processes need to counteract common decision-making habits that help people ignore data, even when it is right in front of them.

FIGURE 5 Perceived Sufficiency of Investments in Evaluation



Foundations Are Trying to Do "Broad Scope" Evaluation

Foundations have different levels at which they might examine their grantmaking: individual grants, clusters of grants, initiatives, portfolios, whole program areas, and the overall foundation. Numerous foundations now have evaluation activities targeted at many or all of these levels.

Foundation demand for evaluation always has existed at the grantee level, but it has increased substantially at the initiative level in the last two decades as strategic philanthropy has taken hold and many foundations have begun funding initiatives focused on their own goals. Almost every foundation that participated in the 2012 benchmarking research was evaluating at the initiative level, and it was one of the top two most time-consuming activities for evaluation staff in small, midsize, and large foundations. (See Figure 2.)

Especially in recent years, demand also has increased even farther up the grantmaking chain to include evaluation of whole program areas (which may include multiple initiatives) and assessment at the overall foundation level. In 2011, the Center for Effective Philanthropy found that almost half of foundation CEOs surveyed said they were conducting foundationwide performance assessment (Buteau & Buchanan, 2011).

Evaluation investments, however, still tend to be focused more at the grantee end of the scale. Foundation respondents in the 2012 benchmarking research said that the sufficiency of investments in evaluative work was weighted toward smaller-scope evaluation at the individual grantee and initiative levels. (See Figure 5.)

But in addition to questions about the amount of evaluation investments foundations are making, benchmarking interviews suggested that larger questions exist about how these investments should look as the scope for evaluation gets broader, as well as the expectations that should surround them. For example, as the use

Data gathered from individual grants and initiatives, where the majority of evaluation resources are focused, do not easily roll up into a neat package that illustrates foundation impact. Yet expectations are high for what broad-scope assessment might offer, even considering that for many foundations such assessment is trying to find meaningful impact data across disparate programming areas like education, health, the environment, international development, and the arts.

of initiative-level evaluation has increased, so have expectations of what it should deliver. While outcome-focused and impact evaluation were less common among foundation initiatives a decade ago (Walker & Grossman, 1999), initiative evaluations are now expected to be much more results-oriented, in spite of the many challenges of doing so with complex and often long-term initiatives (Patton, 2011; Rog & Knickman, 2002). Methods for evaluating initiatives have come a long way in the last two decades, and foundations are learning to apply different approaches to different types of initiatives (Britt & Coffman, 2012). But as numerous case studies about initiative-level evaluations produced for the Evaluation Roundtable have revealed, there is still much room for improvement as foundations attempt to meet expectations around generating data on initiative results while also generating data that help foundations and grantees learn and adapt along the way (e.g., Fiester, 2010; Parker, 2011).

As foundations move up the grantmaking chain toward assessment of whole program areas and the overall foundation, answering evaluative questions about results and impact and finding data that lead to meaningful learning gets even more complex. One challenge is that strategy – with a clear goal and clear and sound theory of change – does not really exist at this level. It becomes too high-level or diffuse to fit together in a way that is more meaningful than just a broad categorization of activities and results. Another challenge is that data gathered from individual grants and initiatives, where the majority of evaluation resources are focused, do not easily roll up into a neat package that illustrates foundation impact. Yet expectations are high for what broad-scope assessment might offer, even considering that for many foundations such assessment is trying to find meaningful impact data across disparate programming areas like education, health, the environment, international development, and the arts, where the problems and strategies needed to address them can be quite different. Yet in the Center for Effective Philanthropy's 2011 survey, foundation CEOs said one of the top reasons for doing assessment at the whole foundation level was to understand the external impact that can be attributed to the foundation's grantmaking (Buteau & Buchanan, 2011).

To try to meet the complex challenges of broad-scope evaluation and the high expectations surrounding it, benchmarking survey and interview data revealed foundations are turning to various tools. Recently, for example, dashboards that gather data on key foundation-level metrics have been a popular choice. As a study on foundation use of dashboards noted:

> Some foundations create dashboards with an operational focus, looking at day-to-day-level data like the number of requests in the pipeline or metrics to track the efficiency of their process. The vast majority include program spend information to see the money allocated compared to budgeted for a program, and a few have created dashboards to track indicators of program impact. (Idealware, 2013, p. 3)

For operational and program-spend data, dashboards can make good sense. Aggregating data on key metrics like the number of requests in the pipeline or grantmaking-process efficiency can tell foundation leaders a lot about where and how the grantmaking process might be improved. But when dashboards are expected to offer data on impact, the concern is that they are oversimplifying and overestimating impact and leaving foundation leaders with few answers about what to conclude from the data. Because it is not possible to aggregate impact metrics across program areas, dashboards might instead, for example, focus on population-level metrics that make too big a leap between an individual foundation's grantmaking and its attributable impact.

The same concern can be raised around another emerging trend – the use of "big data." Big data refers to the search for meaning in massive, multiple data sets that are cross-referenced and sortable, the purpose being to uncover more than any one data set can tell (Boland, 2012). Questions about how to use big data for evaluation came up frequently in 2012 benchmarking interviews. While exploring this is worthwhile, foundations should keep in mind that the answers to complex questions about broad-scope impact, and especially about how a foundation can use this information to maximize its impact, are unlikely to be found in big data alone. As Jacob Harold, Guidestar president and former head of the philanthropy program at the Hewlett Foundation, put it,

> We're not quite ready for it [big data]. Instead, we need to get "medium data" right first. … Medium data is a humbler but essential prerequisite: structured information about who you are, what you're trying to do, and what's happening. (Harold, 2013, para. 1, 2)

The benchmarking data in Figure 5 show that foundation respondents felt evaluation investments at the broad-scope level were insufficient. But before more investments are targeted at this level, it is important to reexamine the purpose of broad-scope evaluation. If it is to provide a picture of foundation-level aggregate impact, serious questions exist about whether that is even a useful or possible endeavor, especially with the tools foundations are using to address it.

Leadership Is on Board, But Evaluation Use and Incentives Remain an Issue

How foundation CEOs and boards communicate about and use evaluation affects a foundation's culture around data, information, and learning. Benchmarking data are mixed on the extent to which foundation leaders support evaluation. On the positive side, 86 percent of participating foundations said management often or frequently communicates to staff that it values the use of evaluative information. In addition, three-fourths said management often values efforts that illustrate the shortcomings of the foundation's work, a finding that is particularly evident when evaluation directors report to the CEO. Likewise, at least three-fourths said foundation boards show moderate to high support for the importance of evaluative information, role of evaluation staff, and evaluation's ability to provide a third-party perspective.

Both survey and interview data also revealed, however, that while foundation leaders communicate the value of evaluation, this support does not often translate into the meaningful integration of evaluative information into the business of philanthropy. Only a little more than one-third (38 percent) of foundations said that management regularly models the use of evaluation in its own decision-making. In addition, less than half (48 percent) said management regularly addressed problems identified in evaluations.

Staff also appear to have few incentives for integrating evaluative information into programmatic decision-making. When asked how frequently management considers the effective use of evaluative information as important in assessing staff performance, half of respondents said rarely or never; another quarter did not know.

These data raise questions about what constitutes meaningful leadership buy-in for evaluation. The message about evaluation's importance for strategic philanthropy has taken hold, but more thinking is needed about how managers and boards can drive organizational behavior about evaluation through incentives and modeling.

As suggested earlier, there also appear to be problems with the evaluative questions foundation leaders are asking and the information being supplied to them in response (often dashboard-style data). To make executive-level decisions about foundation strategy, leaders need to ask much more than just, "Did we move the needle?" They also need to dig deeper and ask why (or why not), how, and with whom. When these types of questions are asked, dashboard-style data alone become insufficient. Dashboard metrics often answer questions only about beginning and end states, regardless of what happened in between and regardless of whether the change reasonably can be attributed to the foundation's efforts or those of its grantees. Dashboard metrics also assume that strategic objectives and theories are correct rather than provide information about whether they are correct or whether the foundation should change course (Patrizi, 2012). Ultimately, when foundation leaders fail to use evaluation in their decision-making, it appears to be less about their willingness to use it than it is about achieving a better fit between the decisions leaders need to make, the right questions to inform those decisions, and the production of data and information to answer those questions.

Conclusion

Looking back at the history of evaluation in foundations, it is clear that some profound changes have occurred. Commitment to evaluation in philanthropy is now deeply rooted and no longer considered a trend that might fade. In fact, 2012 benchmarking data indicate that foundation commitment is increasing and that investments in evaluation are expanding (although so are the expectations that surround it). While individual foundations from time to time still do make substantial shifts in their evaluation direction and resourcing, the majority of foundations remain consistently committed to the endeavor.

Evaluation also now is connected much more closely to programming and strategy. Evaluation and program staff work side by side and share responsibilities, evaluation directors hold high-level positions, and evaluation is seen as a support for strategy rather than just as something to judge it. This is a profound shift in how evaluation is positioned in foundations.

At the same time, a number of issues about evaluation in foundations remain longstanding problems. The main challenge is still how to increase evaluation use, and the most common response to that challenge is still to focus on the supply side. Foundations are producing more data and information than ever. They are also presenting it in more ways than ever; there are fewer thick evaluation reports merely gathering dust on shelves than there used to be.

But while the supply side of the equation is important to addressing the challenge of use, the demand side needs more attention. Rather than assume that more data is better, this research reveals that foundations should ask themselves a set of questions that might help them to maximize the utility of their evaluation investments:

- Are resources (dollars and staff) being spread too thin across too wide a range of evaluation activities or across too many units of analysis? Is it better to prioritize and go deeper where it counts most?
- What does a dedication to "learning activities" really mean for the foundation?
- What evaluative questions should CEOs and boards ask that could actually inform executive-level decisions?

References

Boland, S. (2012). Big data for little nonprofits. *Nonprofit Quarterly.* Available online at http://www.nonprofitquarterly.org/management/21410-big-data-for-little-nonprofits.html

Brest, P. (2012). A decade of outcome-oriented philanthropy. *Stanford Social Innovation Review, 10*(2), 42-47.

Britt, H., & Coffman, J. (2012). Evaluation for models and adaptive initiatives. *The Foundation Review, 4*(2), 44-58.

Buteau, E., & Buchanan, P. (2011). *The state of foundation performance assessment: A survey of foundation CEOs.* Cambridge, MA: Center for Effective Philanthropy.

Center for Evaluation Innovation. (2013). *Evaluation in foundations: 2012 benchmarking data.* Washington, DC: Author. Available online at www.evaluationinnovation.org

Fiester, L. (2010). *Measuring change while changing measures: Learning in, and from, the evaluation of Making Connections.* Baltimore: Annie E. Casey Foundation.

Foundation Center. (2013). *Top funders: Top 100 U.S. foundations by asset size.* Available online at http://foundationcenter.org/findfunders/topfunders/top100assets.html

Frank, C. J., & Magnone, P. F. (2011). *Drinking from the fire hose: Making smarter decisions without drowning in information.* New York: Portfolio/Penguin.

Hall, P. D. (2004). A historical perspective on evaluation in foundations. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation* (pp. 27-50). San Francisco: Jossey-Bass.

Harold, J. (2013, March 20). Nonprofits: Master "medium data" before tackling big data. *Harvard Business Review Blog Network.* Available online at http://blogs.hbr.org/cs/2013/03/nonprofits_master_medium_data_1.html

Heimann, F. F. (1973). *The future of foundations.* Englewood Cliffs, NJ: Prentice Hall.

Idealware. (2013). *Data at a foundation's fingertips: Creating and building dashboards.* Wayne, PA: Technology Affinity Group.

Parker, S. (2011). *Evaluation of the David and Lucile Packard Foundation's Preschool for California's Children grantmaking program.* Washington, DC: Center for Evaluation Innovation.

Patrizi Associates. (2010). *Use of evaluation information in foundations: Benchmarking data.* Available online at http://www.evaluationroundtable.org/documents/et-benchmarking.ppt

Patrizi, P. (2012, December). *The North Star and the swamp: Getting better at learning from strategy.* Report presented at a convening of the Evaluation Roundtable, New York City.

Patrizi, P., & McMullan, B. J. (1999). Realizing the potential of program evaluation: A new study looks at how 21 foundations "do" evaluation. *Foundation News and Commentary* (May-June), 30-35.

Patrizi, P., & Thompson, E. H. (2011). Beyond the veneer of strategic philanthropy. *The Foundation Review, 2*(3), 52-60.

Patton, M. Q. (2011). *Developmental evaluation: Applying complexity concepts to enhance innovation and use.* New York: Guilford Press.

Patton, M. Q., & Patrizi, P. A. (2010). Strategy as the focus for evaluation. In P. A. Patrizi & M. Q. Patton (Eds.), *Evaluating strategy: New directions for evaluation, No. 128,* (pp. 5-28). San Francisco: Jossey-Bass.

Preskill, H., & Beer, T. (2012). *Evaluating social innovation.* Washington, DC and Seattle: Center for Evaluation Innovation and FSG Social Impact Consultants.

Rog, D., & Knickman, J. (2002). Strategies for comprehensive initiatives. In M. T. Braverman, N. A. Constantine, & J. K. Slater (Eds.), *Foundations and evaluation* (pp. 223-235). San Francisco: Jossey-Bass.

Thompson, E. H., & Patrizi, P. (2010). *Necessary and not sufficient: The state of evaluation use in foundations.* Philadelphia: Evaluation Roundtable.

Walker, G. & Grossman, J. B. (1999). *Philanthropy and outcomes: Dilemmas in the quest for accountability.* Philadelphia: Public/Private Ventures.

Julia Coffman, M.S., *is founder and director of the Center for Evaluation Innovation. Correspondence concerning this article should be addressed to Julia Coffman, 1625 K Street NW, Suite 1050, Washington DC 20006 (email: jcoffman@evaluationinnovation.org).*

Tanya Beer, M.A., *is associate director of the Center for Evaluation Innovation.*

Patricia Patrizi, M.S., *is principal at Patrizi Associates.*

Elizabeth Heid Thompson, B.A., *is an independent consultant.*