The Center on Philanthropy & Public Policy

INCREASING EVALUATION IMPACT: A SURVEY OF US PHILANTHROPIC FOUNDATIONS

Howard P. Greenwald

Research Paper - 32
July 2009

UNIVERSITY OF SOUTHERN CALIFORNIA

INCREASING EVALUATION IMPACT:
A SURVEY OF US PHILANTHROPIC FOUNDATIONS

Howard P. Greenwald

Research Paper - 32
July 2009

This research was supported by the California Community Foundation Endowed Research Fund at The Center on Philanthropy and Public Policy

Howard P. Greenwald: Professor, School of Policy, Planning, and Development, University of Southern California; Lewis Hall 228; Los Angeles, CA 90089

ABOUT THE CENTER ON PHILANTHROPY AND PUBLIC POLICY

The Center on Philanthropy and Public Policy promotes more effective philanthropy and strengthens the nonprofit sector through research that informs philanthropic decision making and public policy to advance public problem solving. Using California and the West as a laboratory, the Center conducts research on philanthropy, volunteerism, and the role of the nonprofit sector in America's communities.

In order to make the research a catalyst for understanding and action, the Center encourages communication among the philanthropic, nonprofit, and policy communities. This is accomplished through a series of convenings and conversations around research findings and policy issues to help key decision makers work together more effectively to solve public problems and to identify strategies for action.

The opinions presented in this paper represent those of the authors and not those of The Center on Philanthropy and Public Policy.

The Center on Philanthropy and Public Policy
School of Policy, Planning, and Development
University of Southern California
Lewis Hall, Room 210
Los Angeles, California 90089-0626

Printed in the United States of America

About the Author

Howard P. Greenwald is Professor of Management and Policy at the University of Southern California School of Policy, Planning, and Development. He is a specialist in program evaluation, organizational performance, health services research, and chronic disease epidemiology. He holds a Ph.D. in Sociology from the University of California at Berkeley. He has served as a faculty member at the University of Chicago Graduate School of Business, research scientist at Battelle Memorial Institute, Chairman of the Network for Healthcare Management, Director of the Health Services Administration Program at the University of Southern California, and Commissioner on the Accrediting Commission for Education in Health Services Administration. His current research activities include studies of innovation and effectiveness in formal organizations, the political process of policy-making, vehicle stops by law enforcement officers, long-term quality of life among cancer survivors, and outcomes of multisite interventions designed to improve the quality of life in communities. His forthcoming book, *Health Care in the United States: Organization, Management, and Policy* (San Francisco: Jossey-Bass, 2009), presents an integrated view of the health care system's components and environment. He has written four other books, the most recent of which are *Organizations: Management Without Control* (Thousand Oaks, CA: Sage, 2007) and *Health For All: Making Community Collaboration Work* (Chicago: Health Administration Press), with William L. Beery. He is author alone or in collaboration of approximately 50 peer-reviewed articles in journals such as the *American Journal of Evaluation, the Journal of Clinical Epidemiology, the American Journal of Public Health, the Journal of Women's Health,* and *the Milbank Quarterly*. His opinion pieces have appeared in the *New York Times*, the *Wall Street Journal*, and the *Sacramento Bee*.

Abstract

The objectives of this paper are to (1) assess the function of evaluation in US philanthropic foundations; (2) suggest means by which evaluation can be made to play a more effective role; and (3) identify emerging issues in evaluation. Based on interviews of 24 evaluation directors and top executives of foundations, this paper addresses utilization of program evaluation, its application within foundations, challenges encountered in the evaluation process, issues for professional evaluators, and potential steps that foundations may take to increase the usefulness of evaluation. Most of those interviewed report that their foundations evaluate the programs they sponsor either through in-house resources or outside contractors. However, the funding that foundations allocate for evaluation varies from less than 1 percent to 15 percent of their typical extramural budgets. Most informants said they valued a variety of methods rather than a single approach to evaluation. A majority considered evaluation very important and almost all very or somewhat important. Evaluation findings occasionally led to reduction or termination of funding for individual grantees. More often, however, evaluation served as a basis for deciding to provide technical assistance to grantees experiencing difficulty or for planning future grantmaking. Informants thought evaluation could be made more useful by: reporting results more promptly and in a manner better adapted to foundations' decision-making needs; focusing on principal findings rather than technical details or disclaimers; and, adopting a position of colleagueship with program officers and other foundation personnel. Within foundations, several steps can promote greater impact of evaluation on attainment of goals. These include: dialogue between evaluation and non-evaluation personnel; consensus regarding objectives and methods of evaluation; and, safeguards to the integrity of the evaluation process. Key issues emerging in evaluation of philanthropic foundation activity include (1) an increasing interest among foundations in evaluating not individual programs or interventions but achievements of the foundation as a whole, and (2) the need for evaluation to further develop as a science that produces both rigorous and meaningful findings.

Acknowledgments

The author wishes to thank the following individuals for serving as informants in the research reported here: Fatima Angeles (The California Wellness Foundation), Lester W. Baxter (Pew Charitable Trusts), Theresa Behrens (formerly W.K. Kellogg Foundation), Gale Berkowitz (Lucile Packard Foundation), Martha Campbell (Irvine Foundation), David Colby (Robert Wood Johnson Foundation), Marie Colombo (Skillman Foundation), Nancy Csuti (Colorado Trust), Tom David (Tides Foundation), Astrid Hendricks (The California Endowment), Jackie Williams Kaye (Atlantic Foundation), Tom Kelly (Anne E. Casey Foundation), Diane Littlefield (Sierra Health Foundation), Nancy MacPherson (Rockefeller Foundation), Ann McCracken (Health Foundation of Greater Cincinnati), Kenneth Moore (Moore Foundation), Gary Nelson (Georgia HealthCare Foundation), Ed Pauly (Wallace Foundation), William Richardson (Formerly W.K. Kellogg Foundation), Steven Schroeder (Formerly Robert Wood Johnson Foundation), Edward Sermier (formerly with Carnegie Foundation), Marsha Shotley (Blue Cross Blue Shield of Minnesota Foundation), Mary Williams (Lumina Foundation), Jill Yegian (California HealthCare Foundation). The author accepts responsibility for any and all errors.

Introduction

This paper has three principal objectives: (1) describing the ways in which US philanthropic foundations think of and utilize evaluation findings, (2) suggesting means by which foundations may obtain greater value from evaluation in the future, and (3) identifying emerging issues in evaluation of special concern to foundations. The research reported here examines factors both inside and outside philanthropic foundations that may affect the utility and actual use of evaluation. Inside foundations, the process of program planning, organizational factors, and interpersonal relationships affect the quality of evaluation obtained and its ultimate utilization. External to foundations, the nature of programs to be evaluated and the behavior of outside evaluation specialists affect the quality, relevance, and utilization of evaluation findings.

Philanthropic foundations today may spend several million dollars to evaluate a single intervention or program. Yet, the comments of some observers suggest that evaluation does not currently achieve its full potential for benefitting the work of foundations.[1,2] Some suggest that the role of evaluation, whether for philanthropic foundations or other intervention sponsors, is marginal at best. Bolman and Deal, for example, have written:

> Evaluation consumes substantial time, effort, and money. It typically yields a lengthy report presented with considerable ceremony. Yet rarely are insights or recommendations implemented. Results typically disappear into the recesses of people's minds or the far reaches of administrators' file cabinets.[3]

Other observers have been more optimistic. A number of researchers have viewed potential impact of evaluation not merely as instrumental in the sense of affecting immediate decisions or action but as part of a gradual learning process.[4] Significant research has taken place on factors potentially affecting the utilization of evaluation findings. Efforts in this area have addressed methods of dissemination,[5] decision-making theory,[6] and formal modeling.[7] While making theoretically informed contributions to the evaluation field, however, this research has generally not examined the concerns and behavior of actual users of evaluation.

The work reported here aims primarily at identifying factors that may promote or inhibit the ability of foundations to obtain maximum value from evaluation. Conclusions are based on the outlook, experience, and concerns of key individuals at US philanthropic foundations directly and continuously involved in evaluation. Special attention has been paid to developing recommendations for how vendors of evaluation services can serve foundations better and identifying steps that foundation personnel can take to obtain greater benefit from the evaluation work that they sponsor.

Methodology

Data for this paper were obtained by interviewing evaluation directors or chief executive officers (CEOs) of relatively large US foundations. A purposive sample of evaluation

directors was obtained in a multi-stage process. First, ten individuals were identified as evaluation directors at the 100 largest US foundations by asset size. Second, in the course of ensuing interviews, these individuals were asked to identify other evaluation directors who could contribute to developing an understanding of the use of evaluation by foundations. An additional ten individuals were identified in this snowball sampling procedure. The snowball procedure was discontinued when interview subjects began suggesting the same individuals. Four CEOs or individuals who had recently left such positions were identified to obtain opinions from an alternate perspective.

Semi-structured interviews were conducted via telephone with each of the individuals identified. Persons interviewed were considered to be *key informants* on the thinking and activity of US philanthropic foundations regarding evaluation. Key informant interviews are distinguished from those comprising opinion surveys in that individuals interviewed are asked not just about their personal opinions and behavior but "about patterns of behavior … either observed or expected" in their social surroundings.[8] Unlike interview subjects in opinion surveys, key informants are not selected to be representative of a population in a statistical sense. According to one formulation, key informants "are chosen because they are … knowledgeable about the issue being researched and able and willing to communicate about them."[9]

None of those asked to be interviewed refused. A total of 24 interviews were conducted, including informants associated (presently or formerly) with 22 separate foundations. In two instances, both research directors and former CEOs from the same foundation were interviewed.

Responses to interview questions were aggregated on the basis of perceptual agreement among informants. Perceptual agreement is a method often used in drawing inferences from interviews given by multiple informants about a higher unit of analysis such as an organization. In the study reported here, patterns of convergence among informants in different organizations were sought for the purpose of developing information applicable to an industry, that of philanthropic foundations. Although the resulting findings cannot be safely generalized beyond the sample obtained, they illustrate a range of activities, issues, and solutions capable of supporting practical suggestions to both foundations and evaluators.

Each interview lasted from 20 minutes to one hour. Informants were asked to focus on evaluation defined as "systematic assessment of the implementation or outcomes of a grantee's work, through qualitative or quantitative means." The interview schedule contained items requesting information on;

- Resources committed to evaluation;
- Methodological preferences in evaluation;
- The role of the evaluation director;

- Use made of evaluation;
- Value placed on evaluation by the foundation;
- Changes in foundation thinking or actions due to evaluation findings;
- Ways to make evaluation more useful to foundations.

As a follow-up to the interviews, all informants were asked to review an initial draft of the study findings. The informants were requested to identify any errors in information or quotations from their interviews that may have been made. The informants were encouraged to provide any additional comments they considered relevant, and the comments received were incorporated into the results reported here.

This research concentrates specifically on evaluation of interventions carried out by external grantees to change the thinking, behavior, and surroundings of human beings in a manner consistent with foundation goals. Interventions in this sense may include individual grants or initiatives under which multiple grantees are funded to pursue common objectives specified by the foundation. The findings presented below do not address several activities of foundations that are important but not usually subject to evaluation. Such activities include charitable giving and research funding.

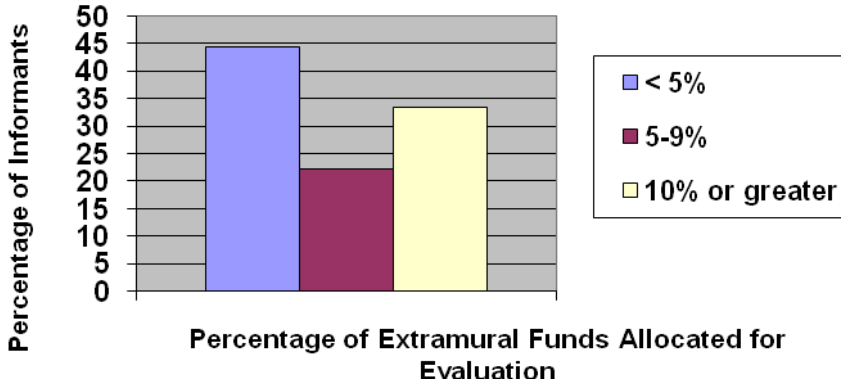
Current Evaluation Practice

All informants indicated that their foundations obtained evaluation of at least some of the interventions they funded. This is not surprising in view of the fact that individuals selected for interviewing were current or past foundation evaluation directors or officers of foundations know to conduct or contract for evaluation. Comments of these individuals suggest variability in allocation of resources to evaluation and, to a lesser extent, evaluation methodology favored.

Allocation of Resources

Of the 22 separate foundations contacted in this study, all divided the evaluation task between in-house personnel and outside contractors. Evaluation was usually done on larger-scale interventions. The larger the scale of an intervention the more likely it was to be evaluated by an outside contractor. When asked the percentage of their yearly extramural funding that was allocated to evaluation, informants gave figures that ranged from less than one percent to 10-15 percent. As indicted in Figure 1, a bimodal distribution prevailed in levels of resources allocated for evaluation.

Figure 1. Extramural Funds Allocated for Evaluation



The range of resource allocation for evaluation may indicate differences in commitment to evaluation across foundations. However, the percentages presented above must be understood in the context of limited data. Four respondents said they did not know the percentage of extramural funds their foundations allocated for evaluation. Several others indicated that funds for evaluation sometimes appeared as budget items under other labels or were allocated to intervention grantees and thus not reflected in a global evaluation budget.

Methods in Use

In general, informants did not identify a particular methodology that they favored for evaluation. Most informants said that the substance of an intervention should determine the methodology used. Some illustrative comments included the following:

> Context and nature of program define best methodology. We've used them all.
>
> We use all approaches. All are valuable, and most evaluation uses mixed methods. We use a combination of quantitative and qualitative methods and triangulate to identify converging lines of evidence.
>
> I tend to like to use as many (methods) as possible; you cannot be set on just one (and be able to) get the best overall picture.

Informants expressed interest in several formal methods, some traditional and some of relatively recent introduction. Among more recently-introduced methods were logic model case studies, of interest in part because of their duel applicability in both evaluation and planning. Informants expressed interest in networking and GIS-based methods. More traditional methods in widespread use included survey research.

Experimental and quasi-experimental methods enjoyed less popularity, informants commenting that these were difficult to apply to the interventions they funded. Although randomized controlled trial methodology was generally deemed infeasible, experimentation was considered valuable for specific purposes. Regarding an intervention intended to promote voting among young people, for example, an informant said that his foundation had used experimental methodology to compare alternative approaches.

Participatory evaluation was considered important by a small number of informants. Some cited value in this method beyond evaluation *per se*. Informants, for example, commented that participatory evaluation and tools used in its application promoted involvement and receptiveness among stakeholders. An informant of this opinion commented,

> We see value in youth-led evaluation. Innovative approaches around digital storytelling, photojournalism are user-friendly evaluation methods. Kids are much more responsive to this than to reports – and their social reality changes very fast, so written material goes out of date.

Purposes Of Evaluation

Prima facie, the purpose of evaluation would be simply to determine the efficacy of individual interventions. But comments in the interviews also cited less obvious purposes as well. Themes of this nature emphasized efficacy of the foundation as a whole rather than that of individual grantees. The following comment is illustrative:

> We're interested in the overall impact of our grantees' work on achievement of the foundation's goals. We're interested in knowing how grantees' work feeds into the benchmarks for the foundation's theory of change.

Other than reporting on efficacy, evaluation was seen as serving more fundamental purposes in the development, operation, and social acceptance of the foundation. Purposes cited by informants included promotion of accountability, calibration of interventions, and organizational learning.

Accountability

Accountability proved a repeating theme in informants' comments regarding the uses of evaluation. On an elementary level, evaluation was seen as a means of determining whether a grantee was acting in a manner consistent with its commitments. More fundamentally, the theme of accountability was expressed as a belief that a foundation needed to demonstrate that it was doing useful things with its resources. As one informant commented, "the purpose of evaluation is to inform a foundation's work, so (foundations) can become better grantmakers and have better accountability to their missions and the communities they serve."

Informants cited a number of stakeholder groups to which they considered themselves accountable. Stakeholders internal to foundations included members of governing boards and other staff personnel. Respondents associated the concept of accountability with functions such as quality control and continuous quality improvement in grantmaking. Evaluation was seen as an essential resource in the process of improvement.

External stakeholders to which interview subjects considered their foundations accountable included "the public" and "the community." The evaluation director of a prominent foundation considered it important to demonstrate his organization's efficacy to groups such as "grantees, thought leaders, business leaders." He commented:

> (As evaluation director), I am responsible for holding (the foundation) accountable. We're responsible for saying how we're doing as an organization. We are also responsible for seeing whether our strategies are having impact.

This informant emphasized large-scale, foundation-wide evaluation as most relevant to transparency and accountability. He commented that during his tenure the foundation had done many intervention-specific evaluations but that the "key evaluation question was how we did as a whole."

Calibration of Interventions

As an objective of evaluation, informants often cited its use as a means of learning the features of an intervention that were working and those that were not. Several mentioned that they were more interested in helping an apparently-foundering grantee than "pulling the plug." Aid to such a grantee might take the form of additional funding or technical assistance. Few individual grantees appear to be defunded as a result of evaluation.

Learning about implementation of and fidelity to objectives, as well as acceptability to the affected community, seems particularly important in multi-year programs or initiatives in which awards are made to successive cohorts of applicants. According to one informant,

> Evaluation has not affected the funding of individual grantees. But lessons from evaluation are immediately incorporated in technical assistance and reviews of new grant applicants. A strong feedback loop is established between interim results and further work by later grantees.

As an example, this informant described an intervention intended to assist school officials and build their leadership capacity:

> We wanted to determine the characteristics of people who seek services and the characteristics of people who don't participate. Results of this kind are not the same as evaluation of grantee performance. We wanted to find out what leaders need to know … These people want to solve a particular set of problems (and we needed a better understanding of what the problems are).

Organizational Learning

Several interview subjects emphasized that evaluation helps them build a model of how the world works and how to change it. Accordingly, evaluation contributes to a continuous process that builds understanding of factors that promote, facilitate, or challenge achievement of a foundation's goals. Greater understanding of how their initiatives fit with the educational establishment, politics of a profession, or history and structure of a community raises the likelihood that an intervention will be effective.

Development of such understanding constitutes a form of *organizational learning.*[10] Organizational learning is a process by which an organization over time develops core beliefs and principles of operation. Both successful and unsuccessful experiments and ventures constitute valuable lessons in the learning process. Such lessons are incorporated into the organization's culture and transmitted to new generations of members.

The comments of several informants provide illustration. One characterized evaluation as tightly linked with an iterative learning process in her foundation. This process helps the foundation clarify its thinking about what it really wants to accomplish and to develop strategy for future grantmaking. In describing a long-term program to which her foundation had committed significant resources, another informant commented that evaluation of the program had helped the foundation "learn to execute a fundamental strategy" regarding social change.

Even disappointing evaluation findings were seen as having value in promoting organizational learning. A former foundation president described an intervention intended to create a "huge national network of home helpers." However, the program proved infeasible because there were not enough home helpers available nationally to complete the hoped-for network. In this instance, evaluation contributed to the foundation's understanding of the social reality into which its intervention was inserted. Generally, the informant concluded, evaluation helps to "assess the foundation's understanding of the field (in which it contemplates intervening) and whether (the intervention) is "planned and framed appropriately, not naïve, underpowered, (or otherwise unlikely to succeed)."

Although distinct purposes can be identified for evaluation, the three stated here are logically interconnected. Accountability in the sense of a foundation's ability to achieve demonstrable results is advanced by calibration of interventions through evaluation findings. Ideally, lessons learned about the feasibility and impact of interventions would be incorporated through an organizational learning process into the foundation's future grantmaking decisions.

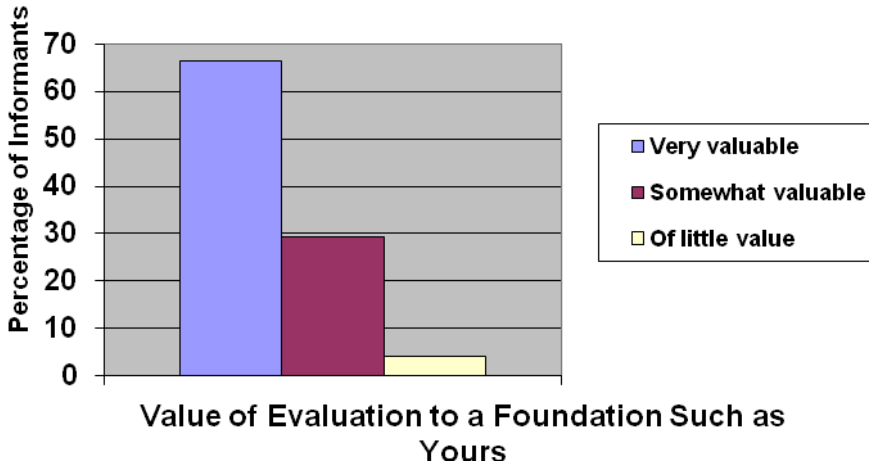
Perceptions of Value and Impact

Value of Evaluation

As indicated in Figure 2, informants generally found evaluation as they had seen it done in their foundations to have been of value. Of 24 respondents to the survey, 16 considered evaluation to have been very valuable, seven somewhat valuable (including two who answered "very to somewhat"), and one of only a little value. Reasons for the predominantly favorable assessment of evaluation are illustrated as follows:

> I don't think we could operate without evaluation. It informs our process. (It helps inform us about) what to look at when senior project officers go on site visits.
>
> (At our foundation) evaluation is integrated (into the thinking of) program teams as they develop grants and is aligned with strategies. We are not just responsible for external evaluation but for integrating it into the core of the foundation.
>
> … good evaluation can lead to wider adoption (of a program), e.g., by government.

Figure 2. Value of Evaluation



Bias regarding the value of evaluation is, of course, expectable among directors of evaluation, the position held by most people interviewed here. However, the foundation CEOs interviewed for this study tended to respond similarly. Both evaluation directors and CEOs cited a broad range of effects that evaluation had had on the thinking and actions of their foundations.

It is important to note that a foundation's allocation of relatively small budgets for evaluation of the interventions it funds does not necessarily reflect the perception of evaluation *per se* as low. Instances may be cited in which outcomes of an intervention are not in significant doubt, and thus do not require major expenditures for evaluation. According to one informant,

> A foundation doesn't have to allocate funds that could be used for service delivery (for) an expensive research methodology to produce evidence that an evidence-based practice works!

Impact of Evaluation

Every informant asked was able to describe one or more instances in which evaluation had had a concrete effect on his or her foundation's thinking or actions. Some of the instances described related largely to thinking and actions within the foundation. Others had impact outside the foundations that were viewed as advancing the foundation's goals.

The clearest impact of evaluation may be seen in instances when a foundation incorporated evaluation findings into a decision to continue, modify, or discontinue a program. An informant, for example, reported that positive evaluation findings made the difference between continuation and abandonment of a program aimed at promoting health among immigrants:

> (Evaluation looked at) outcomes targeted by grantees such as knowledge, awareness, screening, and health status, and showed that (the program) actually has improved outcomes. We are continuing … the program. If there had been results that showed it didn't work we perhaps would have looked at another program or changed components. (Evaluation findings have) verified that the current focus has provided value to the community so we continue to move ahead with the present program.

Informants cited several instances in which evaluation promoted modification of an existing program or adoption of a new approach to achieving the original objective. Examples were cited in a variety of subject areas:

> In (an intervention to) strengthen higher education in Russia, we changed direction (as a result of evaluation findings), placing less emphasis on research and more on travel and interaction among faculty.

> (In an initiative to preserve Amazon wilderness we realized) … that we were not being successful working toward this objective…(and) needed to redeploy resources into what is now called protected area consolidation (promoting conservation while fostering ecologically compatible socioeconomic development).

> Through a thoughtful evaluation we've learned that working with people in public school districts is a limited way to go. We revised (the program's approach). We went to a strategy of working with individual schools and independent and parochial schools as well as public schools.

A policy impact was cited in a program called *Pathways to Teaching Careers*, an educational intervention, but only after evaluation had contributed to the foundation's learning process:

> The initiative involved partnerships between organizations that train teachers and school districts in need. We found a non-traditional source of personnel, teacher aides. Many couldn't complete college due to financial constraints but had qualifications (relevant to becoming teachers). Evaluation findings in this area were presented to staff members of a Congressional committee and resulting legislation provided public support for the *Pathways* model.

Another informant described the influence of lessons learned through evaluation on the efficiency of foster care programs:

> We had a big impact on (reducing the length of time that children spent in) foster care in early 90s. (Evaluation studies) found that (agencies involved in foster care) including the courts weren't coordinated. A large part of the problem was that (the courts) kept putting off hearing dates to move (the process) into the next phase – judges were detached from the kids. In several jurisdictions we … we got judges to (realize) that they had a panel of (actual) children and they were responsible for moving them along … The judges agreed to look at data (on their individual performance to see whether they were lagging).

A foundation's work on end-of-life issues illustrates how initially negative evaluation findings can lead to interventions of great impact. As an informant described evaluation of a nurse training intervention in this area,

> The (evaluation) study came back with unanticipated, serious results. This was a study that had an intervention and control group. (We expected that) patients would experience less pain, have their wishes honored more often, be intubated less frequently... The experimental group didn't do any better (than the control group). Neither group's outcome was satisfactory.

According to the informant, the foundation started looking for alternative approaches and developed multiple, well-recognized interventions which helped initiate "a social movement to advance end-of-life care."

Challenges to Evaluation

Despite the positive examples cited above, informants described a number of challenges regarding evaluation likely to occur in foundations. Challenges may arise from the technical difficulties expectable when any formal research technique is applied in a real-world or action-research setting. Additional challenges may be associated with internal thinking and social dynamics of a foundation. Still other challenges appear to be emerging due to a shift in the focus and objective of evaluation at some foundations.

Methodological Challenges

Some of the challenges cited by informants arose from perceived limitations on methodological validity and objectivity. Others resulted from problems in execution of evaluation design related to conventions of confidentiality or insufficient cooperation by an agency. Comments regarding social limitations on the use of evaluation technology included:

> (In a mental health intervention that we funded) lots of confidentiality issues (got) in the way of assessing individual outcomes. So no client-level evaluation was conducted.

Other comments highlighted limitations on application for evaluation of methodology widely used by social scientists for more abstract research purposes. Data available from administrative agencies such as the US census were seen as difficult to apply to the specific communities in which interventions were taking place. Surveys were considered useful but expensive. Quantitative studies were considered inappropriate for issues related to policy.

Natural science methodology was seen as even more difficult to apply for evaluation purposes. As noted above, experimental methods with the rigor of randomized controlled trials are widely considered infeasible. Other research methods cannot provide evidence of causation with nearly the same strength. Hence, evaluators can seldom clearly attribute observation of developments consistent with an intervention's objectives to the intervention itself. Expressing a widely-held opinion, one informant commented that evaluation should concentrate on demonstrating an intervention's *contributions to* rather than *causation of* an outcome.

A number of informants expressed concern regarding the strength and cohesiveness of the evaluation field itself. The diversity of methods and approaches used by evaluators seems to have reduced the confidence of some foundation personnel in evaluation. As one informant commented,

> Tremendous variation exists in use of the term "evaluation." The term doesn't mean much now (because it includes so many) widely different approaches. There needs to be … more clarity and consistency in use of the term and a more standardized set of skills and competencies.

A hunger for new methodologies seemed evident. As an informant commented,

> We like designs that move away from more constrained, limiting strategies. We want to see more qualitative approaches that have both rigor and validity.

Social and Interpersonal Relationships

Differences in outlook between evaluation personnel and governing board members may reduce the role evaluation findings play in decision-making. A board member may devalue evaluation because the reports of evaluators do not meet their expectations. Some members of a foundation's governing board may expect numerical evidence of promptly-achieved outcomes. Such findings are rare in foundation program evaluation. Board members may discount results that contradict their hopes and impressions. Illustrative comments from an informant who played both an evaluation director and board member role include these:

> EO Wilson has written that man would rather believe than know. Foundations are bastions of self-congratulatory work. It is very easy to *believe*. People don't like to give bad news.

And further,

> I'm lucky. I have board that wants evaluation and a staff that listens. I've seen lots of other foundations that don't have that same set of conditions. I'm amazed that some foundations don't change things as a result of the evaluation reports they receive.

Staff relationships can also result in challenges to evaluation. According to a former foundation president, program officers – individuals involved in the actual development and management of interventions -- are not always responsive to evaluation results. This informant commented that program officers "are seriously overloaded and can't really be thoughtful about evaluation." Describing the importance of an oral tradition in philanthropy, he expressed the impression that "stories get told from program officer to program officer, and these have more weight than an evaluation report."

The structural relationships between evaluation and other foundation functions as well as the interplay of personalities that occurs in any organization may likewise reduce the quality of evaluation or diminish attention paid to it. In describing her foundation, an informant cited both phenomena:

> The organizational structure was very difficult. Foundations struggle with this. (In my foundation) a separate evaluation unit reported to the senior director for programs. There was built-in tension, with the evaluation unit responsible … having to negotiate with program directors, who held the budget for evaluation. There was back and forth (competition) for responsibility. Tensions around these struggles exacerbated the problem of using evaluation findings. (It was a challenge) to do the job and at the same time be accepted as colleagues. Personalities of key individuals added to difficulty.

Commitment and clarity of foundation leadership regarding evaluation also affects the quality of evaluation obtained and the benefits the foundation may draw from evaluation. Lack of consensus among board members can have an adverse impact. Members of a foundation board described by one informant, for example,

> ... never got clarity about what they wanted to get from evaluation, and the CEO was reluctant to engage them on this. (The CEO) was very conflict-averse. Individual board members had different conceptions of evaluation. One was experimental-oriented. Another said he just wanted stories. There was no clarity about why we were doing evaluation. Was it for accountability to the board? Were we doing it to influence policy?

In a similar set of comments, the evaluation director of another foundation reported that she "never got clarity from the board about the objective of evaluation."

Sensitivity to accountability issues by decision-makers appears to vary among foundations, which, in turn, affects commitment to significant and meaningful program evaluation. The comments of an informant regarding a large foundation illustrate how relatively weak concern for accountability may reduce commitment to evaluation:

> The problem is incredibly deep and difficult. How many people have you met who are willing to hold themselves accountable when there is no reason to be so? When you have exclusive control over $4 billion (do you) ask yourself how well you are doing? If you can pick all board members, ask yourself whether they will really be able to ask you anything. Will you ever evaluate your own performance? Everything follows from that basic issue.

Challenges to evaluation from outside foundations cannot be ignored. Informants raised issues about the validity of the information that eventually reaches decision-makers. Grantees, it was said, often do not want to share bad news with either program officers or evaluators. Evaluators themselves may be reluctant to transmit negative findings. Informants raised this possibility for both outside consultants and in-house evaluation personnel. Evaluation specialists of both types value good relations with program officers and thus may understate negative results.

Finally, social and interpersonal relationships affect the level of resources available for evaluation. An informant expressed the opinion that foundations whose benefactors are still alive and contributing funds are disinclined to allocate major resources to evaluation. The informant commented, "donors don't want to see funds used for evaluation; rather, they want them used for programs."

Emerging Trends and Challenges

The kind of evaluation that foundations find most valuable may be in flux. Some of the potential change in emphasis follows from new questions that philanthropic foundations and their critics have begun to ask. Other dimensions arise from a desire for new evaluation methodologies. Trends of both kinds pose challenges for evaluation.

Several informants suggested that foundations have begun seeking a broader perspective from evaluators than they have in the past. Some commented that it is uneconomical to

expend resources for evaluation on individual projects. According to this perspective, more value can be obtained by evaluating clusters of projects or sets of initiatives. Also of emerging prominence is evaluation covering performance of the foundation as a whole.

Specific comments illustrate a growing interest among foundations in evaluation efforts that cover a bigger picture than the past. According to one informant:

> The trend is toward evaluation of the cluster and initiative. Yet another (area of foundation interest) is evaluation on the strategy level. (Evaluation needs to help determine whether) the foundation as a whole is achieving its overall strategy. You can have an effective program that doesn't get you to a goal. Our big goal is to increase number of college degrees from 40 to 60 percent of working population by 2025. We want to evaluate a program's contribution to moving us (in the direction of) this goal.

Describing the objectives of in-house evaluation personnel, an informant commented:

> We're responsible for saying how we're doing as an organization. We are also responsible for seeing how our strategies are having impact. We pursue this objective through strategic maps, logic models, retrospectives, etc.

Informants reported a variety of activities intended to assess and report on progress in achieving overall goals. Associated methods include surveys of grantees and outside stakeholders. Several foundations now maintain scorecards and dashboards which regularly update their performance. One foundation hired a group of writers to summarize clusters of individual projects and their outcomes in a commercially published monograph.

Challenges associated with these broadened evaluation objectives include issues of measurement. Overall goals may be clear to foundation officials. But individual grantees are of necessity encouraged to identify their own specific objectives and to develop individual plans of action. Developing aggregate measures reflecting the success of an entire portfolio of initiatives will require significant effort and creativity.

Making information about achievement of broad foundation goals meaningful for decision-makers will constitute an additional challenge. Indicators reflecting the performance of a foundation as a whole are necessarily more abstract than measures regarding an individual grantee, program, or initiative. Nevertheless, governing board members will need to use such information to help make practical decisions about strategy and grantmaking. In the perspective of one informant, foundations need to

> … look for better knowledge capture around lessons learned, sharing findings among initiatives, (and) turning data into knowledge so we can learn to stop making the same mistakes (over and over again).

The possibility that both governing board members and public officials may make more stringent demands on non-profits (including foundations) in the future may result in additional challenges to evaluation. Increased demands for accountability may require evaluators to seek more direct evidence of efficacy than they have in the past. As if to forewarn of such pressure in the years to come, an informant observed that the contractor hired by his foundation to evaluate an urban education program

> … didn't analyze key issues (in helping us) understand whether (the program) was working. Never looked at student data. Same baloney as in many foundations. Lots of crap and never get to the central point about what we should be measuring no matter how difficult it is. I don't give a crap about whether (evaluators report that a school) administrator feels good. There's no accountability here.

The Evaluation Director's Role

Examination of the evaluation director's role provides additional perspective on the challenges identified above. In some foundations, the role of the evaluation director does not appear to be fully accepted by others. Cross-pressures reported by evaluation directors highlight some of the social dynamics within foundations that affect the potential use and usefulness of evaluation findings.

Evaluation directors described their role as including several concrete elements. Several indicated that they wrote evaluation RFPs and took the lead in selecting evaluation contractors. Some commented that they held training sessions on evaluation for other staff. Evaluation directors coordinate evaluation activities taking place in connection with separate projects to ensure that issues of concern to foundation decision-makers are addressed.

More broadly, evaluation directors characterized themselves as liaisons between project officers and governing boards. In this fashion, some exercise important functions in the foundation's strategic planning. One interview respondent commented, for example, that her role emphasized "building an evaluation system to pull together all the pieces, to connect with the strategy level." Another noted that his responsibilities included reporting concerns arising from evaluation to the board. Another highlighted the importance of articulating themes and evaluation findings across distinct interventions.

On a more fundamental level, evaluation directors characterized their role as one of advocacy for evaluation. As one commented, "(I help) establish and maintain a culture that is thirsty for evaluation." Others commented that their "mission" if not their formal role included integrating measurement and evaluation with grantmaking and a foundation's strategic direction.

Comments by informants suggested that advocacy for evaluation may put evaluation personnel at odds with others. Program and evaluation specialists may have different perspectives associated with distinct organizational subcultures. The perspective of the former president of a leading US foundation illustrates this diversity of outlook:

> Program staff saw (evaluation) as a distraction. They are action people and so are grantees – sure that what they're working on is going to be a good thing. … From a program director's point of view they don't want to be doing a lot of assessment except for the most basic – they want to be out having training sessions and helping people do what they do better. From this perspective, the whole idea of evaluation in a formal sense is disruptive.
>
> From another angle, program directors are not interested in having results that show their investment in staff and money was not effective. They aren't scientists -- they know something will work by gut. (They are) almost hostile (to evaluation). (In my experience) program directors were passionate and didn't think what they did could or needed to be measured.

Potential hostility to evaluation was not merely a subjective matter. In at least some foundations, unfavorable evaluation can put program officers at actual risk. Program officers often help initiate an intervention with the governing board and select specific grantees. The degree of success achieved by these programs, as measured by evaluation, affects their career prospects. "Projects that fail tend to have staff turnover," commented one informant.

Improving Evaluation I: Recommendations for the Evaluation Field

The interview schedule included an item asking informants how evaluation "could be made more valuable to a foundation such as yours." Responses to this item appear most directly relevant to outside consultants but may be applied to internal evaluation personnel as well. The recommendations received address some of the challenges to evaluation identified above and reflect the stresses often associated with the evaluation director's role. Recommendations were expressed in three interrelated areas.

Deliver Evaluation Results in a Timely Manner

When informants were asked about how evaluation could be made more valuable to foundations, more prompt delivery of findings was their most frequent answer. According to several, evaluation findings are most valuable early or at the midpoint of a program. During an implementation period, evaluation may pinpoint problems in launching an initiative. Timely evaluation may identify grantees that have encountered problems but may achieve their objectives with the aid of additional resources such as technical assistance.

Evaluation may be particularly valuable at the midpoint of an initiative. Often, foundations seek to assemble a second cohort of grantees at this time. Input from evaluators may contribute to revision of an initial RFP regarding qualifications, resources, and expectations of potential, new grantees.

Informants, however, expressed the impression that evaluators tend to concentrate detailed, rigorous, outcome-oriented final reports. The fact that the initiative has often been completed by the time the final report appears greatly reduces the usefulness of the evaluation. As one informant commented, "evaluation comes too late for decision-making." According to another who reflected on significant foundation experience, "Some (reports) came too late in the day to advise us for anything concrete. We had to keep moving."

Another perspective on timeliness is captured by an informant's term, the *real time thought proces*s. According to this informant, "Evaluation takes place as a learning process." When this process takes place, "we don't need a final report because (evaluation) has been telling our story all along." Further, "in an iterative manner the real time process approach should improve an intervention as time goes on." Illustrating this principle, the informant recalled evaluation of an anti-violence initiative:

> We spent $6 million contracting with Rand, Stanford, and Johns Hopkins to get an outcome evaluation. Findings were negative. So we asked *why*? The contractors were unable to provide an answer. But we did learn from pieces of the evaluation. For example, community fellows who were funded lightly in the initiative accomplished a lot (as individuals).

Make Findings Clear and Definitive

Second most prominent were comments urging evaluators to produce findings that are clearer and less guarded than they receive at present. Informants remarked that evaluators often produce reports that were excessively technical. Several officials commented that the evaluation reports they receive are too jargon-laden and "academic" to be meaningful to board members and other stakeholders. Comments illustrating the importance of clear, readily-comprehensible findings included the following:

> … people who are most eager to find out what we're doing get through the technical complexity. But we'd like more accessibility.

> The most useful evaluation findings have been oriented toward who will be using the results and for what – a very different perspective from academicians ...

> (Evaluation findings should be) more digestible and delivered faster. After benchmark events, a bullet-point memo on findings should be delivered, raising issues and perhaps recommendations. (Evaluators need to write) reports that are not jargonistic. There must be scientific precision … but tell the story in a way that people can understand.

Several informants recognized potential challenges in reconciling scientific rigor and ready communication of results, but emphasized the importance of communicability. According to one, for example:

> (Evaluators should) be able to boil things down to elevator speech findings…If you want attention from anyone in the hierarchy you need to do this. Consultants are often technical but you need to distill things even to (the degree that they) sacrifice precision in service of communication.

Among informants, an appreciation for definitive findings was closely associated with the value they placed on accessible and concise communication. According to several, the qualifiers and disclaimers that characterize many evaluation reports reduce their comprehensibility and potential impact. An informant illustrated this impression thus: "It is hard for staff to read a very detailed evaluation with footnotes and caveats and to figure out what it means."

Work in Partnership

In addition to the product that evaluators may deliver, informants commented on the importance of the process involved. Evaluators who worked well with both foundation personnel and grantees were seen as providing the best value. Informants accorded the highest rank to evaluators who had the capacity not only to produce reports but to serve as full team members in promoting the initiative's success.

An ability to work with and advance the objectives of grantees received emphasis. Most fundamentally, the ability to carry out effective group process with grantees was considered important. Skill in group process was considered important for obtaining evaluation data or other information desired by foundation staff. Some informants emphasized the contributions to developing evaluation capacity among the grantees.

An ability to understand and address the interests and needs of foundation personnel was also characterized as important. Respondents commented that "The quality of this work depends on the ability of the evaluator to facilitate reflection and feed back results." The *real time thought process* concept introduced earlier is relevant here. Evaluators were thought to generate the greatest value when they helped foundations go through a self-reflective learning process aided by formative feedback consistent with the foundation's needs. To aid this process, an informant commented on the importance of an evaluator's making time to "actually talk about evaluation and not just rely on written reports."

Sympathy for the commitment of foundation personnel to an intervention was also considered a positive attribute. As one respondent commented,

> The wonderful thing about working in a foundation is that it employs smart people who really care about their work. Sometimes an evaluation consultant can get into nerdy place. Instead, they need to appreciate the passion that program managers have about their work. Evaluators need to support this passion –evaluation should be in service to the goals of the initiative – to help it achieve impact – this approach has a better chance of success than being a dispassionate third party. The conflict of interest issue (raised by diminution of the boundary between evaluator and client or

subject) isn't operative for this kind of change that foundations today seek. The evaluator has to be a part of the change effort by being part of the feedback loop.

Improving Evaluation II: Recommendations for Foundations

No interview item explicitly addressed the internal steps foundations might take to obtain better value from evaluation. However, recommendations of this nature may be inferred from comments made by informants on the objectives of and challenges to evaluation. Interrelated recommendations concern dialogue between evaluation and non-evaluation specialists, leadership consensus on evaluation, and safeguarding of the evaluation process. Some relatively simple, instrumental steps are also likely to promote more useful evaluations.

Institute Effective Dialogue

The fact that evaluation and non-evaluation specialists (including program officers) should have different outlooks is natural and potentially constructive. Quotations cited include contrasting impressions of the approaches of program officers and evaluators as, for example, "passionate" versus "nerdy." Effective dialogue between program and evaluation specialists can bring the perspectives of program and evaluation personnel closer together and enable the work of each to better contribute to the foundation's goals. Such dialogue should involve program officers, higher-level program staff, and governing board members. The fundamental objective of dialogue should be development and maintenance of partnership between evaluation specialists and others in a foundation.

Dialogue should enable evaluation specialists to educate program officers and board members about the potential contributions and limitations of evaluation. It should also provide evaluators with clear direction about the foundation's expectations and help maintain their commitment to fulfilling these expectations. It is instructive to recall an informant's complaint about an evaluation contractor's failure to collect student-level data in an urban education intervention. When confronted, the contractor responded "we (only) did what we were told." According to the informant, evaluation specialists often "intellectualize" rather than "grapple" with challenges in assessing impact. Dialogue should promote solutions to such challenges.

Develop Leadership Consensus

Ultimately, the usefulness of evaluation depends on its ability to meet the needs and expectations of foundation decision-makers. In the abstract, these needs and expectations would seem straightforward. Evaluation would appear to serve foundations best by providing systematic studies utilizing standardized criteria to assess efficacy. Informants in the present study generally implied that evaluation of this kind was crucial in advancing an understanding among decision-makers about the interventions that were most likely to work.

However, consensus around these principles does not appear to exist across foundations. The history, culture, and idiosyncratic characteristics of a foundation may result in different concepts and approaches to evaluation. From a living benefactor's perspective, for example, the purpose of evaluation may be confined to assurance that funds are being used in a manner consistent with his or her intent.

Consensus regarding purposes and methods of evaluation may not exist even within an individual foundation. It is useful to recall the comment of an informant cited earlier that one of his board members was experiment-oriented while another "just wanted stories." Evaluation personnel require a consensus among decision-makers about the methods and data that are considered valuable for their executive functions. Differences in expectations, whether explicit or hidden, mean that at least some decision-makers will receive findings that they consider of limited value. Lack of consensus may cause confusion and result in poor focus among evaluators.

Safeguard the Evaluation Process

Both comments of informants and theory in the social sciences[10] suggest that the structural settings in which evaluation specialists are placed may reduce the value of results they provide. Dependence by evaluators on program personnel for resources may motivate evaluators to produce more positive results than they otherwise might. This principle holds for both in-house evaluation personnel and outside contractors.

Foundations will obtain more useful results from evaluators if they ensure that evaluators are not punished for reporting negative findings, either by losing favor with operations personnel or, in the case of outside contractors, losing future business. Foundations might consider instituting practices for obtaining independent input on the quality of the evaluation they receive, alternative interpretation of evaluation findings, and overall success in attaining their goals. The Council of Economic Advisors has been cited as a model for "tapping knowledge in the social sciences" for advising top-level US policymakers.[12] Under this model, a small group of highly qualified individuals directly responsible to the foundation CEO might be established and encouraged to independently review data, conduct studies on its own initiative, and consult with other outside experts. Each member might receive a non-renewable term appointment. In this fashion, foundation leadership would obtain input relatively uncontaminated by organizational or economic pressure.

Appropriately Schedule and Specify Evaluation Tasks

Foundations can take instrumental actions in the areas of organization and contracting to help them obtain more useful contributions from evaluators. In the case of both internal evaluation personnel and outside contractors, foundations can benefit from bringing evaluators into the intervention planning process at an early stage. Early involvement of evaluators can help configure an intervention in ways that make it more readily evaluable, for example, by identifying outcomes that lend themselves to concrete measurement. Calls for proposals by potential grantees can then specify requirements

compatible with the foundation's decision-making needs and schedule. RFPs for outside evaluation contractors can also be formulated to reflect the value foundations place on timely results and formative feedback.

Discussion

The findings reported above should be viewed with the understanding that the study was not intended to obtain information that could be generalized across all US foundations or even among those of any given size or orientation. Rather, the study aimed at illustrating a range of activities, approaches, and issues potentially present in any foundation's efforts to obtain and make use of evaluation. Hence, conflict between program and evaluation personnel reported by some informants must not be thought of as necessarily prevalent. Rather, the potential for such conflict in any foundation must be recognized.

It is also important to acknowledge potential bias in the sample used here, composed as it was largely of foundation evaluation directors. A sample of program officers and other foundation personnel specialized in areas other than evaluation may have yielded less favorable opinions regarding the value of evaluation. However, comments on the features of an evaluation effort that affect its usefulness to a foundation appear less subject to biased reporting.

Variations among foundations on several important dimensions are notable. In areas that some foundations identified as problems, others reported outstanding practices and achievements. Although resource allocation for evaluation by some foundations appeared quite small, for example, evidence surfaced that some foundations have made evaluation an object of generous funding and significant attention. One informant suggested the existence of a sort of "evaluation elite," whose members use practices such as:

> … commissioning numerous, large-scale evaluations; making evaluation reports public, resulting in hundreds of thousands of downloads of these reports; using timely evaluation findings to strengthen foundation grantees' projects; and using evaluation results to contribute to a more informed public policy discourse.

Another informant reported that in one year her foundation expended 30 percent of its extramural budget on evaluation.

Similarly, although social relations within foundation may give rise to a risk of positive bias in interpreting evaluation findings, some informants reported practices indented to promote independence in the evaluation process. One informant, for example, reported that prior to an evaluation's going into the field, her foundation convened a scientific review panel including experts in both content and evaluation. As she described the process,

> … the investigator presents the proposal and the review panel comments on it. It is usually a lively dialogue that helps everyone avoid the pitfalls that you usually recognize after the study is completed.

Despite these examples, comments by informants interviewed in this study suggest that evaluation today does not contribute as much as it might to advancement of foundation goals. A number of practical steps have been recommended above for both foundations officials and evaluation professionals to help remedy this situation. However, informant opinions suggest a number of challenges that are likely to resist resolution for years to come.

First, evaluation as a science needs to progress farther than it has. Evaluation techniques now in use tend to borrow heavily from the social sciences and less often from the natural sciences. The rarity of clear, unqualified results highlighted by informants in this study suggests essential limitations in the currently-available methods.

Second, the progression of emphasis suggested by informants from evaluation of specific interventions or programs to assessment of the foundation as a whole represents a crucial development. Increased emphasis on assessment of the foundation's overall contributions to society is tied directly to the mandate for accountability. Emphasis on this mandate transforms evaluation from a sometimes *ad hoc* to a central function within foundations. Associated with this increased importance, however, are challenges such as the need to develop indicators that at the same time reflect achievements of multiple, individual programs, yet at the same time transmit simple, meaningful information to decision-makers.

Although it is hoped that rapid progress will take place in addressing these challenges, foundations need to make the most of the evaluation capabilities that are presently available. Under one option, foundations might increase concentration on objectives that are clearly measurable and feasible given the resources allocated for their achievement. Another approach might be to concentrate on initiatives that lend themselves to experimental or quasi-experimental evaluation, procedures better able to establish causal relationships than many others in widespread use. A third alternative, which is compatible with the two, would view each successive evaluation effort as a step in a learning process, adding to an information base about which interventions are most feasible and which, even if not clearly successful, to contribute to the foundation's overall objectives. Decision-makers adopting this alternative would need to take an eclectic approach to findings reported by evaluators and employ a degree of informed guesswork in moving forward.

References

1. Wallace N. 2008. "Timely evaluations still rare in the non-profit world." *Chronicle of Philanthropy*. 21(5): 23.
2. Lenkowsky L. 2007. "Evaluations can be dangerous." *Chronicle of Philanthropy*. 19(20): 35.
3. Bolman LG, Deal TE. 2008. *Reframing organizations: artistry, choice, and leadership*. San Francisco: Jossey-Bass.
4. Weiss CH, Buchvalas. 1977. "The challenge of social research to decision making." In: CH Weiss (ed.), *Using Social Research in Public Policy Making*. Lexington, MA: Lexington Books.
5. Lawrenz F, Gullickson A, Toal S. 2007. "Dissemination: handmaiden to evaluation use." *American Journal of Evaluation*. 28(3): 275-289.
6. Henry GT, Mark MM. 2003. "Beyond use: understanding evaluation's influence on attitudes and actions." *American Journal of Evaluation*. 24(3): 293-314.
7. Christie CA. 2007. "Reported influence of evaluation data on decision-makers' actions: an empirical examination." *American Journal of Evaluation*. 28(1): 8-25.
8. Seidler J. 1974. "On using informants: a technique for collecting quantitative data and controlling for measurement error in organizational analysis." *American Sociological Review*. 39: 816-831.
9. Kumer N, Stern LW, Anderson JC. 1993. "Conducting inter organizational research using key informants." *Academy of Management Journal*. 36(6): 1633-1651.
10. Levitt B. March JG. 1988. "Organizational learning." *Annual Review of Sociology*. 14: 319–338.
11. Greenwald HP. 2008. *Organizations: management without control*. Thousand Oaks, CA: Sage.
12. Wilensky HL. 1967. *Organizational intelligence: knowledge and policy in government and industry*. New York: Basic Books.