Article

Understanding Dimensions of Organizational Evaluation Capacity

Isabelle Bourgeois[1] and J. Bradley Cousins[2]

Abstract

Organizational evaluation capacity building has been a topic of increasing interest in recent years. However, the actual dimensions of evaluation capacity have not been clearly articulated through empirical research. This study sought to address this gap by identifying the key dimensions of evaluation capacity in Canadian federal government organizations. The methodology used, based on Leithwood and Montgomery's *Innovation Profile* approach, featured semistructured interviews with evaluation experts and a validating exercise conducted in four government organizations. The framework developed as a result of the study identifies six main dimensions of evaluation capacity (human resources, organizational resources, evaluation planning and activities, evaluation literacy, organizational decision making, and learning benefits), each one broken down into further subdimensions. The evaluation capacity of organizations on each of these dimensions and subdimensions can be described using four levels: low, developing, intermediate, and exemplary. The study found that government organizations vary in terms of their capacity from one dimension to the next, and indeed, from one subdimension to the next.

Keywords

evaluation capacity, evaluation capacity building (ECB), organizational learning, evaluation practice, evaluation utilization, government

Introduction

Interest in evaluation capacity building (ECB) has increased in recent years, following an initial treatment of the issue in a volume of *New Directions for Evaluation* published by Compton, Baizerman, and Stockdill in 2002. Much of this work has focused on ECB in organizations and there is a growing body of conceptual and empirical work on the topic (see, e.g., Cousins, Goh, Clark, & Lee, 2004; Preskill & Boyle, 2008a). Yet, although knowledge is advancing about building the capacity

[1] École nationale d'administration publique, Gatineau, Québec, Canada
[2] University of Ottawa, Ottawa, Ontario, Canada

Corresponding Author:
Isabelle Bourgeois, École nationale d'administration publique, 283 Alexandre-Taché Boulevard, Section D, Gatineau, Québec, Canada.
Email: isabelle.bourgeois@enap.ca

of organizations to do evaluation and, to a lesser extent, *use* evaluation, little attention has been directed toward defining organizational evaluation capacity itself. In this article, we develop and empirically validate a framework for organizational evaluation capacity and consider implications of the framework for ongoing research and practice.

Results-based management (RBM) is an important feature of a new public management government framework applied in service organizations around the world. Managing for results requires a comprehensive system of performance measurement and program evaluation to foster increased accountability in public organizations (Jorjani, 2008; Mayne, 2009). Despite RBM's potential, in practice many challenges exist in its implementation. For example, in the Government of Canada, the responsibility for performance measurement is placed in the hands of program managers because of their substantive knowledge (Treasury Board Secretariat, 2010). However, program managers often have neither the appropriate expertise nor guidance to undertake complex performance measurement exercises. This results in a scarcity of high-quality performance measurement data. Similarly, in the United States, the passage of the Government Performance and Results Act (GPRA) in 1993 and the implementation of the Program Assessment Rating Tool (PART) in 2004 required federal agencies to focus on establishing quantifiable measures of progress and reporting on their success. Although promising, these initiatives have not fully achieved their objectives; studies show that even if they have resulted in an increased availability of performance information, questions remain as to the tool's use for budgetary allocation and program decision making (Mark & Pfeiffer, 2011; Mathison, 2011). More recent initiatives, such as the Performance Improvement Council (PIC), aim at making the PART process more transparent and incorporating input from various sources. These new initiatives further recognize the need to increase the capacity of organizations and individuals to use data to make fundamental program decisions (Mark & Pfeiffer, 2011). Other countries have also moved in the direction of increasingly more sophisticated performance measurement or centralized national evaluation functions, but have not necessarily been successful at integrating performance data and evaluation findings into budgetary allocation processes (see, e.g., Talbot's presentation of the United Kingdom's performance and evaluation system, 2010, and a discussion of the Spanish context by Feinstein & Zapico-Goni, 2010).

Aside from budgetary allocations and ongoing program administration, one of the main uses of performance measurement data in RBM systems is for periodic evaluation studies. Authentic engagement with evaluation, however, may be easier said than done. In Canada, for example, given increased requirements for evaluation coverage (as per the Treasury Board's *Policy on Evaluation*, 2009) and a relatively conservative level of resources allocated to the evaluation function, departmental evaluators must use available data whenever possible to increase their efficiency. The implementation of ECB initiatives in this and other federal government contexts, therefore, offers a potential bridge between the technical expertise required to conduct evaluative activities and the substantive knowledge of program managers and staff.

ECB refers to the changes undertaken by organizations to integrate evaluation practice and use at all levels (Boyle, Lemaire & Rist, 1999; Cousins et al., 2004; Sanders, 2002; Stockdill, Baizerman, & Compton, 2002). One of the most commonly used definitions of ECB is provided by Stockdill and her colleagues (2002):

> . . . a context-dependent, intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its appropriate uses are ordinary and ongoing practices within and/or between one or more organizations/programs/sites. (p. 8)

Added to greater concerns about evaluator recruitment and training in the federal community, ECB has become an issue of interest in recent years (Mayne, 2009; Preskill & Boyle, 2008a, 2008b). This

is also true of other jurisdictions; for example, Compton and MacDonald (2008) propose ECB as a strategy to strengthen evaluation services and program effectiveness in the face of fluctuating program funding.

In their comprehensive review of the literature on the integration of evaluation into organizational culture, Cousins and his colleagues (2004) identify two types of ECB: direct ECB, which involves planned ECB activities that occur either within or outside of actual evaluation projects (e.g., training on statistical data analysis), and indirect ECB, which results from involvement of stakeholders in processes that produce evaluation knowledge. In essence, indirect ECB is akin to participatory evaluation, that is, evaluations that are conducted in partnership between those trained in evaluation logic and methods and members of the program or stakeholder organization community (Cousins & Chouinard, 2012). However, these ECB processes differ from participatory evaluation approaches in two ways: They are typically integrated into the organization's practices and they are ongoing rather than episodic or event-driven (Preskill & Torres, 1999; Rowe & Jacobs, 1998; Stockdill et al., 2002).

ECB processes have been linked to two consequences for organizations: evaluation use and organizational learning (Cousins et al., 2004). Evaluation becomes better understood and more useful in organizations that implement intentional ECB strategies. In this way, ECB initiatives foster the development of a culture of systematic self-assessment and reflection (Cousins et al., 2004) that, in turn, can lead to increased organizational learning, referred to as "the vehicle for utilizing past experiences, adapting to environmental changes and enabling future options" (Berends, Boersma, & Weggerman, 2003, p. 1036). Thus, ECB represents one of the ways through which individual-level learning may be transferred to the organizational level (Berends et al., 2003; Popper & Lipshitz, 2000) and sheds light on how organizations can move beyond single-loop (or incremental) learning into double-loop learning (Argyris & Schon, 1978).

Organizational Factors Contributing to the Success of ECB

A number of factors or conditions leading to successful ECB in organizations have been identified in recent years. In order to clarify and organize these factors, we have classified them into the four categories outlined below.

- *External environment.* External accountability requirements often create a demand for evaluation results and so act as a motivator for developing evaluation capacity (Gibbs, Napp, Jolly, Westover, & Uhl, 2002; Katz, Sutherland, & Earl, 2002; Mackay, 2002; Stockdill et al., 2002; Sutherland, 2004; Toulemonde, 1999).
- *Organizational structure.* The systems and staffing structures of organizations mediate organizational members' ability to interact, collaborate, and communicate with each other (Preskill & Torres, 2000). Successful ECB depends on the flexibility of organizational roles, since individuals must be able to step away from their main responsibilities to participate in evaluation activities (Torres & Preskill, 2001).
- *Organizational culture.* The culture of an organization reflects the traditions, values, and basic assumptions shared by its members and that establish its behavioral norms. The culture of an organization involved in ECB must encourage questioning of organizational processes and experimenting with new approaches (Goh, 2003; Preskill & Torres, 1999; Rowe & Jacobs, 1998; Torres & Preskill, 2001; Toulemonde, 1999).
- *Organizational leadership.* Managerial support is necessary to the implementation and sustainability of evaluation capacity within an organization (Cousins et al., 2004; Goh, 2003; Goh & Richards, 1997; King, 2002; Milstein, Chapel, Wetterhall, & Cotton, 2002; Owen & Lambert, 1995).

Although there is general support for these categories in the literature, a stronger empirical basis is warranted.

State of Research on ECB

As we have shown, the factors likely to influence the success of ECB in an organization, as well as its ultimate consequences, have been identified in the theoretical evaluation literature. In addition to the anecdotal reports of ECB that have been published (see, e.g., Diaz-Puente, Yague, & Afonso, 2008; Garcia-Iriarte, Suarez-Balcozar, Taylor-Ritzler, & Luna, 2011; Lawrenz, Thomas, Huffman, & Covington Clarkson, 2008; Taut, 2007; Volkov, 2008), work has been done to identify the stages through which organizations move as they develop their evaluation capacity (Bourgeois & Cousins, 2008), and how ECB might best be conceptualized (Huffman, Thomas, & Lawrenz, 2008; Preskill & Boyle, 2008a; Taylor-Powell & Boyd, 2008). However, few empirical studies have focused on how evaluation capacity is manifested in organizations and how it can be assessed (one recent example is found in Nielsen, Lemire, & Skov, 2011). Such information would advance our knowledge and provide a backdrop for further work. Thus, in this article we attempt to identify the key dimensions of evaluation capacity in organizations, operationalized through a framework based on the *Innovation Profile* approach developed by Leithwood and Montgomery (1987). From a practical perspective, this framework offers organizations a model for its members to reflect on their capacity development activities. The framework can also be used as the basis for the development of an instrument focusing on organizational self-assessment of evaluation capacity. Accordingly, we addressed the following research questions in the current study:

1. What are the essential dimensions of evaluation capacity in Canadian federal government organizations?
2. How are minimal and exemplary performance on each of these dimensions characterized?
3. What are the steps required to move from minimal to exemplary performance?

Method

Data collection encompassed three phases, reflecting an adaptation of the innovation profile approach (Leithwood & Montgomery, 1987). Conceptually, this approach—which was developed in the education sector within the context of implementing planned changes in classroom practices—focuses on growth defined by observable change from a current state of practice toward an ideal state. The process involves identifying concrete behavioral manifestations of the current state and building a series of manageable steps for multiple dimensions of the desired innovation. These steps should be challenging enough to represent observable change from the previous state, but be feasible in order to enable step attainment or success in moving from one step to the next (Leithwood & Montgomery, 1987). The descriptions developed for each behavioral change are generally based on a qualitative data collection process. Application of the innovation profile approach thus results in a multidimensional matrix describing growth in performance or, in the case of this study, evaluation capacity development in organizations.

The innovation profile strategy was used by Cousins, Aubry, Smith-Fowler, and Smith (2004) as an alternative approach to process evaluation in their study of mental health case management (Cousins et al. refer to the approach as *key component profiles.*). We argue that it is well suited to the study of organizational evaluation capacity because of its focus on the incremental steps required to move from low to high capacity and its flexibility, defined in terms of the inclusion of varying numbers of levels across dimensions as well as its accommodation of a wide array of dimensions (and subdimensions). The three phases undertaken as part of the current study are summarized below.

Phase 1: Identification of Key Dimensions of Evaluation Capacity (Divergent Phase)

The first phase focused on identifying the key dimensions of evaluation capacity through an in-depth literature review and a series of expert interviews. An important aspect of the literature review involved moving beyond descriptions of capacity building initiatives undertaken in various organizations to definitions and features of evaluation capacity itself.

Once the literature review was completed, we conducted semistructured interviews with expert informants who have a broad view of evaluation in the Canadian federal government. We recruited four individuals for the first phase of the study; two were external consultants who have worked with several departments and agencies on evaluation studies and two were former or current senior officials of a central agency of the government of Canada who have worked on interdepartmental evaluation issues and are familiar with the challenges faced by different departments and agencies as they develop their evaluation capacity. Their point of view, as insiders of the federal evaluation community but outsiders with respect to the evaluation function of specific departments and agencies, informs their overall vision of how evaluation capacity appears in various organizations. The purpose of these interviews was to obtain these experts' definitions of evaluation capacity as well as to solicit their views on behavioral manifestations of capacity.

In our content analysis of the literature review and interview data, potential dimensions and markers of evaluation capacity were used to identify the main categories for coding purposes. We summarized the results of this analysis in a draft framework of evaluation capacity.

Phase 2: Review and Feedback on Draft Framework (Convergent Phase)

The second phase of data collection focused on confirming the key dimensions of evaluation capacity derived from Phase 1. We once again used key informant interviews with the four experts consulted in the first phase of the study. We asked participants to review the draft framework and provide feedback on its clarity and contents. Based on this review, we could confirm existing dimensions and subdimensions or identify challenges that warranted changes to the framework.

Phase 3: Triangulation of Findings Included in the Framework

The third phase was a validation exercise undertaken to finalize the draft evaluation capacity framework. It focused on key informant interviews with evaluators and decision makers from four federal government departments and agencies. The participating organizations were selected on the advice of the experts consulted previously and were chosen to ensure varying levels of evaluation capacity as assessed by the experts. The representatives were asked to implement the framework in their own settings and provide feedback on its utility in terms of organizational reflection and improvement. We contacted three individuals in each organization: the Head of Evaluation, a senior evaluator, and a decision maker. We conducted 11 interviews in this phase of the study.

As with the previous interviews, we used a qualitative content analysis to identify trends in the data. Because of the increased complexity associated with the use of four different organizations and three different organizational roles, data coding and analysis were more detailed than in the first two phases and took these types of variables into account. First, the data were aggregated by organizational role; this analysis enabled us to validate and further refine the categories of evaluation capacity included in the draft framework. Second, data were aggregated and analyzed by organization; the findings from this analysis have been reported elsewhere (Bourgeois & Cousins, 2008).

Results

The final version of the framework, presented in Tables 1–6, provides a summary of our key findings. A more detailed description of these results follows.

Structure of Framework

The framework presents the dimensions of evaluation capacity as identified in Canadian federal government organizations. Several structural elements were utilized to ensure clarity and consistency. Six main dimensions emerged from the three data collection phases, which we divided into two broad categories: "capacity to do" evaluation and "capacity to use" evaluation. Most participants focused on the "capacity to do" category, likely because the dimensions included here are easier to control and speak to the more operational facets of evaluation. Each dimension is further organized into a number of subdimensions; again, these were based on interview data and focus on more specific descriptions of the dimension. The final components of the framework distinguish the differing levels of evaluation capacity: "low capacity," "developing capacity," "intermediate capacity," and "exemplary capacity."

The first main dimension (see Table 1), Human Resources, addresses the composition of the evaluation unit itself and is divided into five subdimensions. The first subdimension, Staffing, refers to the balance of evaluation positions within the organization and whether these are sufficient to manage the workload identified in the evaluation plan. It also includes career progression for evaluators, which deals with employee retention, and succession planning, two issues crucial to capacity building and maintenance. The second and third subdimensions focus on the technical and interpersonal skills required of evaluators. Skills related to the identification of evaluation issues, the use of appropriate data collection methods, the generation of evidence-based recommendations, and project management are part of the technical abilities required of evaluators. "Softer" skills such as building client trust, communicating evaluation messages in a clear and transparent way, and meeting program stakeholders' informational needs are part of the communications and interpersonal skills used by evaluators. The fourth subdimension involves professional development and includes elements related to both internal and external professional development activities, as well as the development of learning plans for evaluation staff members and ongoing assessments of the skill set that exists within the evaluation unit. Finally, the fifth subdimension refers to the quality of the leadership within the evaluation unit. Good leaders should have both evaluation and management experience, be able to translate the information needs of senior managers into concrete project plans, and act as mentors or coaches for team members.

Participants focused heavily on the Human Resources dimension during the interviews, especially those directly involved with evaluation. This observation suggests that, in their view, the essence of evaluation capacity may be more heavily aligned with the Human Resources dimension, rather than a more balanced perspective including all six dimensions.

The second dimension (Table 2) is Organizational Resources. Three subdimensions are included: budget, ongoing data collection, and organizational infrastructure. Budget refers to the stability of the evaluation budget and whether it provides sufficient funding to complete the activities outlined in the evaluation plan. Ongoing Data Collection speaks to the performance measurement systems that are in place within the organization and that produce information that is fed into evaluation studies. Organizational Infrastructure is the stability of the governance structure, the existence of organizational evaluation policies, and the organizational supports that help or hinder the work of evaluators, such as procurement services.

The third dimension (Table 3) focuses on the activities undertaken by evaluators as part of their regular duties. The development of an organization-wide evaluation plan is key among the subdimensions that make up this section. It is characterized by the development of an evaluation plan in consultation with other stakeholders, the inclusion of a risk assessment process in the identification of evaluation priorities, ongoing intelligence gathering, and a systematic review of the evaluation unit itself. Evaluators in most departments use consultants to some extent, so it was included as a subdimension. Information sharing within the unit was included here as well, since evaluation staff

Table 1. Capacity to Do Evaluation, Dimension 1: Human Resources.

Level	Staffing	Evaluation logic and technical skills	Communications and interpersonal skills	Professional development	Leadership
Exemplary capacity	Evaluation unit is optimally staffed (i.e., positions are created and filled based on operational requirements outlined in long-term evaluation plan) Appropriate balance of senior and junior evaluator positions, given organizational requirements (i.e., size of organization, proportion of work done in-house, etc.) Career progression process is in place to facilitate promotions within unit and ensure succession planning through employee retention	Evaluation issues are clearly identified and linked to ongoing organizational concerns and priorities Innovative use of methods and approaches to data collection (e.g., routine use of complex survey methodology) Recommendations made in evaluation reports are clearly linked to evaluation findings Evaluation projects are generally well managed and problems are usually identified and resolved quickly by senior evaluators	Evaluation unit has clearly established client trust within the organization Evaluation reports and other products deliver open, clear and transparent messages Clients feel that evaluators understand key organizational issues and respond to them appropriately (i.e., informational needs of program managers are met through various evaluation activities)	Assessment of skill sets among staff is done regularly and learning activities are arranged to fill gaps within unit All team members have personalized learning plans All staff members engage in external professional development activities directly related to their work (e.g., conferences) Organization develops its own internal professional development activities (e.g., brown bag sessions, resources, in-house seminars)	Evaluation unit headed by individual with strong evaluation and management background Leader effectively reconciles expectations of senior management with operational requirements and resources of team Leader guides, mentors or coaches team members as part of his or her regular duties

(continued)

Table 1. (continued)

Level	Staffing	Evaluation logic and technical skills	Communications and interpersonal skills	Professional development	Leadership
Intermediate capacity	Evaluation unit is fully staffed (i.e., all positions are filled and adequate to meet operational requirements of annual evaluation plan) Appropriate balance of senior and junior evaluator positions, given organizational requirements (e.g., size of organization, proportion of work done in-house) Career progression process is available for some levels of evaluation professionals wishing to gain experience; part of succession planning	Evaluation issues are clearly identified and reflect the concerns of program managers Use of evaluation methods respects accepted standards and yields defensible findings (e.g., well-developed interview protocols) Recommendations made in evaluation reports are clearly linked to evaluation findings Evaluation projects are fairly well managed and problems are usually identified and resolved with help from evaluation unit leader	Evaluation unit is committed to building and maintaining client trust Evaluation reports deliver open, clear, and transparent messages Clients feel that evaluators understand key organizational issues (i.e., informational needs of program managers are taken into account in evaluation design)	Most team members have personalized learning plans Most staff members engage in external professional development activities directly related to their work (e.g., conferences) Some internal learning resources are made available to staff members (e.g., journals, books)	Unit headed by individual with some evaluation and management experience Leader effectively reconciles expectations of senior management with operational requirements and resources of team Leader guides team members as often as possible

(continued)

Table 1. (continued)

Level	Staffing	Evaluation logic and technical skills	Communications and interpersonal skills	Professional development	Leadership
Developing capacity	Evaluation unit is partially staffed (i.e., less than 50%; difficult to meet operational requirements outlined in annual evaluation plan) Few senior evaluation positions in the unit Career progression occurs in an ad hoc manner (i.e., through the usual competition process when turnover occurs) No succession planning efforts are underway	Evaluation issues are clearly identified Use of evaluation methods respects accepted standards and yields defensible findings (e.g., well-developed interview protocols) Recommendations made in evaluation reports are usually linked to evaluation findings Some project management issues may arise due to lack of experience or resources	Evaluation unit is working toward building client trust Evaluation reports are generally written in a clear manner (e.g., clients suggest edits to the reports or ask clarification questions) Clients feel that evaluators are open to learning about issues related to program management (i.e., program managers are able to make suggestions regarding evaluation questions)	Some staff members engage in professional development activities delivered by external organizations to improve generic skill sets (i.e., not necessarily directly related to evaluation)	Unit headed by individual who is new to the area of evaluation and/or has limited management experience Leader is not generally involved in senior management discussions and therefore assigns work based only on operational or treasury board requirements Leader coordinates team activities but is not involved in guiding team members in their work
Low capacity	Evaluation unit has several vacant positions (i.e., more than 50%; no link between evaluation plan and staffing actions) No senior evaluator positions in the unit No career progression process for evaluators due to small number of available positions in organizational chart No succession planning efforts are underway	Evaluation issues are not always identified Weak evaluation methods do not always yield defensible findings (e.g., reliance on external data sources with no verification) Recommendations made in evaluation reports are not linked to evaluation findings Problems with project management process arise often due to lack of experience or resources	Evaluation unit has not yet been able to develop client trust Evaluation reports often raise clarification questions from clients (e.g., clients do not understand the chain of results or specific pieces of evidence) Clients feel that evaluators are not open to learning about issues related to program management (i.e., program managers are not able to make suggestions regarding evaluation questions)	Staff members do not engage in professional development activities delivered by external organizations or provided in-house	Head of evaluation position is vacant

Table 2. Capacity to Do Evaluation, Dimension 2: Organizational Resources.

Level	Budget	Ongoing data collection	Organizational infrastructure
Exemplary capacity	Evaluation budget ensured through continuing funding specifically allocated to evaluation unit (i.e., is not shared with other corporate units) Evaluation budget is allocated based on evaluation plan (i.e., specific amount of budget based on plan)	Performance measurement system fully implemented across organization (or across programs in the case of large organizations) High-quality information is collected by the performance measurement system Performance data feed directly into results-based management (including evaluation studies)	The organization has a stable governance structure and clear accountability lines for results Organizational policies on evaluation and performance measurement have been developed and implemented Organizational culture ensures that performance data feed into a structured planning and reporting process (e.g., use of evaluation in RPP, budgeting) Existence of organizational supports that provide needed services to evaluators in a competent and timely manner, allowing them in turn to produce timely and useful evaluation reports (e.g., procurement, communications, HR, access to information)
Intermediate capacity	Evaluation budget is stable but shared with some other corporate groups (i.e., policy or audit) Evaluation budget is appropriate given the evaluation plan (i.e., plan does not determine the budget, but it is sufficient to complete planned activities)	Performance measurement system in place for major programs or activities Incomplete implementation of performance measurement system (e.g., sporadic data collection, missing variables) Performance data can be adapted to suit the informational requirements of results-based management (including evaluation studies)	Organizational infrastructure showing some maturity (e.g., clear governance structure and stability in senior management ranks; organizational policy on evaluation has been developed and implemented) Existence of a structured planning process that includes consideration of evaluation issues and findings (e.g., use of evaluation in RPP, budgeting) Organizational supports do not always provide needed services to evaluators in a timely manner and results in delays for evaluation projects
Developing capacity	Evaluation budget provided on a case-by-case basis for each new project Evaluation plan not considered in budget allocation	Performance measurement is done on a program-by-program basis Ad hoc implementation of performance measures, with uneven quality Performance data are difficult to integrate into results-based management (including evaluation studies)	Organizational infrastructure still under development (e.g., no clear governance structure, no policies on evaluation are in place, no structured planning process) Organizational supports often fail to provide needed services to evaluators
Low capacity	Organizational budget does not include funds for evaluation activities	Performance measurement is not done in the organization, either by programs or other corporate groups	Organizational infrastructure still under development (e.g., no clear governance structure, no policies on evaluation are in place, no structured planning process) Evaluators do not have regular access to organizational supports

Note. RPP = Report on Plans and Priorities.

Table 3. Capacity to Do Evaluation, Dimension 3: Evaluation Planning and Activities.

Level	Evaluation plan	Use of consultants	Information sharing	External supports	Organizational linkages
Exemplary capacity	Evaluation plan follows a 5-year cycle and is updated annually Evaluation plan developed in consultation with all senior managers and includes needs assessment exercise Evaluation plan includes thorough risk assessment process Ongoing intelligence gathering allows for changes to be made to plan when necessary (i.e., integrate emerging needs) Evaluation plan includes systematic review of evaluation unit itself (i.e., establishing and measuring service standards for the unit, assessing impact on organization)	Appropriate balance of evaluations designed and conducted by evaluation staff, or by consultants for specific expertise (e.g., when a project's scope is too large for the evaluation unit's resources, when dealing with complex interdepartmental evaluations or when specialized expertise is required) High-quality work produced by consultants when they are involved	Major decisions on evaluation projects are discussed within the unit to benefit from staff members' knowledge and experience Evaluators actively gather information on new developments in policy and strategic planning Knowledge management issues and processes are discussed regularly and common standards are followed by staff members	Evaluators make frequent use of external supports to evaluation such as professional associations, published standards, and so on. Evaluators are actively involved in broadening their external networks by liaising with evaluators in other organizations, engaging academic experts, and gathering information on the priorities of central agencies (i.e., Treasury Board, Privy Council) External stakeholders responsible for the delivery of federal government programs are involved in evaluation activities (where applicable)	Evaluators in regular contact with program managers through formal or informal ties Ready access to deputy head (i.e., lead administrator) on all aspects of evaluation; clear interest in evaluation information demonstrated by Deputy Head Evaluation unit located in close proximity to key organizational areas such as policy development, strategic planning and performance measurement units

(continued)

Table 3. (continued)

Level	Evaluation plan	Use of consultants	Information sharing	External supports	Organizational linkages
Intermediate capacity	Evaluation plan follows a 5-year cycle and is updated annually Evaluation plan developed in consultation with most senior managers Evaluation plan includes some assessment of risk Provisions made for reviewing the plan on an ongoing basis	Medium and large evaluations are designed in-house and conducted by external consultants Evaluation staff involved in conducting the evaluation (i.e., contribute to the field work done by consultants) Smaller studies are conducted in-house by evaluation staff Consulting work produced is generally considered high quality	Evaluators share their progress and other information with their colleagues at regular staff meetings Evaluators are generally aware of new developments in policy and strategic planning Knowledge management standards have been developed and are generally followed within the unit	Evaluators make use of external supports to evaluation such as professional associations, published standards, and so on. Evaluators keep themselves informed of new developments in external organizations of interest (i.e., universities, other departments and agencies, central agencies) External stakeholders involved in the delivery of federal government programs are kept informed	Evaluators communicate with their program clients through ongoing, formal mechanisms Deputy Head receives regular reports about evaluation activities but is not directly involved in evaluation Evaluation unit located in close proximity to some key organizational areas such as policy development, strategic planning and performance measurement

(continued)

Table 3. (continued)

Level	Evaluation plan	Use of consultants	Information sharing	External supports	Organizational linkages
Developing capacity	Evaluation plan exists for 1 or 2 years Evaluation plan developed in consultation with program managers or senior managers No consideration of risk in planning evaluation schedule	Virtually all evaluations are designed and conducted by external consultants Evaluation staff is mainly involved in managing contracts and overseeing the work done with little substantive input Varying quality levels in work produced by consultants	Evaluators share their progress with their supervisors and other staff members in a sporadic, informal manner Evaluators are not aware of new developments in policy and strategic planning Knowledge management standards exist but are not followed	Evaluators have access to basic external supports, such as a professional association or published standards but do not often make use of them Evaluators do not generally liaise with external organizations or experts External stakeholders are informed of evaluation reports once they are published (where applicable)	Evaluators communicate with program clients on specific issues related to projects Deputy Head is made aware of evaluation findings only through formal requirements (e.g., internal evaluation committee) Evaluation unit is removed (either physically or structurally) from key organizational areas such as policy development, strategic planning, and performance measurement
Low capacity	No evaluation plan is developed by staff Evaluation projects occur on ad hoc basis	All evaluation work is contracted out Evaluation staff manage contract work done by consultants with no substantive input Some problems with the quality of the work produced by consultants	Evaluators do not typically share progress with other staff members Evaluators are not aware of new developments in policy and strategic planning Knowledge management standards do not exist within the unit	Evaluators do not make use of basic external supports, such as professional associations and published standards Evaluators do not liaise with external organizations or experts External stakeholders are not informed of evaluation activities	Evaluators communicate infrequently with program clients Deputy head tends to delegate responsibility for evaluation Evaluation unit is removed (either physically or structurally) from key organizational areas such as policy development, strategic planning, and performance measurement

members spend a considerable amount of time sharing with their colleagues information related to their progress on certain files or on general project management issues. Evaluators in some organizations also establish linkages with external supports such as professional associations, program stakeholders, and other organizations likely to provide assistance, such as the Treasury Board Secretariat. In addition, evaluation staff may establish linkages within their own organizations through formal or informal ties in order to remain informed regarding policy decisions likely to affect their work and to better share the results of evaluations conducted by members of the unit.

The fourth dimension is the first one included under the overarching ''capacity to use'' evaluation category and reflects a less operational perspective (see Table 4). It focuses on Evaluation Literacy within the organization and is divided into two subdimensions: Involvement in evaluation and results-management orientation. Involvement in evaluation is the participation of program staff and other stakeholders in the evaluation process. Participatory evaluation theory holds that the greater the involvement of stakeholders in all phases of an evaluation, the greater the instrumental, conceptual, and process use of evaluation (Cousins & Chouinard, 2012). Therefore, in order to build evaluation capacity, organizations must pay attention to the involvement of staff members in the evaluation process. Results-management orientation refers to the larger organizational culture and the messages that are brought forward by senior managers. A results-management orientation can be manifested through the development of results chains for programs and the implementation of performance measurement strategies.

The fifth dimension (Table 5) focuses on the integration of evaluation information with organizational decision-making processes. At the outset, management processes such as the development of Memoranda to Cabinet (MC) and Treasury Board (TB) submissions should consider evaluation in order to ensure that sufficient resources are provided for the eventual evaluation of new initiatives. At the final stage of the evaluation process, the findings and recommendations made in an evaluation study should be clearly linked to budget allocation and other high-level organizational and policy decisions. An organization with exemplary capacity searches out evaluation information in its decision-making process and relies on this information on an ongoing basis.

Finally, the sixth dimension, Learning Benefits, addresses the types of uses that can be made of evaluation information within an organization (see Table 6). At a more operational level, the evaluation findings can be used as a basis for action and change through the implementation of evaluation recommendations (instrumental use). The evaluation findings can also have an impact on stakeholders' understanding of, and attitudes toward, a program by clarifying certain operational aspects or by highlighting specific program results (conceptual use). At a broader level, participation of organizational members in the evaluation process can result in behavioral or cognitive changes within these individuals based on their exposure to evaluation (process use).

Organizational Variation

The specific elements included in each level of evaluation capacity (i.e., the bullets within each cell in the matrix) varied somewhat over the course of the development of the framework. Elements were added as necessary to increase the clarity of the description and to differentiate between levels. It is probable that there would be within-organization variation in the profile of any given organization. The purpose of the framework is to describe organizational evaluation capacity and to provide organizations with a means of generating information that can be used to identify the particular elements that require improvement in order to reach desired levels of evaluation capacity. Therefore, variation that may be observed within an organization between its levels of evaluation capacity on different subdimensions is to be expected, and may facilitate discussion of next steps for the organization in terms of developing its capacity.

Table 4. Capacity to Use Evaluation, Dimension 4: Evaluation Literacy.

Level	Involvement in evaluation	Results-management orientation
Exemplary capacity	Organizational staff members generally understand the purpose of evaluation and how it supports the organizational mandate (e.g., staff members understand results-based management principles and practices) Program managers and other staff members are closely involved at key points in the evaluation process (e.g., review identified issues and provide feedback, facilitate data collection opportunities, review draft evaluation reports)	Senior managers promote a results-management orientation for the entire organization and make it a priority by providing time and resources Organizational members share clear ideas about organizational purpose and goals through formal and informal mechanisms (e.g., strategic planning sessions, retreats, regular meetings, brown bag lunch sessions) All programs have a clear results chain (i.e., logic model) Program managers take the lead for the development and implementation of performance measurement strategies; evaluators provide technical expertise when needed
Intermediate capacity	Organizational staff members are familiar with the general principles of evaluation and how it can help them in their work (e.g., they understand the difference between evaluation and audit) Program managers are involved in evaluation projects (e.g., sit on Evaluation Steering or Advisory Committees) and provide program-related feedback on report drafts	Organizational outcomes or expected results are only outlined in official documentation but are not included in communications from senior managers Organizational members share clear ideas about organizational purpose and goals through formal mechanisms such as strategic planning sessions and meetings Some programs have a clear results chain (i.e., logic model) Program managers work with evaluators in the development and implementation of performance measurement strategies, but evaluators lead these projects
Developing capacity	Little awareness of evaluation or its purpose within larger organizational context Little involvement from program staff and managers (i.e., brief comments on draft evaluation reports)	Organizational outcomes or expected results are not articulated clearly for all organizational members; most are not aware of results management principles and practices Some programs are engaged in developing results chains such as logic models Program managers not involved in the development or implementation of performance measurement strategies; evaluators conduct these processes with little input from programs
Low capacity	No discernible awareness of evaluation or its purpose within larger organizational context No involvement of program staff and managers	Organizational outcomes or expected results have not been developed Programs do not have results chains such as logic models The organization does not support the development of performance measurement strategies

Table 5. Capacity to Use Evaluation, Dimension 5: Integration With Organizational Decision Making.

Level	Management processes	Decision support
Exemplary capacity	Program and policy staff integrate evaluation into other areas of their work (e.g., they routinely request the involvement of evaluators in management processes such as the preparation of Memoranda to Cabinet and Treasury Board Submissions)	Evaluation findings and recommendations considered in budget allocation and other high-level organizational and policy decisions Demand for evaluation evidence originates from all levels of the organization
Intermediate capacity	Program and policy staff are aware of the evaluation services that can be provided and sometimes contact evaluation staff for advice	Evaluation findings and recommendations usually considered in program management decisions and some policy decisions Program managers are interested in and use evaluation as a management support tool (i.e., evaluation as provider of ongoing management information)
Developing capacity	Evaluation unit operates separately from program units and is not generally involved in management processes; program and policy staff unaware of the potential contributions of evaluation staff	Little consideration of evaluation findings and recommendations in organizational and policy decisions No specific demand for evaluation services other than to meet the requirements of central agencies
Low capacity	Evaluation unit does not involve or inform program units of its activities	Evaluation findings and recommendations are not used in organizational and policy decisions No demand for evaluation services exists within the organization

Validation Exercise

In the third phase of the study, four different federal government departments were asked to assess their organization based on the dimensions and subdimensions developed in the first two phases. The purpose of this exercise was 2-fold: First, it helped us identify missing elements and verify the clarity of the wording used; second, it enabled us to test the framework as a complete organizational self-assessment of evaluation capacity (as reported in Bourgeois & Cousins, 2008). Overall, data obtained in this phase of the study validated the framework: The capacity levels of the participating organizations that had been identified by the experts consulted in the first phase of the study were consistent with the results produced through the application of the framework. Further, participants felt that the framework enabled them to document specific resource requirements based on their vision of evaluation in their respective organizations, and provided them with a guide for measuring the success of their ECB activities. Participants expressed an interest in obtaining a final version of the framework for use in their organizations, and stated that a self-assessment tool based on a more quantitative measure of organizational evaluation capacity could be useful. A longer term recommendation for both research and practice, therefore, is the transformation of the framework into an instrument for assessing such capacity. This work is currently underway. Broader methodological issues to be addressed include the instrument's reliability, as well as the weightings of subdimensions based on their importance to the organization, as was done by Cousins, Aubry, Smith-Fowler, and Smith (2004).

This last element is important because the structure of the framework assumes that the dimensions and subdimensions are equally weighted. In practice, this may not be true. One can imagine, for example, an organization at an early stage of evaluation capacity development being more

Table 6. Capacity to Use Evaluation, Dimension 6: Learning Benefits.

Level	Instrumental/conceptual use	Process use
Exemplary capacity	Evaluation findings are used consistently as a basis for action and change (i.e., evaluation recommendations are appropriate and implemented in a timely manner) Evaluation findings and reports often have an impact on stakeholders' understanding and attitudes about programs	Strong evidence of behavioral or cognitive changes occurring in stakeholders by virtue of their proximity to evaluation Evidence that organizational members routinely apply evaluation logic to other organizational issues (e.g., by questioning basic assumptions and using systematic inquiry to identify solutions to organizational problems) Formal or informal processes to share lessons learned during evaluations are in place and involve the entire organization (e.g., seminars, brown-bag lunch sessions, brochures on recent studies)
Intermediate capacity	Evaluation findings are sometimes used as a basis for action and change (i.e., evaluation recommendations are sometimes implemented) Evaluation findings and reports can have an impact on stakeholders' understanding and attitudes about programs	Some evidence of behavioral or cognitive changes occurring in stakeholders by virtue of their proximity to evaluation Evidence that organizational members sometimes apply evaluation logic to other organizational issues (e.g., using an inquiry-based process to identify organizational issues and their solutions) Lessons learned through evaluations are shared with organizational members directly involved with the program (e.g., letters, formal presentation of report)
Developing capacity	Evaluation findings are rarely used as a basis for action and change (i.e., evaluation recommendations are usually not implemented) Evaluation findings and reports rarely have an impact on stakeholders' understanding and attitudes about programs	Little evidence of behavioral or cognitive changes occurring in stakeholders by virtue of their proximity to evaluation No evidence that stakeholders apply evaluation logic to other organizational issues Evaluation projects are not shared once completed; evaluation reports disseminated only to internal evaluation committee
Low capacity	Evaluation findings are never used as a basis for action and change (i.e., evaluation recommendations do not usually make their way to those with the ability to act upon them) Evaluation findings and reports do not have an impact on stakeholders' understanding and attitudes about programs (because they are rarely aware of the evaluation)	No evidence of behavioral or cognitive changes occurring in stakeholders by virtue of their proximity to evaluation No evidence that stakeholders apply evaluation logic to other organizational issues Evaluation projects are not shared once completed; evaluation reports not disseminated outside of the evaluation unit

interested in focusing on the capacity to do evaluation rather than the capacity to use it. Once evaluation systems and functions are developed, implemented, and to some preliminary degree, institutionalized, we might expect more pronounced interest in improving organizational capacity to use evaluation.

Conclusion

Although much has been published on ECB, the actual characteristics and attributes of evaluation capacity itself have rarely been defined and described based on empirical data. This study concluded that evaluation capacity in Canadian federal government departments and agencies can be described functionally and operationally through six main dimensions that reflect an organization's ability to do evaluation and use evaluation: human resources, organizational resources, evaluation planning and activities, evaluation literacy, organizational decision making, and learning benefits. Each of these dimensions was broken down into a number of subdimensions, with evaluation capacity being assessed using four levels: low, developing, intermediate, and exemplary. Although the Leithwood and Montgomery (1987) approach permits variation across dimensions in terms of the number of levels, interview respondents felt that a common structure across all dimensions would provide a clearer picture of evaluation capacity and make the resulting framework more useful. The number of subdimensions varies from one dimension to the next, in an attempt to develop a comprehensive framework of evaluation capacity.

The study yields important clues as to what a theory of change of evaluation capacity might look like, by suggesting that organizational development in this domain does not occur in linear fashion across a series of elements or dimensions. In addition, the framework enhances our understanding of the potential impacts of targeted organizational improvement initiatives by showing the steps required to move between levels of capacity. These lessons extend well beyond a discussion of organizational evaluation capacity.

Continuing research may focus on expanding the scope of the framework to other types of organizations or government organizations in different jurisdictions and contexts. It seems likely that the dimensions and subdimensions identified here would generalize well, given the commonalities in application of measurement and evaluation systems in governance frameworks that embrace RBM and new public management. It would be instructive to examine the applicability of the framework to the voluntary sector. Preliminary findings from other research on evaluation capacity suggest that governmental and nongovernmental (voluntary sector) organizations differ significantly in their capacity to conduct and use evaluation. Despite higher ratings of capacity to *do* evaluation in government settings, the capacity to use it was seen as lower than in the voluntary sector (Cousins, Goh, Elliott, & Aubry, 2008). This finding may be at least partly attributable to the fact that many voluntary organizations, due to their smaller scale, would directly assign managers and decision makers to evaluation roles, rather than having a self-standing evaluation unit or function. One can imagine process use being higher in such instances, since evaluation would be more integrated into the organizational decision-making function. In any case, additional research is required to determine the applicability and relevance of the framework across organizational sectors.

The context within which this study was undertaken poses certain limitations to the interpretation of its findings. The focus on Canadian federal government organizations, in particular, generated findings that are applicable to these organizations but may not be appropriate in other contexts. Further, the small number of participating organizations has resulted in some data loss, especially in the case of the low capacity organization, in which a suitable evaluation user could not be found who might offer a balancing perspective to the assessment of the Head of Evaluation and senior evaluator.

As discussed previously, the major practical implication of this study is the potential transformation of the proposed framework into an instrument for assessing evaluation capacity in government organizations. Such an instrument could serve as a valuable self-reflection tool within organizations, generating serious discussion and debate about evaluation capacity, and optimal strategies for improving it. As is the case with innovation profiles, the use of such a tool would best be restricted to formative, developmental challenges within the organization, as opposed to more summative, accountability-oriented demands. Ongoing research on the use of such a tool and its associated

benefits and drawbacks would further knowledge development in this area, and represents another valuable avenue to pursue.

Acknowledgments

Authors' Notes

Declaration of Conflicting Interests

Funding

References

Argyris, C., and Schon, D. (1978). *Organizational Learning: A Theory of Action Perspective*. Reading, MA: Addison-Wesley.

Berends, H., Boersma, K., & Weggeman, M. (2003). The structuration of organizational learning. *Human Relations*, *56*, 1035–1056.

Bourgeois, I., & Cousins, J. B. (2008). Informing evaluation capacity building through profiling organizational capacity for evaluation: An empirical examination of four Canadian federal government organizations. *Canadian Journal of Program Evaluation*, *23*, 127–146.

Boyle, R., Lemaire, D., & Rist, R. C. (1999). Introduction: Building evaluation capacity. In R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity: Lessons from practice* (pp. 1–19). New Brunswick, NJ: Transaction.

Compton, D. W., Baizerman, M., & Stockdill, S. H. (2002). The art, craft and science of evaluation capacity building. *New Directions for Evaluation*, *93*, 47–61.

Compton, D. W., & MacDonald, G. (2008). Using evaluation capacity building (ECB) to interpret evaluation strategy and practice in the United States National Tobacco Control Program (NTCP): A preliminary study. *Canadian Journal of Program Evaluation*, *23*, 199–224.

Cousins, J. B., Aubry, T. D., Smith Fowler, H., & Smith, M. (2004). Using key component profiles for the evaluation of program implementation in intensive mental health case management. *Evaluation and Program Planning*, *27*, 1–23.

Cousins, J. B., & Chouinard, J. A. (2012). *Participatory evaluation up close: An integration of research-based knowledge*. Charlotte, NC: Information Age.

Cousins, J. B., Goh, S. C., Clark, S., & Lee, L. E. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, *19*, 99–141.

Cousins, J. B., Goh, S., Elliot, C., & Aubry, T. (2008, May). *Government and voluntary sector differences in organizational capacity to do and use evaluation*. Paper presented at the annual meeting of the Canadian Evaluation Society, Québec, Canada.

Diaz-Puente, J., Yague, J. L., & Afonso, A. (2008). Building evaluation capacity in spain: A case study of rural development and empowerment in the european union. *Evaluation Review*, *32*, 478–506.

Feinstein, O., & Zapico-Goni, E. (2011). *Evaluation of government performance and public policies in Spain* (Evaluation Capacity Development Working Paper Series, 22). Washington, DC: Independent Evaluation Group, World Bank.

Garcia-Iriarte, E., Suarez-Balcazar, Y., Taylor-Ritzler, T., & Luna, M. (2011). A catalyst-forchange approach to evaluation capacity building. *American Journal of Evaluation*, *32*, 168–182.

Gibbs, D., Napp, D., Jolly, D., Westover, B., & Uhl, G. (2002). Increasing evaluation capacity within community-based HIV prevention programs. *Evaluation and Program Planning*, *25*, 261–269.

Goh, S. (2003). Improving organizational learning capability: Lessons from two case studies. *Learning Organization*, *10*, 216–227.

Goh, S., & Richards, G. (1997). Benchmarking the learning capability of organizations. *European Management Journal*, *15*, 575–583.

Huffman, D., Thomas, K., & Lawrenz, F. (2008). A collaborative immersion approach to evaluation capacity building. *American Journal of Evaluation*, *29*, 358–368.

Jorjani, H. (1998). Demystifying results-based performance measurement. *Canadian Journal of Program Evaluation*, *13*, 61–95.

Katz, S., Sutherland, S., & Earl, L. (2002). Developing an evaluation habit of mind. *Canadian Journal of Program Evaluation*, *17*, 103–119.

King, J. A. (2002). Building the evaluation capacity of a school district. *New Directions for Evaluation*, *99*, 63–80.

Lawrenz, F., Thomas, K., Huffman, D., & Covington Clarkson, L. (2008). Evaluation capacity building in the schools: Administrator-led and teacher-led perspectives. *Canadian Journal of Program Evaluation*, *23*, 61–82.

Leithwood, K. A., & Montgomery, D. J. (1987). *Improving classroom practice using innovation profiles*. Toronto, Canada: OISE.

Mackay, K. (2002). The World Bank's ECB experience. *New Directions for Evaluation*, *99*, 81–99.

Mark, K., & Pfeiffer, J. R. (2011). *Evaluation of government performance and public policies in Spain* (Evaluation Capacity Development Working Paper Series, 26). Washington, DC: Independent Evaluation Group, World Bank.

Mathison, S. (2011). Internal evaluation, historically speaking. *New Directions for Evaluation*, *132*, 13–23.

Mayne, J. (2009). Building an evaluative culture: The key to effective evaluation and results management. *Canadian Journal of Program Evaluation*, *24*, 1–30.

Milstein, B., Chapel, T. J., Wetterhall, S. F., & Cotton, D. A. (2002). *New Directions for Evaluation*, *99*, 27–46.

Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity–results and implications of a Danish study. *American Journal of Evaluation*, *32*, 324–344.

Owen, J., & Lambert, K. (1995). Roles for evaluators in learning organizations. *Evaluation*, *1*, 237–250.

Popper, M., & Lipshitz, R. (2000). Organizational learning: Mechanisms, culture, and feasibility. *Management Learning*, *31*, 181–196.

Preskill, H., & Boyle, S. (2008a). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, *29*, 443–459.

Preskill, H., & Boyle, S. (2008b). Insights into evaluation capacity building: Motivations, strategies, outcomes, and lessons learned. *Canadian Journal of Program Evaluation*, *23*, 147–174.

Preskill, H., & Torres, R. T. (1999). *Evaluative inquiry for learning in organizations*. Thousand Oaks, CA: Sage.

Preskill, H., & Torres, R. T. (2000). The learning dimension of evaluation use. *New Directions for Evaluation*, *88*, 25–36.

Rowe, W. E., & Jacobs, N. F. (1998). Principles and practices of organizationally integrated evaluation. *Canadian Journal of Program Evaluation*, *13*, 115–138.

Sanders, J. R. (2002). Presidential address: On mainstreaming evaluation. *American Journal of Evaluation*, *23*, 253–259.

Stockdill, S. H., Baizerman, M., & Compton, D. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. *New Directions for Evaluation*, *99*, 7–25.

Sutherland, S. (2004). Creating a culture of data use for continuous improvement: A case study of an Edison project school. *American Journal of Evaluation*, *25*, 277–293.

Talbot, C. (2010). *Evaluation of government performance and public policies in Spain* (Evaluation Capacity Development Working Paper Series, 24). Washington, DC: Independent Evaluation Group, World Bank.

Taylor-Powell, E., & Boyd, H. H. (2008). Evaluation capacity building in complex organizations. *New Directions for Evaluation*, *120*, 55–69.

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation*, *28*, 45–59.

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

Toulemonde, J. (1999). Incentives, constraints and culture-building as instruments for the development of evaluation demand. In R. Boyle & D. Lemaire (Eds.), *Building effective evaluation capacity: Lessons from practice* (pp. 153–177). New Brunswick, NJ: Transaction.

Treasury Board Secretariat. (2010). *Supporting effective evaluations: A guide for developing performance measurement strategies* (draft). Ottawa, Ontario, Canada: Government of Canada.

Volkov, B. (2008). A bumpy journey to evaluation capacity: A case study of evaluation capacity building in a private foundation. *Canadian Journal of Program Evaluation*, *23*, 175–197.