

Building and Cultivating Internal Evaluation in the Not-for-Profit Sector: A Critical Evaluation of Our Roles

Lymari Benitez, Ph.D., & Kira Krenichyn, Ph.D.

Paper for the American Evaluation Association 2011 Annual Meeting

Introduction:

In this paper, we will reflect on our own experiences as internal evaluators trying to change the role of “outside expert” that evaluators have traditionally assumed, usually with low stakeholder participation. We are advocating for internal evaluation that relies on stakeholders’ involvement and provokes opportunities to critically reflect upon data. We are interested in building and cultivating evaluation processes that are integrated in daily work and that promote growth informed by measurable outcomes and results. Although our experiences are different (one of us being the only internal evaluator at a small organization, and the other is a part of a large evaluation team), we both agree that successful internal evaluation needs to be participatory and requires support from senior management. We also agree that evaluation activities should ultimately play a key role in creating learning organizations. Thus, this paper will address the questions: What challenges have we, as internal evaluators, experienced? What strategies have we used to answer those challenges, and with what? What can we, as internal evaluators, do to further organizational learning goals—what are our roles in furthering those goals?

Internal Evaluation in the Not-for-Profit Sector

In recent decades, the role of evaluators in the not-for-profit sector has grown. Increased delivery of social services and associated demands for reporting and accountability have created the need for nonprofits to build internal capacity for program evaluation. Government initiatives (for example, Investing in Innovation Fund, Choice Neighborhoods and Promise Neighborhoods) and many foundations now require rigorous evaluations of the programs that they fund. The philanthropic sector has also embraced the notion of accountability and requires reports, logic models, surveys and site visits among others. These documents and activities are used to assure that money is well spent. This reflects an interest in making organizations accountable for the outcomes they hope to obtain and assuring funders that the initiative or program funded is a good investment.

In the government sector, The Government Performance and Results Act of 1993 (GPRA) (<http://www.whitehouse.gov/omb/mgmt-gpra/index-gpra>) requires that government funded programs identify clear and obtainable goals and measure their progress and success against these goals. In addition, in 1999 the federal government instituted performance measurements for all agencies (Hoole & Patterson, 2008). These requirements place stronger attention on results. To illustrate, the Promise Neighborhoods (PN) grant application, a program sponsored by the Department of Education and private funders, seeks to improve children’s and community outcomes in distressed urban communities and includes a set of indicators that grantees will use to measure their progress and success through rigorous evaluation. This process will provide these organizations with enough information to promote learning about the overall impact of the program and identify successful strategies supporting the achievement of positive outcomes (<http://www2.ed.gov/programs/promiseneighborhoods/index.html>). Examples like this imply that organizations receiving these grants would need to design and develop systems to collect, track and analyze data that can be used for learning purposes. In other words, the grant requires grantees to build their evaluation capacity¹ and be accountable for funds received.

¹ Evaluation capacity building (ECB) is defined as the process of improving an organization’s ability to use evaluation to learn from its work and improve results.

As documented, reporting to funders is one of the most frequently reported purposes of evaluation (Carman & Fredericks, 2008; Fine Thayer, & Coghlan, 2000), and in one study (Fine et al., 2000) reporting was the only evaluation function more often associated with lower versus higher participation—compared to other functions such as continuous improvement, monitoring, or promoting the organization. Reed and Moriaru (2009) likewise found that funders were the highest priority audience for evaluation. The time that organizations devote to these activities could become a burden if the organization does not have the resources to fulfill these requirements.

However, not all not-for-profits view evaluation as tool to generate data for funding reports. Carman and Fredericks (2010) found that older organizations reported fewer difficulties with carrying out evaluation beyond funders' reports. This finding connects with literature that indicates evaluation is related to organizational life cycles, and the age of an organization may also relate to other organizational variables, such as ability to adapt to change (for example, older organizations may be more mature to engage in data discussions). Organizations that reported fewer obstacles also described ways that they incorporated evaluation into overall management practices, such as regularly reporting to the board, including measures in annual reports, and routinely collecting data. Hoole and Patterson (2008) described organizations where evaluation had successfully been employed for organizational learning. Organizations made serious commitments of staff and resources, developed learning teams that included a mix of stakeholders, and promoted ongoing dialogue.

To deal with these funding demands, nonprofit organizations are pressured to identify strategies that best fit their needs and allow them to comply with expectations. Some organizations rely exclusively on external consultants while others usually assign staff members the responsibility of gathering evaluation data. Carman and Fredericks (2008) conducted a survey of non-profit organizations in Indiana with the goal of describing how evaluation looks in practice and found that the majority of nonprofits (80% of those surveyed) relied on internal staff for evaluation activities. The level of expertise and knowledge of staff members responsible for evaluation activities was diverse. In the majority of the organizations, program staff members with basic knowledge (or no knowledge) of evaluation methodologies engaged in evaluation activities as part of their program responsibilities, while in only a few, staff members with knowledge in evaluation assumed the formal role of internal evaluators. In Carman and Fredericks (2008) study only 4% of the organizations reported having a dedicated evaluation staff. Although the study did not use a national representative sample, our experience seems to indicate that it reflects the state of evaluation in the non-profit sector (at least in NYC).

Role of Internal Evaluator: Challenges and Strategies

Although internal evaluation units are not the norm in the nonprofit sector, we work for organizations that choose to invest in internal evaluation. In both organizations internal evaluators have assumed the role of advisors to management, contributors to the strategic planning and vision of the organization, monitors of program quality and developers of reports for outside funders. All these are possible because, as internal evaluators, we are more in-tune with the day-to-day practices of our organizations. We are responsible for understanding how goals are set and measure as well as report our progress to different stakeholders. We are in a position where we produce and translate information with a full understanding of the context and nuances of each organization. Therefore, the role of an internal evaluator should include facilitating the involvement of multiple stakeholders and the process of organizational learning. However, there are many challenges to consider.

1. Meeting accountability needs

Accountability and funders' demands/requirements are often the impetus for evaluation, but organizations may then begin to use evaluation in other ways. Conversely, evaluation for reporting results may continue to be felt as a drain and distraction (Carman & Fredericks, 2008). No matter what, accountability remains an important purpose of evaluation across the board; not-for-profits should be held accountable to all stakeholders, including the individuals and communities that they serve.

But there can be challenges as long as an organization is only responding to accountability demands, and until there is a more learning-oriented approach. Staff misconceptions about evaluation processes and how results are used can hinder their disposition to participate in evaluation activities. Evaluation activities may be perceived as reflecting senior management agendas that are far removed from daily practices. Also, staff members may perceive that evaluation results are not used or used only to make radical decision like firing people. Anxieties and confusion also seem to be a part of many conversations about evaluation, perhaps because stakeholders feel that they are being asked to be social scientists but don't want to be or don't feel prepared to be. Managers also sometimes seem to place fundraising (while an important goal) above all else, forgetting the organization's mission or the importance of learning from results, focusing instead on concerns such as how to best find positive results or whether a potential funder will make unreasonable requests for evaluation.

Organizational leaders may also be inexperienced in using data, and individuals throughout an organization may feel threatened or uncomfortable with engaging in dialogue about findings (Torres & Preskill, 2001). If participation is strictly utilization-focused, it can become tokenistic—a mechanism of control over stakeholders, a means for forcing acceptance of evaluation process and findings, and more for the benefit of the organization than for the stakeholders. On the other hand, if evaluation is too focused on empowerment and transformation, it can slow down a project with results that are too abstract and immeasurable, or it can neglect the needs of majority groups, achievement of organizational missions, and broader social purposes (Herrera, Springett, & Kagan, 2009).

When there is a lack of knowledge or misunderstandings of the goals of evaluation, it is important for internal evaluators to incorporate into their evaluation plan the voices of the evaluation recipients. It is also important for leaders supporting evaluation to make sure that these processes are not seen as threatening to their staff and results are used to help improve the work (Hoole & Patterson, 2008). In many cases, this implies identifying current needs of the program. A good strategy to check if program needs are addressed within the evaluation plan is to discuss a draft of the intended evaluation. This discussion provides an opportunity to align senior staff and program staff priorities. In many cases, these priorities will be very similar and related to the expected outcomes of the program. If they are not, then the internal evaluator has gained valuable information that should be taken into consideration during the evaluation.

Because on many occasions timelines for evaluations and program reporting are different, it is also important to identify when data needs are critical to a program. In one of our organizations, the internal evaluator is the person in charge of engaging in conversations with program staff and identifying critical data points and times for reporting. This information is then used to assure that all internal as well as external evaluation plans include ways of feeding back data to the programs.

2. Commitment of evaluation resources

Doing evaluation right requires commitment of human and financial resources to ensure that the right research questions are asked, the appropriate social science research methods are employed, and results are interpreted clearly, correctly, and fairly. Building evaluation capacity may be a challenge because of existing demands on staff time, resources, and expertise. Organizations frequently report too few resources and too little support to carry out evaluation for organizational learning and change. Lack of staff time is frequently reported as a barrier to doing evaluation, along with limitations to other resources such as expertise, money, technology, data collection capacity, and standards for identifying outcomes (Carman & Fredericks, 2010; Reed & Morariu, 2009). Innonet's recent survey of over 1,000 organizations found that too few organizations (25%) spent the recommended portion of their budgets on evaluation, and one in eight spent no money on evaluation (Reed & Morariu, 2009). Carman and Fredericks (2008) found that 29% spent no money on evaluation. Without the commitment of resources, it seems to follow that there is insufficient capacity for producing legitimate results, and worse, pressure to report funders can lead to time wasted on "input and output" and production of unused and even misleading data (Hoole & Patterson, 2008). Furthermore, technical quality is "only one of the constructs" of successful evaluation, and an over-reliance on outside expertise can leave staff members feeling unprepared and unqualified to participate in evaluation (Herrera et al., 2009).

A study of three organizations in the Twin Cities area found that mandated evaluations drove the agenda in all three organizations, and the key resource of time was limited at all three. At the same time, each organization was able to build evaluation capacity and therefore address the challenges of evaluation while meeting accountability demands. Based on their experiences, a framework for building evaluation capacity includes internal and external that support evaluation; structures including a purposeful plan, infrastructure to support evaluation activities (e.g., a mechanism for framing questions, data collection systems, ability to carry out research), and peer learning structures; and resources including professional development, personnel, time, and fiscal resources. Internal and external environments must be supportive of and open to change, and individuals should be rewarded for creativity and risk-taking. Leadership should publicly support evaluation (e.g., holding evaluation meetings), and champions are needed (like one organization's "Evaluation Platoon"). Although this framework includes a host of internal and external supports, the authors point out that many structures can be build deliberately and that "people and organizations can learn evaluation by doing it" (King & Volkov, 2005).

Ideally, evaluation resources include a combination of both internal and external time and expertise. A related challenge is the difficulty of finding external evaluators who have interest and experience in using evaluation for organizational learning and change and who can support the development of internal evaluators or program staff (Torres & Preskill, 2001). Helping program staff to articulate a program theory is an important piece of building evaluation capacity, because it can help them to have a better sense of what is measurable and what is not, how to match measurement to program scope, and how their measurable outcomes can connect to ultimate, "saving the world" level outcomes. It is also important to identify which staff members are the primary intended users, or PIUs, of an evaluation, and whether and how they can affect changes in programs using evaluation processes and results (Monroe, Fleming, Bowman, Zimmer, Marcinkowski, & Mitchell, 2005).

3. Promoting a culture of evaluation and learning

A lack of an understanding of what program evaluation entails impedes the development of trusting relationships among internal evaluators and program staff. This relationship is key for successful implementation of internal evaluation processes as internal evaluators need to rely on program staff to gather and track data and other stakeholders for interpretation of findings and learning opportunities. One case study of a community-based health initiative in the UK concluded that the evaluation was unsuccessful, because it failed to establish sufficient communication among stakeholders and did not engage strategic decision makers in learning. Even though clients and program managers were engaged and had positive responses to the evaluation process, funding agencies did not buy in and were uncommunicative from the beginning of the evaluation project. They did not give feedback on final reports, and they did not commit to continue funding the initiative. The study's authors concluded that the structural barriers—differing cultures of the stakeholder organizations, restructuring and competing priorities, and a culture of short-term funding—were too difficult to overcome. Two-way communication needs to be established early on, and evaluation must be seen as a longer-term learning process (Reeve & Peerbhoy, 2007). Moreover, internal evaluators may be perceived as “spies” and greeted with suspicion. Program staff may not necessarily see the internal evaluator as a “colleague” and do not welcome evaluation activities as part of their day-to-day work. They think these activities disrupt their program delivery and are only a tool use to attract external funders.

A learning approach should include the stakeholders who are closely connected to organization in order to produce the most valid results (e.g., results that are based on the right research questions, and that are more likely to be put to appropriate uses). A learning approach should also include expertise in evaluation to ensure that the right methods are used and that results are interpreted correctly and fairly—that the results are valid. Thus, when findings are ready for discussion, it is important to include senior and program staff in the process of finding explanations of what is happening (e.g., why aren't we getting the expected outcomes). The internal evaluator has enough knowledge to identify meaningful ways of reporting findings as well as recommendations to the different stakeholders. For example, to a senior manager a finding can be explained based on its impact on the strategic plan while to program staff explanations can be related to implications for direct services.

These findings also suggest the need for very high-level changes, but some practical strategies might be used to address barriers and promote an evaluation culture. Conversations and meetings describing the evaluation process and communicating its goals are essential to get staff buy-in and cooperation. Examples of practical strategies are user-friendly reports that include recommendations and simple documents outlining stakeholder involvement and responsibilities.

4. Expanding the role and functions of evaluation

Research on evaluation (Carman, 2004, 2007; Carman & Fredericks 2008, Price 2011) has documented that non-profits think of evaluation in very broad terms and the role of internal evaluators to reporting and monitoring activities. Very few organizations in these studies gathered outcome data that can be used for comparisons and assessing impact. Most of the evaluation data consistently gathered by organizations relates to program participation (number of people served) and performance monitoring. This type of data is extremely important for service organizations as it provides opportunities to assess their service capacity, inform decision making, and increase capacity and accountability for outcomes. However, these data do not answer more complex questions related to why certain outcomes are obtained or whether the program is producing expected outcomes (Price, 2011).

In addition, the recent emphasis on outcomes has probably been too narrow, and other evaluation approaches need to be supported. Carman and Fredericks (2008) offer some critique of the current proliferation of “outcomes” as the focus of evaluation. Funders want to see the impact and therefore provide trainings in outcomes evaluation, but the content of the training is not always relevant to a given organization’s mission, and other evaluation approaches can be just as useful for learning and making improvements. For example, in one of their case studies, a cost-benefit analysis of a crafts sales program showed that it was costing the organization more money to sell crafts made by program participants than the participants were making; it would have made more sense for the organization to give the money directly to the participants, so they eliminated the program. Other research has found that the emphasis on outcomes has steered organizations away from seeking funding that would require reporting outcomes that they are worried they won’t be able to report. Funders need to be willing to support evaluation for learning purposes, and organizations need to push hard to ensure that evaluation meets their needs and measures what they will really impact (Hoole & Patterson, 2008).

Torres & Preskill (2001) identify a set of “catalysts” for building organizational learning, which takes place through stages leading to adoption and then refinement, often as a result of a crises or realization that evaluation is needed or findings are going unused. These catalysts are at a macroscopic level; in addition to an organization’s realization that a learning approach is needed, or the support and acceptance of staff, the not-for-profit field also needs guidelines for using evaluation, a diverse and capable pool of evaluators, educational and professional development opportunities, learning opportunities within organizations, and the support and vision of funders and legislators (e.g., accepting shorter-term indicators of progress toward achieving a mission). This implies that strategies for achieving organizational learning would be at the broader scale, for example including advocacy for funding of evaluation efforts, creation or expansion of intermediary organizations, and networks of not-for-profits who can support one another.

Internal evaluators play a key role in reporting and monitoring activities and are equipped to facilitate conversations and processes that incorporate data into the decision making process. However, their role should not be limited to these activities. By bringing monitoring and reporting issues to the attention of decision makers, they can influence conversations related to why specific changes can lead to program improvements and learning. They can outline the foundations for more complex evaluations and spear collaborative initiatives with external experts. They can play a lead role in plotting an organization’s stages of growth and developing theories of logic models, not just for programs but for evaluation itself.

References

Carman, J. G., Fredericks, K. A. (2008). Nonprofits and evaluation: Empirical evidence from the field. *New Directions for Evaluation*, Fall (119), 51-68.

Carman, J. G., & Fredericks, K. A., (2010). Evaluation capacity at nonprofit organizations: Is the glass half-empty or half-full? *American Journal of Evaluation*, 31 (1), 84-104.

Fine, A. H., Thayer, C. E., & Coghlan, A. T. (2000). Program evaluation practice in the nonprofit sector. *Nonprofit Management & Leadership*, 10 (3), 331-339.

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.

Monroe, M. C., Fleming, M. L., Bowman, R. A., Zimmer, J. F., Marcinkowski, J. W., & Mitchell, N. J. (2005). Evaluators as educators: Articulating program theory and building evaluation capacity. *New Directions for Evaluation*, Winter (108), 57-71.

Price, K.M. (2011). The evolution of understanding: Positioning evaluation within a comprehensive performance management system. In S. Mathison (Ed.), *Really New Directions in Evaluation: Young Evaluators' Perspectives*. *New Directions for Evaluation*, 131, 103-109.

Reed, E. & Moriarty, J. (2010). State of evaluation: Evaluation practice and capacity in the nonprofit sector. Washington, DC: Innovation Network.

Reeve, J., & Peerbhoy, D. (2007). Evaluating the evaluation: Understanding the utility and limitations of evaluation as a tool for organizational learning. *Health Education Journal*, 66, 120-131.

Torres, R. T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present, and future. *American Journal of Evaluation*, 22, 387-395.