

A BUMPY JOURNEY TO EVALUATION CAPACITY: A CASE STUDY OF EVALUATION CAPACITY BUILDING IN A PRIVATE FOUNDATION

Boris Volkov
University of North Dakota
Grand Forks, ND

Abstract: Evaluation capacity building (ECB) processes were explored through a case study of the ECB efforts of a private foundation that had solid financial resources, intelligent staff, an internal evaluation officer, and access to prominent evaluation consultants. The overarching research question was: What are the factors that affect the ECB process in an organization? The research design included in-depth interviews, surveys, document analysis, and observations involving a sample of the foundation staff and external consultants. This article illuminates the complexity of the organization's experiences with building evaluation capacity and provides insights on what must be considered when developing a viable evaluation function in organizations.

Résumé : Les processus de renforcement des capacités d'évaluation (RCÉ) ont été explorés par une étude de cas sur les efforts effectués dans le cadre du RCÉ par une fondation privée, une organisation avec des ressources financières solides, un personnel intelligent, un agent d'évaluation interne, et l'accès à des experts-conseils reconnus en évaluation. La question primordiale de recherche était : Quels sont les facteurs qui affectent le processus de RCÉ dans une organisation? La recherche était conçue de manière à inclure des entrevues approfondies, des sondages, l'analyse documentaire, et des observations d'un échantillon du personnel de la fondation et des experts-conseils externes. Cet article révèle la complexité des expériences de l'organisation dans le renforcement des capacités d'évaluation et offre des commentaires sur ce qui doit être considéré en élaborant une fonction d'évaluation soutenable dans une organisation.

■■■■■■ Evaluation capacity building (ECB) has increasingly become a catchphrase in the field of program evaluation. Greater requirements of public accountability have produced the realization

that organizations must increase their own capability to continuously monitor and evaluate their projects, programs, and strategies. ECB has gained prominence internationally, especially within public sector organizations that operate educational, social, and health-related programs (Naccarella, Pirkis, Kohn, Morley, Burgess, & Blashki, 2007). Many profit and nonprofit organizations have launched their own evaluation departments or, at a minimum, have established full- or part-time positions for internal evaluators to help with evaluating and monitoring their programs. Reputable governmental institutions (e.g., the National Science Foundation and the Centers for Disease Control and Prevention) have also shown increasing interest in developing evaluation capacity in the organizations they sponsor and the communities with which they partner.

According to Boyle, Lemaire, and Rist (1999), “evaluation capacity refers to the human capital (skills, knowledge, experience, etc.) and financial/material resources, and evaluation practice to the actual ‘doing’ of evaluation” (p. 5). The obvious link between ECB practice and program evaluation is that the goal of ECB is to reinforce and sustain effective program evaluation practices. This goal can be achieved by increasing an organization’s capacity to design, implement, and manage effective evaluation projects; access, build, and use evaluative knowledge and skills; cultivate a spirit of continuous organizational learning, improvement, and accountability; and create awareness and support for program evaluation and self-evaluation as a performance improvement strategy in the internal and external environments in which it functions (King & Volkov, 2005). As argued by Taut (2007b), evaluation use for learning, participation, and ownership and evaluation capacity building are closely linked.

Haarich and del Castillo Hermosa (2004) argue that “[E]valuation capacity building, as an approach for the development of evaluation systems, is the integrated and planned development of skills, resources and infrastructures and the intentional shift towards an evaluation culture in an organization, department or government” (p. 2). Evaluation capacity building simultaneously addresses demands for accountability and for interactive participation. Cousins, Goh, Clark, and Lee (2004) state that, when integrated into the policies and ongoing activities within an organization, evaluation has the potential to transform itself into a learning system that cultivates shared values and understanding among organization members.

Although evaluation is increasingly becoming a vital part of organizational life, the evaluation field is still short of practical tools for its

effective institutionalization. Cousins et al. (2004) suggest that research in this area is underdeveloped. According to Stockdill, Baizerman, and Compton (2002), the evaluation profession as a whole lacks a well-developed theory and associated indicators for understanding evaluation capacity at an organizational level, particularly its inherent change over time and “ongoingness.” Despite its becoming more common, evaluation capacity building has not even been clearly defined, resulting in varying conceptualizations of evaluation capacity building (Compton & Baizerman, 2007; Naccarella et al., 2007; Taut, 2007a). To date, there is still only limited literature about evaluation capacity building in organizations, including a few published case studies of ECB efforts. Accordingly, not much literature exists on evaluation capacity building in the foundation world. It is certain, however, that evaluation is increasingly attracting attention in the philanthropic field—its goals, opportunities, uses, various approaches to implement it, and associated challenges (Slater, Constantine, & Braverman, 2004). Mark and Beery (2004) believe that foundations might substantially benefit from using evaluative thinking and evaluation-generated information throughout the entire planning and grant-making process. As stated by Kramer, Graves, Hirschhorn, and Fiske (2007), a number of philanthropic organizations are currently engaged in “a fundamental transition from a narrow view of evaluation limited to traditional third-party outcome studies to include more performance-centered approaches encompassing a wide range of activities that provide foundations and their grantees with current information and actionable insights” (p. 9). As a result of such an interest in evaluation, many foundations have established internal evaluation positions in addition to employing evaluation services of external consultants.

The purpose of this case study was to look in depth at the ways evaluation capacity building materializes itself in organizational practices. In meeting this purpose, the overarching research question was: What are the factors that affect the ECB process in an organization? The study presented an opportunity to describe and interpret ECB practices in a nonprofit organization and improve the ECB knowledge base for a wide range of stakeholders in organizations that “find themselves at various stages along a continuum of readiness for evaluation capacity building” (Cohen, 2006, p. 92). The lessons learned from the case may help such organizations to be more successful and avoid some pitfalls in launching and maintaining the work to increase their long-term capacity to conduct and use program evaluations.

METHODS

The study used a qualitative research design, specifically an interpretive case study approach. Data collection involved analysis of documents and secondary data, structured interviews (with organization staff and management, as well as external and internal evaluation experts), and participant observation. Three criteria were used to identify the case: (a) the organization had to be a not-for-profit organization (a less-researched type of organizations in terms of evaluation capacity); (b) there had to be an ongoing, routine program evaluation function in the organization (which would reflect the presence of at least some evaluation capacity in the organization and would also allow observation of the role of this structural component in the overall ECB process); and (c) one or more high-level administrators in the organization had to be committed to institutionalizing or broadening the evaluation function (as a critical element of an ECB-friendly internal organizational context, according to King & Volkov [2005] and Volkov & King [2007]).

The site selected for this study was a nonprofit organization (which further will be called “the Foundation”), a private foundation in a large city in the US Midwest (a geographic factor facilitating the Midwest-based author’s physical access to the organization and its members). The Foundation met the three criteria and was an especially interesting place to study evaluation capacity building processes. A case study can be selected for its very uniqueness, for what it can reveal about a phenomenon, and for knowledge we would not otherwise have access to (Merriam, 1998). The Foundation was an example of such an atypical case. Not usually mandated to formally appraise their activities, philanthropic foundations—unlike many other organizations—“are free to conduct evaluation and to demand it as an activity by their grantees *only* if they see the value in it—or, to state it more simply, only if they want to” (Slater, Constantine, & Braverman, 2004, p. xlvi). Whereas many nonprofits struggled to fund external evaluations or do their own internal evaluation studies, the Foundation had enough money to hire prestigious evaluation firms to do large-scale evaluations whenever its leadership and staff believed it would be of value.

ECB work in the Foundation was initiated as a stand-alone process, as opposed to ECB being a part of ongoing evaluation procedures; the focus was on the linkage among evaluation, learning, and knowledge. When the study began, there was a Research and

Evaluation (R&E) unit in the organization (including a Research and Evaluation Lead and a contracted part-time evaluation consultant). There were three different divisions/large-scale programs in the Foundation, each of which had different lines of work and its own team of personnel.

The author negotiated access to the organization and collected data between June and December 2006. Unrestricted access was provided to proceedings, reports, and other relevant documents generated during evaluation projects and ECB efforts, as well as to the steering group and other ad hoc meetings for various lines of evaluation work in the organization. The author used in-depth interviewing with 22 ECB participants: organization staff, management, and external consultants. Following the “interview guide approach” (Patton, 1990), a standard interview protocol was used, consisting of open-ended questions with the focus on the participants’ experiences with and perceptions of ECB in their organization. The key interview topics included factors and conditions leading to and supporting the ECB process; obstacles to the ECB process; outcomes of ECB (both organizationally and personally); interviewees’ concerns, attitudes, and needs while being involved in ECB efforts; and factors that could make the Foundation’s ECB sustainable. The Foundation documentation also provided valuable data regarding the profile and history of the Foundation; strategic priorities; planning and implementing ECB; plans for and uses of external evaluation; and the nature of the staff involvement in evaluation and ECB work.

As a brief review of the strengths and weaknesses of the research, it ought to be acknowledged that, as with all exploratory research, the findings of this study are tentative. Obviously, using a qualitative case study design does not allow us to generalize the findings of this study to other contexts (which was not the actual purpose of the study). That being said, this study provides in-depth information about the real-life organizational dynamics played out during the process of developing evaluation capacity. The Foundation’s “atypical case” status also makes it interesting for analysis (some readers may rightfully argue, though, that the fluid contexts of contemporary organizations make the line between “typical” and “atypical” rather blurred when describing organizational processes). Another strength to mention is the data source and methodological triangulation that helped the author minimize the chances that the research findings would be based on idiosyncratic data.

CASE DESCRIPTION

The Organizational Context

The Foundation operated as a traditional grant-maker for nearly 60 years, making about 120 grants in 39 categories. In 1998, the Foundation's board of directors and senior management made a dramatic shift by deciding to focus their mission on significant changes in a small number of communities, which presumably would allow the foundation's financial and human assets to achieve greater impact. Through that innovative approach, the Foundation started working directly with communities by providing financial resources and technical assistance to help them with community revitalization. The organization was committed to helping communities via identifying, sharing, and advocating strategies and tools with lasting impact. Communities worked inclusively to develop their vision and strategies. The Foundation formed partnerships with eligible communities, that is, those in need of revitalization, to work intensively for up to 10 years to reach their goals and encourage them to identify what they need and how they want to accomplish their goals. The Foundation's clients, or partners, came from rural, urban, American Indian, and rural Latino communities in a number of states.

To achieve its mission, the Foundation administered a few distinctive programs. The program theories stated that if provided with an opportunity to build people's skills, connections, and knowledge base, communities would take the opportunity to design, lead, and implement sustainable strategies that would revitalize communities. Through one program, the Foundation formed 10-year partnerships with up to 12 communities from different states. The Foundation provided technical assistance and financial resources to these communities so that each could develop and implement sustainable strategies. At the time of this research, the Foundation partnered with 10 diverse communities. Another program worked to research, develop, or find the products communities need to advance their initiatives. The program linked communities with existing tools, strategies, or models, and also collaborated with other groups to create new products. The third program helped small rural communities strengthen their leadership systems. The program offered comprehensive community leadership programs and activities within communities that, in general, experienced significant decline and need in population, income, and resources.

Evaluation and ECB Activities at the Foundation

Launching ECB

Up to the year 2002, there was little consistent evaluation work at the Foundation. When closing out grants, the foundation officers carried out some kind of assessment and reported their findings in a final report on a case-to-case basis. Later, when the Foundation switched from its earlier, traditional grant-making operation to a new mission of working with communities, no evaluation targeted that new line of work. At some point in time, an external evaluation consultant developed an intricate program logic model in consultation with the Foundation president/CEO, but the document was eventually put away and forgotten, perhaps due to its overwhelming complexity and the staff's inability or reluctance to utilize it.

As the Foundation began work with the communities, community staff had to come up with plans for revitalizing their community. The foundation staff expected communities to include evaluation as part of their strategic plans to ensure they would learn from their work. The expectation was that the evaluation would be tied to the strategies that were in a plan, and, as people learned from the work, it would change the way they did the work. In theory, evaluation of the community work was a strong component of the Foundation's efforts; in practice, however, evaluation was always hard to do. It was difficult for the communities to implement the strategies they had designed in their plans. Once a plan was approved by the Foundation and people began to implement it, they often found that they had to alter their strategy somewhat. Evaluation was always onerous, for several reasons. First, it required a lot of work. Second, it was not sufficiently clear what kind of evaluation report the Foundation expected. Third, while there were some efforts to strengthen evaluation and help those communities document their work, some Foundation staff reportedly resisted evaluation by saying that there were no appropriate positions in communities and that it was not possible to expect them to do that much work. These serious challenges contributed to the evaluation work's lagging well behind the community development activities.

Another issue to be highlighted is the fact that the Foundation explicitly hoped to become a learning organization. There was a shared understanding that because they were doing some things different than other foundations, the ability to learn from what was being done—the

ability directly linked to evaluation—was important. Around 2001, the management started thinking and talking more about learning questions that the foundation staff would use for continuous improvement, what they were learning about the work in communities, and how they needed to use that knowledge. There was some effort to formalize that process. The management developed several learning schemes for how the foundation could learn from its work and feed that back in for decision-making, among other things. Those plans were presented to the Board and to staff, with days of in-service dedicated to them, but reportedly not much from the plans was ultimately implemented. Apparently, despite the initial good intentions, there was not sufficient motivation in the leadership and staff to make the goal of becoming a learning organization a key priority.

Over time, a group of staff began to advocate for establishing a position at the foundation to guide the research work or at least to set some quality standards. Also, there was more pressure from the Board because of the foundation's novel, yet-to-be-proven approach and the level of money involved. The Board wanted evaluation of the foundation's programs. These factors played a role in establishing a Research and Evaluation (R&E) Lead (or an Evaluation Lead) position in 2003. The R&E Lead position included a set of research responsibilities. There was also an expectation that the foundation would start evaluations of its three programs. The set-up provided for an evaluation liaison who worked with the Evaluation Lead in each of those programs. The evaluation liaison in a program was the decision maker, whereas the R&E Lead was an internal evaluation consultant providing technical assistance. However, he was also responsible for the quality and integrity of evaluation, which served the overall purposes of the foundation.

The ECB started in an explicit way in 2003 when the R&E Lead began work on program evaluations with staff. One of the first activities that he initiated was to help the program staff agree on a theory of change, or a program theory. They laid it out in a comprehensive flow chart, detailing how the program was going to be implemented, what outcomes they were trying to achieve, and how they would evaluate them. In early 2004, the R&E Lead significantly expanded the foundation's evaluation agenda by engaging the management team to determine priorities for evaluation, contracting with outside evaluation consultants, developing electronic data-handling capacity, building staff capacity for evaluation, and building community capacity for evaluation.

Program-level evaluation

In 2005, the Foundation embarked on evaluations to improve the programs and provide the Board with information needed to make mission-related decisions. Each of the programs was evaluated by external evaluators from both major national and smaller local evaluation firms. The evaluation experience was not similar for all programs. For some, it was a stressful process, but for others empowering. The program managers were involved in providing all the information that was needed to conduct evaluation, and they also participated in making program changes as a result of evaluation. They learned that they had to stay very actively engaged in the evaluation process, deciding at what points they were going to hire external experts to do some pieces of the evaluation work and what activities they should keep internal.

At the beginning of 2006, one program's staff started working on a cross-community evaluation. To assist all the program teams in developing their own evaluation systems, they started developing an evaluation toolkit/guide, a handbook that all of the teams could consult for steps to take in developing an evaluation system. The staff, none of whom were trained evaluators, tried to do it on their own; however, they later realized the necessity of hiring external evaluation consultants to effectively implement that work. Hence, both the program team and external experts were involved with developing the evaluation guide, as well as with facilitating evaluation training sessions to orient the foundation's community partners toward the value and use of evaluation and continuous stakeholder participation.

The Internal Learning System and cross-community data collection

In 2004, the R&E Lead came up with the idea of creating a process, or a system, designed to document evidence for internal learning in the foundation. This was called the Internal Learning System (ILS). Special software developed and maintained by the IT department offered an option to link learning and evidence to the program's logic model. The three program logic models were adapted so that each distinct element of the program (assumptions, actions, intermediate outcomes, impact, etc.) would have a separate box and therefore could be linked to associated documents.

After more than a year of using the electronic linking system, it became increasingly evident that the system was cumbersome and

not well structured for showing so-called “value-added capacities” or community-level outcomes. Staff descriptions of activities or achievements were not adequate for useful analysis and, more importantly, synthesis. The entire process required an internal research effort too laborious to make employing the documentation meaningful and useful. Gradually, the Internal Learning System went out of use. It was clear that something different was needed to replace it.

In 2005, just before his position was terminated (quite unexpectedly, at least for the majority of the foundation staff), the R&E Lead developed and proposed a comprehensive plan for cross-community data collection—a long-term study of revitalization in the communities involved in its largest program. The goal of the study was to gather and assess evidence of community development as it was understood and pursued by the Foundation and the involved communities. After the R&E Lead had left the foundation, the plan was discussed a few times by the senior management, program leads, and outside evaluation and community work experts. People realized that it was not an easy task to come up with proper and accurate indicators. Some staff believed that the indicators would not be relevant for the other programs. No one on the staff had time, sufficient expertise, or the motivation to take charge of the study. The initiative for launching the cross-community data collection slowly died.

Utilizing external evaluation expertise

Besides hiring external evaluation firms to formally evaluate their programs, the foundation also asked for advice from a number of evaluation consultants. These consultants provided input on the program logic models and evaluation potential, conducted several evaluation trainings for the staff, consulted with each of the program areas, and evaluated some projects. Perhaps the most remarkable use of evaluation expertise was the Evaluation Advisory Circle, a group of nationally recognized evaluation specialists who were brought in to review and advise on the emerging evaluation of the foundation programs. Altogether, there were three two-day sessions on an annual basis. The Circle meetings involved 8 to 10 evaluation and philanthropy experts from around the US, as well as the active participation of a number of Foundation staff members invited to attend any or all discussions of the Circle. The staff took an active role in orienting the Circle to their program evaluations and answering questions from the Advisory Circle members. In this regard, according to one outside evaluation practitioner, the foundation was “certainly to be commended for openness to continuous learning and seeking input

and feedback on what they are doing.” Unfortunately, the results of those meetings were not fully utilized and integrated in the evaluation work at the foundation.

Staff development in evaluation

Four program people in the three programs worked directly with evaluation as program evaluation liaisons. Anyone who wanted to go to conferences and learn more about evaluation was encouraged to do so. For instance, the Foundation funded one evaluation liaison to take several courses toward an evaluation certificate at a local university (unfortunately, she left the foundation for another position shortly after completing it). A number of the staff members went through different sorts of evaluation training and workshops.

Consulting outside evaluation experts was not helpful merely for getting feedback on evaluation, but it was also valuable capacity building for the staff. Their talks about the role and place of evaluation at the foundation helped to move understanding of and personal capacity in evaluation to a new level. Monthly “learning review” meetings were used as opportunities to share information on evaluation and to involve people in looking at their work evaluatively. One staff member reported that he was “excited that we are stepping up to the plate and making evaluation a priority now and really focusing on a clear, unifying message of what we mean when we say ‘evaluation’ and ‘ECB’ and the expectations that go along with that.”

Unfortunately, despite that excitement, by summer 2007 the brain drain in terms of the evaluation expertise combined with organizational memory reached a critical point. In January 2006, the Evaluation Lead’s job was terminated. Reportedly, at least two of the reasons were the insufficient rapport between the Evaluation Lead and the Foundation leadership and program management, as well as complex politics around program evaluation. There were no plans in sight to hire new evaluation staff at that point in time. Shortly afterwards, one program’s Evaluation Liaison left for another job. The Director of Community Activities and Learning (whose job responsibility, among many others, was to develop and implement evaluation and learning systems for community activities) resigned by the end of 2006. A community liaison, who participated in evaluation trainings and ECB work in communities, left the foundation. A member of the Evaluation Development Team (who worked on building evaluation capacity in the communities) found out about his termination in March 2007. Two program Leads involved with the evaluations of

their programs quit at the beginning of 2007. Finally, in the summer of 2007, the president of the foundation himself resigned. The foundation's existing internal evaluation capacity was in jeopardy, due to the leadership's lack of commitment to meaningfully integrating evaluation into the organizational life, the rather complicated internal politics around program evaluation issues, and the loss of the extant evaluation expertise because of the increased staff turnover through resignations and terminations.

A Bumpy Journey to Evaluation Capacity

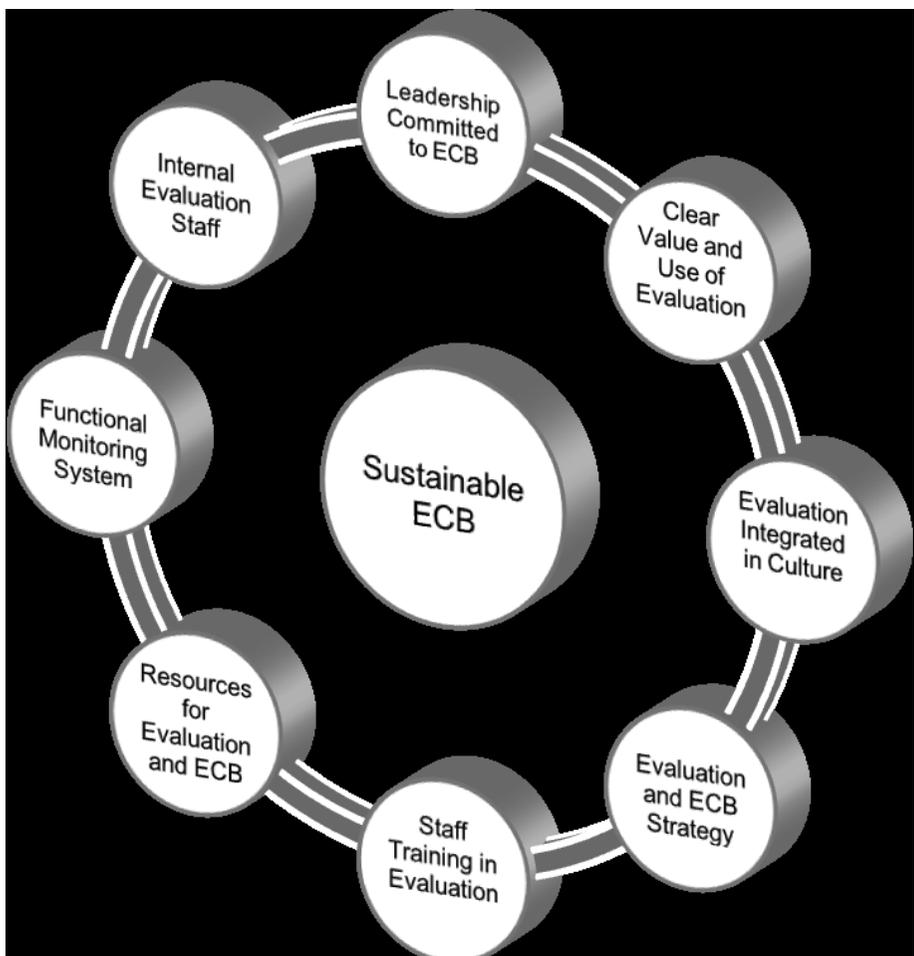
The Foundation's efforts to develop internal evaluation capacity comprised a series of ups and downs. As the Foundation took a new mission and an innovative philanthropic approach of working directly with communities, evaluation was supposed to acquire an important role in its operations. After a rather long latent period in starting internal evaluation work, there were a number of developments in terms of building the evaluation function in the organization. The case study revealed several factors related to the process and outcomes of ECB at the foundation. The most compelling factors can be metaphorically presented as the elements of an ECB "wheel" that can properly spin and progressively gain momentum only when those embedded factors (or "driving forces") are in place and act in unison (see Figure 1). This graphic representation emerges from the research data; is supported by a literature analysis; and integrates important elements of the three core categories of ECB, namely, organizational context, ECB structures, and resources for ECB (King & Volkov, 2005; Volkov & King, 2007), which will be discussed later in this article.

In terms of the *leadership committed to ECB*, the study participants wanted to see the foundation leaders truly committed to an ongoing evaluation function. Otherwise, either the ECB process would die because it would not have adequate resources, or the evaluators, evaluation results, and related information would be separated from program decision-making. The management's concern was certainly one of the factors prompting ECB at the Foundation, possibly originating from Board pressure concerning whether they made the right decision in changing the direction of the foundation. Some questions from the president and the program managers included "What are we doing?" and "Is it effective?" As one respondent commented, "You can only ask these questions for so long without creating a mechanism to find that out." The role of the senior leadership of the foundation, including its president, should be paramount in the ECB work, as highlighted in the following excerpt:

If we don't put evaluation as the highest priority that the President is concerned about all the time, then it's not going to be a concern to other people in your organization. So the emphasis on evaluation and concern for it needs to be modeled right at the top.

Another driving force of evaluation capacity building is *clear value and use of evaluation*. A number of the study participants believed that ECB would be sustainable when they could see the good value

Figure 1
The Evaluation Capacity Building Wheel and Driving Forces



of evaluation and the difference it would make in their work and organizational life. When evaluation was used to do continuous improvement and to begin to identify areas of learning that could be used and disseminated, it would sustain itself. After all, the foundation reportedly strove to be a learning organization, hence there was an awareness that evaluation should contribute to learning. One of the program officers explained that

one of the major principles or values for the foundation is the learning to do things better, and it is the model that we are entrenched in. I think that is one of the things that really drives evaluation.

As a matter of fact, the foundation evaluation activities led to the improved learning process and related program changes. For instance, Foundation staff were more thoroughly documenting their work and lessons learned. Some program staff became heavy users of evaluation, using its results to significantly redesign their programs.

Many study participants also talked about the importance of getting *evaluation integrated into the organizational culture*—into the work plans, the staff conversations, and their entire way of working. They lamented over the lack of buy-in, the fear of evaluation, and the fact that individuals in the foundation resisted if they could not find a comfort level with the evaluation process and a way to internalize it into their daily work. It is well known that distrust of the evaluation process is common in organizations (King, 2002). When people do not see it being used as a learning mechanism and instead think it is a control mechanism, their willingness to make the necessary investment of time and other resources may be seriously compromised. There has to be some explicit value for evidence-driven decision-making, for inquiry and trying to figure out the right questions to ask, and for setting up systems and processes for pursuing those questions.

For a number of interviewees, it was essential to ensure that people took responsibility for getting the data for the evaluation, that they had tangible incentives to do that, so it was part of their performance management agreement. Some study participants would expect evaluation skills and familiarity to be reflected in the job position postings for new hires and to become a part of the screening process for new staff.

Nonexistence of an organization-wide *evaluation and ECB strategy* was an issue for almost half of all interviewees. It was an arduous

task to build capacity throughout the foundation without an evaluation philosophy for the organization, “an overarching operative capacity building approach to evaluation that comes from the top down or from the top and diffuses around,” according to one interviewee. As another foundation staff member lamented:

Not a shared vision for evaluation. Not a shared agreement of what success looks like for the foundation. No centralized evaluation function that kind of watches over all the programmatic evaluations and gives us guidance. And evaluation is now driven by panic ... as opposed to being planned and being strategic.

The study participants supported the idea of creating a unified evaluation/ECB strategy within the organization. The leadership and staff of the foundation, in close collaboration with both internal and external evaluation experts, would have to determine an overarching learning and research agenda connected to the overall mission of the foundation. The next step for the foundation would be to figure out routine data collection and analysis systems, capacity that did not exist in any mature way at the foundation.

Also emphasized was the importance of ongoing *staff training in evaluation* theory and methodology. If staff had training in evaluation approaches so that they understood what major evaluation questions were and the kinds of data needed to address those questions, they could incorporate that knowledge into better program development and implementation.

More than one third of the study participants did notice some more advanced understanding of evaluation as well as increase in evaluation and research skills in the staff. Community activities and learning reviews and the evaluation consultants' meetings with people raised the level of awareness of evaluation and its role in the foundation's work. Some people were intensely studying and directly learning about program evaluation, whereas others just picked up some evaluation knowledge because it was in the environment. Still, much more was required to further strengthen the staff's individual capacity to think evaluatively.

Obviously, having the sufficient *resources for evaluation and ECB* (e.g., to provide evaluation training and bring in consultants) was a positive factor. There was strong support from the senior leadership, who allocated significant funds for evaluation. The foundation did a

thorough job of selecting evaluators for each of its programs and selecting some independent leaders in the evaluation field who advised them on different areas around program evaluation. However, there was some controversy around having abundant financial resources. On the one hand, access to evaluation resources such as external evaluators was undeniably helpful. Influence of external evaluation experts constituted a factor of support for ECB. On the other hand, heavy reliance on external evaluators was detrimental to building evaluation capacity. As it was decided to commission external evaluations for each of the program areas, one evaluation consultant commented that it sent the message to the regular program staff that they did not have to attend to evaluation because somebody else had that responsibility. As a result, they lost all incentive and momentum to build the internal capacity to truly engage in evaluation. Instead, it became a contract activity:

We've always had a luxury of "why don't we just go hire a consultant, and she will come and do it for us." It's an easy way, but it is also a detached way.

Another factor was related to setting up a *functional monitoring system* that would ensure quality and feedback for people to do their jobs effectively and provide something close to real-time information about successes, challenges, and other sorts of interim milestones. It should be a practical evaluation system linked to the organizational mission and theory of change, something that would be translated relatively easily and that would be flexible and useful both for the foundation and the communities it supported. The foundation was struggling to design and utilize such a system.

Finally, the presence of the *internal evaluation staff* was necessary. The Evaluation Lead, through practicing and promoting evaluation, strove to make it a greater priority at the foundation, and a number of staff members were genuinely frustrated by the sudden termination of the Evaluation Lead position. Consequently, one of the barriers to ECB in the organization was the lack of staff with evaluation responsibility who would understand how to integrate evaluation into programs and to carry through on staff training activities and staff development in a supervisory way that would reinforce the importance of evaluation. The foundation needed a dedicated staff member who would possess strong evaluation skills, as well as a deep understanding of grant-making and program administration, to strengthen and promote the evaluation function in the organization and its partners.

DISCUSSION

For the purposes of comparative analysis, I will look at the above factors through the lens of the grounded framework and checklist for evaluation capacity building, featuring three core categories of ECB: organizational context, ECB structures, and resources for ECB (King & Volkov, 2005; Volkov & King, 2007). First, I examine the foundation's ECB as it relates to its organizational context (i.e., the internal and external organizational context, power hierarchies, administrative culture, and decision-making processes). In terms of the internal organizational context of ECB at the foundation, the leadership of the organization did not adequately support and share responsibility for ECB. There were evaluation champions in the organization, but they had quite different ideas about the purposes and directions for evaluation work and never came to a mutual agreement.

The organization's interest in and demand for evaluation information did increase. It transpired mostly because of the pressure from the Board and some program managers concerned that the lack of information impeded accountability and improvement of their programs. The internal political environment was ambiguous about introducing or supporting change. For example, evaluation-related changes took a rather long time to be initiated and implemented. Due to the lack of evaluation and monitoring information available, there were limited opportunities for sufficient input in decision-making or ensuring that people in the organization were able to use data to make decisions. There were some opportunities for socializing around evaluation activities (for example, working with outside evaluators collaboratively, as well as discussing evaluation findings during the community activities and learning reviews and team meetings), but there was no consistent integration of evaluation into ongoing staff activities.

The organization's external environment was in part presented by Board mandates and accountability requirements. The Board expectations were actually a key factor that prompted the ECB efforts at the foundation. Unfortunately, it was challenging to receive support for evaluation work from stakeholders that partnered with the Foundation. The sense of growing responsibility, external expectations, and related concerns have increasingly become a critical issue for grant makers "who are sandwiched between nonprofit groups that can't deliver what they promised (or can't find convincing ways to talk about their work if they have) and a powerful foundation board that wants solid evidence of big results" (Callahan, 2007, ¶ 24).

Second, I discuss some elements of the case related to ECB structures, that is, the mechanisms within organizations that facilitate the development of evaluation capacity. Notwithstanding some attempts by the R&E Lead, a purposeful long-term ECB plan for the organization was never developed. There was no appropriate conception of evaluation for the organization tied to the Foundation's mission. Accordingly, a strategy for conducting evaluations that would apply existing evaluation frameworks, guidelines, and professional standards was absent. The evaluation process did not have any prominent place in the major organizational policies and procedures. Although there were evaluation teams or task forces by programs, the Foundation never established a diverse and capable evaluation oversight group (composed of members of the staff, board of directors, and community) to initiate, assess, and advance evaluation processes continually in the organization.

Regarding the infrastructure to support specific components of the evaluation process and communication systems, there was some work at the foundation to build individual ability and readiness to implement evaluation activities. Some staff members attended evaluation workshops and trainings. Meetings with external evaluation experts also contributed to expanding the staff's evaluation knowledge. Again, that work was not sufficiently consistent and focused. Another component of the ECB structures—availability of an internal reporting/tracking system—was on the foundation's radar screen for a few years. Regrettably, the foundation's monitoring system was not adequately developed and properly utilized and, as a result, remained irrelevant to the needs of the staff.

The purposeful socialization into the organization's evaluation process was mostly unsuccessful. While some limited professional development and coaching in evaluation was provided, no clear expectations for people's evaluation roles were established, no tangible incentives for participation in the evaluation process were offered, and no "learning evaluation by doing it" was effectively promoted.

Finally, the foundation's evaluation capacity building efforts should be illuminated as regards ECB resources (availability and use of evaluation resources). Considering staff access to evaluation resources, the internal evaluator tried to teach evaluation processes by engaging staff in evaluation-related activities. External consultants were invited and discussed the various aspects of evaluation with staff members. Nevertheless, there was little rapport between the internal evaluation personnel and the foundation leaders and program

managers. Following the termination of the R&E staff position, no new evaluation staff positions were planned to be hired. It should be acknowledged, however, that access to external evaluation expertise still remained, through contracting with outside consultants.

On the one hand, there seemed to be generous fiscal support from the administration for program evaluation (e.g., the expensive contracts with elite evaluation shops). Obviously, there was solid organizational capacity in terms of basic resources for the evaluation process (copying, equipment for data collection and analysis, computers, software, etc.). On the other hand, the leadership did not have plans to provide long-term, explicit, dedicated funding for evaluation capacity building activities. Similarly, adequate staff time was not allocated to collaborate on evaluation-related activities.

To summarize, the ECB wheel at the Foundation was evidently out of alignment, with some driving forces not existing at all and others not compelling enough to make a tangible difference. As a result, the organization has had a rather long and bumpy ride to the intended destination of sustainable evaluation capacity—which is yet to be reached.

CONCLUSION

Some caveats must be suggested before making any conclusions. The context of a well-funded foundation is dramatically different from other organizational contexts. The very concept of ECB may be liberally operationalized in various organizational milieus, especially since there is still no consensus among researchers or practitioners on an operational definition of ECB. Certainly, the evaluation capacity building in another organization may be implemented in a different way than at the Foundation. These potential differences suggest that the results of this study may not be applicable to other settings. Thus, caution is needed in extrapolating the results of this study to other contexts, or in drawing general conclusions about the ECB-related factors and relationships.

While the insights drawn from this research may not be applicable to some other organizations, they can help researchers and practitioners better understand the nature of evaluation capacity development, as well as inform future studies and work in this field. The results of this study contribute to the body of ECB research by focusing on ECB process factors.

This case may serve as a cautionary tale; however, it is certainly not my intent to dampen the spirits of the organizations striving to institutionalize evaluation. The lessons learned and insights from this case study can be used to inform such organizations about important issues when incorporating evaluation into the organizational culture, which is “a challenge of enormous proportions, but one ... that is undeniably worthy of figuring out” (Cousins et al., 2004, p. 134).

One key lesson to emphasize is that even in an organization with good intentions and seemingly favourable conditions for building evaluation capacity (solid financial resources, intelligent and hard-working staff, an internal evaluation officer, and access to prominent evaluation consultants), the success of ECB is neither certain nor guaranteed. The case documented serious deficiencies in a number of core ECB areas in the Foundation. Notwithstanding a number of positive developments, there were telling signs that, for the most part, the evaluation capacity building in this organization failed. Regrettably, the leadership of the Foundation did not adequately support and share responsibility for ECB. A purposeful long-term ECB plan for the organization was never developed, and there was no appropriate conception of evaluation for the organization tied to its mission.

Another lesson suggests the importance of the following ECB factors—or driving forces—identified in this case: leadership committed to ECB, the clear value and use of evaluation, evaluation integrated into the organizational culture, an evaluation/ECB strategy, staff development in evaluation, resources for evaluation and ECB, a functional monitoring system, and internal evaluation staff. More specifically, a leadership’s commitment to the ECB process and an organization-wide evaluation and ECB strategy are the variables that need more attention in the ECB professional literature. Indeed, these issues were of serious concern to a large number of the Foundation staff members and external experts.

New for literature on evaluation capacity building is the discussion of both the negative and positive sides of having abundant financial resources in the organization. A related lesson from this case study is that both sustainable ECB and evaluative thinking are among those stubborn things that money just cannot buy. Most certainly, money is needed to support evaluation capacity building activities. However, an abundance of money may actually be detrimental to ECB when the organization is able and willing to hire an expensive expert or firm at any time while neglecting its own evaluation capacity devel-

opment. Further research is warranted on the effects of the financial status of an organization on developing internal evaluation capacity.

More effort should be made to further explore and develop the concept of ECB. The need to assess the value of different definitions of ECB (Compton & Baizerman, 2007) becomes of greater concern for us to be able to build consensus “around its socially constructed meaning in an iterative fashion” (Taut, 2007a, p. 120). I encourage both ECB practitioners and internal evaluators to join this discussion by publishing their understanding and experiences with ECB.

I would suggest conducting a number of comprehensive case studies in various organizations to directly observe their ECB strategies, processes, and outcomes (one possibility could be “action research” studies within organizations interested in ECB). Long-term case studies of organizational ECB can be especially beneficial to learn about how and why specific ECB strategies are working, and which components of the ECB process are being affected by which internal and external pressures. Such cases would present “interesting opportunities for observing the diffusion and localization of evaluation capacity in one or more organizations, and evolution of this evaluation capacity toward an optimal level” (Kuzmin, 2004, p. 183).

Using a qualitative design precluded me from being able to generalize the findings of this study to other contexts. In addition to qualitative studies that investigate the relationship between organization and program attributes and ECB implementation, it may be wise to think about cross-organizational quantitative studies of the numerous factors associated with ECB efforts—both successful and failed—to produce results that would allow some generalization. The true value of such research is in providing organizations, ECB practitioners, and evaluators with a better and realistic understanding of organizational arrangements that sustain or impede the ECB process.

In closing, I want to note that, although this study seems, on the face of it, to provide only descriptive information about the ECB efforts of one organization, the study findings ultimately raise larger issues about the implications of the increased accountability requirements in our society. As more and more organizations institute accountability and evaluation policies, we need to think hard about the related needs of their staff and partners. The transfer of ECB to partner communities and organizations through their multi-dimensional engagement in evaluation may constitute another fascinating area for research.

REFERENCES

- Boyle, R., Lemaire, D., & Rist, R. (1999). Introduction: Building evaluation capacity. In R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity* (pp. 1–19). New Brunswick, NJ: Transaction.
- Callahan, S. (2007, April 5). How to end the evaluation tug of war. *Chronicle of Philanthropy*. Retrieved from <http://philanthropy.com/free/articles/v19/i12/12004501.htm>.
- Cohen, C. (2006). Evaluation Learning Circles: A sole proprietor's evaluation capacity-building strategy. *New Directions for Evaluation*, 111, 85–93.
- Compton, D., & Baizerman, M. (2007). Defining evaluation capacity building. *American Journal of Evaluation*, 22(1), 118–119.
- Cousins, J. B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19(2), 99–141.
- Haarich, S., & del Castillo Hermosa, J. (2004, August). *Development of evaluation systems: Evaluation capacity building in the framework of the new challenges of EU structural policy*. Paper presented at the 44th Congress of European Regional Science Association, Porto, Portugal.
- King, J. A. (2002). Building evaluation capacity in a school district. In D. W. Compton, M. Baizerman, & S. H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building*. *New Directions for Evaluation*, 93, 63–80.
- King, J. A., & Volkov, B. (2005). A framework for building evaluation capacity based on the experiences of three organizations. *CURA Reporter*, 35(3), 10–16.
- Kramer, M., Graves, R., Hirschhorn, J., & Fiske, L. (2007). *From insight to action: New directions in foundation evaluation*. Boston, MA: FSG/Social Impact Advisors.
- Kuzmin, A. (2004). *Exploration of factors that affect the use of evaluation training in evaluation capacity development*. Unpublished doctoral dissertation, Union Institute & University, Cincinnati, Ohio.
- Mark, M., & Beery, W. (2004). Making judgments about what to evaluate and how intensely. In M. Braverman, N. Constantine, & J. Slater (Eds.),

Foundations and evaluation: Context and practices for effective philanthropy (pp. 145–165). San Francisco, CA: Jossey-Bass.

- Merriam, S. (1998). *Qualitative research and case study applications in education*. San Francisco, CA: Jossey-Bass.
- Naccarella, L., Pirkis, J., Kohn, F., Morley, B., Burgess, P., & Blashki, G. (2007). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning, 30*, 231–236.
- Patton, M. Q. (1990). *Qualitative evaluation and research methods*. Newbury Park, CA: Sage.
- Slater, J., Constantine, N., & Braverman, M. (2004). Putting evaluation to work for foundations. In M. Braverman, N. Constantine, & J. Slater (Eds.), *Foundations and evaluation: Context and practices for effective philanthropy* (pp. xxxv–xlvi). San Francisco, CA: Jossey-Bass.
- Stockdill, S., Baizerman, M., & Compton, D. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. In D. W. Compton, M. Baizerman, & S. H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, 93*, 7–25.
- Taut, S. (2007a). Defining evaluation capacity building: Utility considerations. *American Journal of Evaluation, 28*(1), 120.
- Taut, S. (2007b). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*(1), 45–59.
- Volkov, B., & King, J. (2007). *A checklist for building organizational evaluation capacity*. Kalamazoo, MI: The Evaluation Center, Western Michigan University. Available at <http://www.wmich.edu/evalctr/checklists/ecb.pdf>

Boris Volkov, Ph.D., is an Assistant Professor of Evaluation Studies with the Center for Rural Health and Department of Family & Community Medicine at the University of North Dakota School of Medicine and Health Sciences.