

MEASURING ORGANIZATIONAL EVALUATION CAPACITY IN THE CANADIAN FEDERAL GOVERNMENT

Isabelle Bourgeois
École nationale d'administration publique
Gatineau, Québec

Eleanor Toews
Graduate Diploma in Public Policy and Program Evaluation
Carleton University
Ottawa, Ontario

Jane Whynot
Whynot and Associates
Ottawa, Ontario

Mary Kay Lamarche
Master of Arts in Public Ethics
Saint Paul University
Ottawa, Ontario

Abstract: The development of organizational evaluation capacity has emerged in recent years as one mechanism through which evaluators can extend their influence and foster evaluation utilization. However, organizational evaluation capacity is not always easy to define, and internal evaluators sometimes struggle with the identification of concrete activities that might increase their organization's evaluation capacity. This article describes an organizational self-assessment instrument developed for Canadian federal government organizations. The instrument is presented and described, and further details regarding its use and next steps for this area of evaluation research are also provided.

Résumé : Le renforcement de la capacité organisationnelle en évaluation est l'un des moyens disponibles aux évaluateurs pour augmenter leur influence et promouvoir l'utilisation des résultats d'évaluation. Cependant, il peut s'avérer difficile de bien cerner la capacité organisationnelle en évaluation et d'identifier les activités concrètes qui peuvent l'améliorer. Cet article décrit un instrument d'auto-évaluation de la capacité organisationnelle en évaluation, élaboré pour les organismes du gouvernement fédéral canadien. L'instrument y est présenté, et des détails sup-

Corresponding author: Isabelle Bourgeois, École nationale d'administration publique, Gatineau, QC; <isabelle.bourgeois@enap.ca>

plémentaires au sujet de son utilisation et des prochaines étapes dans ce volet de recherche sont fournis.

INTRODUCTION

■ Evaluation practice and use in the Canadian federal government have evolved over time in response to contextual changes and political interests. The revisions made to the Treasury Board *Policy on Evaluation* in 2009, among other drivers, have contributed to an increase in evaluation activity in many, if not most, departments and agencies. The introduction of a requirement related to the evaluation of all departmental program spending on a five-year cycle, for instance, has meant that organizations have had to invest new resources in their evaluation functions. The increased demand for quality evaluation studies through the implementation of various expenditure review mechanisms has also led to additional departmental investments in evaluation. Given these investments, it could be expected that the quantity of evaluations produced, as well as their quality and use, would increase over the course of several years and bolster the influence of evaluation on organizational decision-making. However, the budgetary reductions implemented in early 2012 have resulted in fewer resources being allocated to all government operations, including departmental evaluation functions. In light of increased demand and requirements coupled with fewer financial and human resources, the managers responsible for departmental evaluation functions (typically called Heads of Evaluation in federal organizations) find it increasingly difficult to deliver on their professional commitments.

Enhancing the organizational evaluation capacity of federal government departments and agencies offers one potential solution to decreasing evaluation resources. Evaluation capacity building, or ECB, is defined as “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” (Stockdill, Baizerman, & Compton, 2002, p. 8). In other words, the development of an organization’s evaluation capacity enables the institutionalization of evaluation practice and results in increased evaluation use and organizational learning. It is theorized here and elsewhere (e.g., see Boyle, Lemaire, & Rist, 1999; Cousins, Goh, Clark, & Lee, 2004; Sanders, 2002; Stockdill et al., 2002) that increased organizational

evaluation capacity leads to the production of quality evaluations in a more efficient and timely manner, thus enabling organizations to meet their external reporting requirements (in the case of most Canadian federal government departments and agencies, this implies meeting the requirements of the Treasury Board Policy on Evaluation) and to use evaluation for organizational decision making and continued learning.

Although the evaluation literature is replete with case examples of ECB activities occurring in various types of organizations, very little attention had been paid to the actual dimensions of evaluation capacity and how these might be described and enacted in practice until the development of an organizational framework for evaluation capacity by Bourgeois and Cousins in 2008 (most recently published in its entirety in 2013). The framework, which is based on empirical research conducted on Canadian federal government organizations, identifies six dimensions of evaluation capacity and describes in detail how each of these develop over time. The research work summarized in this article sought to build on the framework and presents its operationalization into an organizational self-assessment instrument. We developed this instrument in response to requests from the federal community; our objective was not to produce a standardized instrument for the purposes of interorganizational comparison. Rather, we attempted to develop an instrument that would enable evaluation managers and other organizational leaders to accurately assess their organization's evaluation capacity and to use this information to identify key areas for improvement and capacity building. A literature review summarizing recent efforts in evaluation capacity building as well as the measurement of evaluation capacity is presented to contextualize our work; a brief discussion focusing on evaluation in the Canadian federal government will then be presented. The specific methods used to develop and validate the instrument are then provided, followed by a presentation of the instrument, a discussion of intended users and uses of the instrument, and next steps.

ORGANIZATIONAL EVALUATION CAPACITY: EXISTING FRAMEWORKS, INSTRUMENTS, AND TOOLS

In recent years, researchers interested in ECB have focused on the development of frameworks and measurement instruments, in addition to the continued description of specific ECB initiatives found in earlier publications on the topic. Although highly informative, these case narratives (e.g., see case narratives by Adams & Dickinson,

2010; Bourgeois, Hart, Townsend, and Gagné, 2011; Díaz-Puente, Yagüe, & Afonso, 2008; Karlsson, Beijer, Eriksson, & Leissner, 2008; Naccarella et al., 2007) often do not offer an empirical knowledge base from which general conclusions can be drawn (Bourgeois, Chouinard, & Cousins, 2008, among others). There have been recent calls for empirically based models of ECB that enable a better understanding of its various components. For example, Taylor-Powell and Boyd (2008) outline a three-component framework meant to structure the various strategies employed in evaluation capacity building in the area of extension education. Their framework includes dimensions such as professional development, resources and supports, and organizational environment—many of these are reflected in our own framework of EC (outlined below). Along the same lines, Preskill and Boyle (2008) illustrate a set of factors that may influence the initiation, design, implementation, and impact that ECB activities and processes have on sustainable evaluation practice. This model focuses on evaluation knowledge, skills, and attitudes, as well as sustainable evaluation practice (described through several elements, such as a strategic plan for evaluation, resources dedicated to evaluation, leadership, culture, etc.). The model is quite broad and appears to focus less on the tangible aspects of evaluation implementation and more on the processual components of ECB. In this way, the model summarizes all of the various elements that may be considered in a study of ECB or in the development of an ECB strategy; however, its level of complexity offers limited support to evaluation managers and practitioners seeking to improve their own organizational practice and evaluation use.

Labin, Duffy, Meyers, Wandersman, and Lesesne (2012) offer a different conceptual framework, based on a systematic review of the ECB literature, named the Integrative ECB Model. This model features some of the same elements as the ones identified previously; however, the model in this instance seeks to provide a comprehensive view of the needs for ECB (reasons, motivations, objectives, resources, and strengths), the activities involved in ECB and its evaluation, and the results or outcomes associated with ECB at the individual, organizational, and program levels. The Integrative ECB Model therefore focuses more on actual capacity building rather than on the components of evaluation capacity, although these are implied in the model and presented in the literature review.

The issue of empirical validation has also emerged in response to the development of various conceptual frameworks and represents a significant evolution in the area of ECB research. Cousins et al.

(2008) present one of the first attempts to validate key elements of ECB through a pan-Canadian survey of internal evaluators, focusing on organizational capacity to do and use evaluation. The questionnaire used in this study was based on a conceptual framework that included components found in other models, such as organizational support structures and organizational learning capacity. The survey enabled Cousins et al. to describe the perceptions of the respondents regarding evaluation capacity in their own organizations as well as the factors that influence ECB, use of evaluation results and processes for decision making. Other authors, such as Tseng (2011), Nielsen, Lemire, and Skov (2011), and Taylor-Ritzler et al. (2013) have moved this empirical work further, through the development of quantitative measurement scales used to validate their own conceptual models of individual or organizational evaluation capacity. In all three of these cases, the conceptual models were validated through individual-level questionnaires based on the key elements or factors related to EC. These elements included, among others, the objectives of evaluation, its structures and processes within the organization, individual competencies, and technology. In all three cases, the conceptual models and ensuing instruments focus on a specific sector of activity, from Danish municipalities (as in Nielsen et al., 2011) to Chicago-based nonprofit organizations (as in Taylor-Ritzler et al., 2013). The application spheres of each of these models may be construed as overly specific; in our opinion, the availability and diversity of such models outline the need for adaptation rather than generalization. We therefore position our own work, conducted with Canadian federal government organizations, as a natural extension of the work already conducted elsewhere, rather than a duplication of efforts. Although all of the models identified above share some fundamental characteristics or elements, each of them is uniquely suited to its particular organizational environment, as is our own.

The main difference between our model and instrument and those outlined earlier is intended use. Whereas other authors have focused on validating their conceptual models of EC/ECB in the spirit of furthering research on the topic, our intention was always one of contributing to evaluation practice by producing a useful instrument to evaluation managers. Rather than following in the footsteps of other researchers, we sought to engage practitioners in the instrument development and validation processes, to develop a tangible management tool that can be used independently of the research team and for multiple purposes. In other words, the organizational self-assessment instrument presented here is meant to enable bet-

ter evaluation use and organizational learning from the ground up. More specifically, by measuring current organizational evaluation capacity and putting in place specific interventions to further develop evaluation capacity over time, Heads of Evaluation will be able to improve both the quality and quantity of evaluations and to foster the development of organization-wide evaluation literacy.

CONTEXT: EVALUATION IN THE FEDERAL GOVERNMENT

As early as the 1960s, the linkage between program effectiveness and good public sector management was recognized by the Canadian federal government. The necessity for formal procedures to measure the effectiveness of public programs resulted in the introduction in 1977 of the first formal Government of Canada evaluation policy in which a platform for examining program effectiveness and efficiency was introduced. Evaluation findings and recommendations were to be used by deputy heads to make more informed decisions on management and resourcing, to be accountable for the programs for which they were responsible, and to provide quality advice to ministers (Segsworth, 2005).

In the three decades following this inaugural evaluation policy, several intervening policies have been implemented across government along with significant changes in evaluation practice (Shepherd, 2012). In 2001, the previous evaluation policy aligned the evaluation function with the practices of New Public Management, through an increased focus on results-based management. The most recent iteration of the Treasury Board *Policy on Evaluation* (Government of Canada, Treasury Board Secretariat, 2009) emphasizes evaluation use within the expenditure management system through the provision of credible and neutral information. Evaluation is one of the smallest functional communities in the Canadian federal government (Government of Canada, Treasury Board Secretariat, 2011), but its relative importance within the expenditure management system elevates the function's visibility and importance.

A significant addition to the most recent *Policy on Evaluation* (Government of Canada, Treasury Board Secretariat, 2009) is a requirement for a neutral assessment of a department's evaluation function. Section 7.1 of the policy states:

Deputy heads are responsible for monitoring compliance with this policy in their departments to ensure its effec-

tive implementation. They are responsible for ensuring that a neutral assessment of their departmental evaluation function is conducted at a minimum of once every five years. (Government of Canada, Treasury Board Secretariat, 2009)

Underpinning the provision of credible and neutral information is the functional capacity of individual departments and agencies. Under the current Management Accountability Framework exercise, department and agency evaluation practices are assessed on an annual basis by the Treasury Board Secretariat (TBS); the resulting information is used to provide insight into the progress of departments and agencies both in implementing the policy and in their organizational capacity to conduct and to use evaluation. This effort is complemented by the production of an annual report on the Health of the Evaluation Function that focuses on the overall strength of the evaluation function across federal government organizations. This report is meant to highlight potential improvements to existing tools and mechanisms, to better enable departments and agencies to meet policy objectives. Both the organizational-level information and the data collected for the function at large through these exercises are used by the TBS to develop guidance and other tools to support the evaluation community.

The organizational EC self-assessment instrument presented in this article provides a comprehensive mechanism to complement the federally driven efforts noted above. The instrument is unique in that it provides a summary of an organization's current evaluation capacity, identifies evaluation capacity building needs, and provides a framework to create an established evaluation function. The information generated by departments and agencies through the use of the instrument can be used by stakeholders to identify priorities, benchmark organizational practices, develop evaluation plans, and address challenges in responding to various policy requirements.

METHODOLOGY: DEVELOPMENT OF THE ORGANIZATIONAL EC SELF-ASSESSMENT INSTRUMENT

Conceptual Framework

As described earlier, the framework of organizational evaluation capacity developed by Bourgeois and Cousins (2008; 2013) constitutes the conceptual centrepiece of the self-assessment instrument

presented in this article. This framework is specially focused on the evaluation capacity of Canadian federal government organizations and describes an organization's capacity to both produce and use evaluation through six dimensions: human resources, organizational resources, evaluation planning and activities, evaluation literacy, organizational decision-making, and learning benefits. Each of these six dimensions is further broken down into subdimensions (19 in total). The subdimensions are described qualitatively in the framework at four levels of capacity: low, developing, intermediate, and exemplary.¹

Instrument Development

Although the framework can be used to some extent to diagnose an organization's evaluation capacity, it is not ideally suited to this task because of its length and level of detail and can therefore be considered more of a reference piece than a useable tool. To fill this gap, and based on a number of requests from Heads of Evaluation, the EC instrument was developed by operationalizing the framework's six dimensions as well as their respective subdimensions. A series of items (i.e., statements outlining a specific aspect of evaluation capacity) were developed through this operationalization process to measure each dimension and subdimension (see Table 1 for examples of the items). The instrument follows the structure of the framework and retains most of its contents, but organizes it in a way to facilitate organizational assessment. Each item is assessed using a four-point Likert scale focusing on the extent to which the statement reflects current organizational practice.

Instrument Validation Study

The review and pilot process used to validate the instrument followed three main steps. The first phase involved a review of the draft instrument by three subject-matter experts knowledgeable in organizational evaluation capacity building. The experts were asked to comment on the face validity of the instrument and to identify any gaps. Overall, the feedback received was positive, and suggestions for specific items focused on improving clarity. The major change that was implemented following the review was the inclusion of an importance scale to the instrument, in addition to the measurement scale described above. Adding the importance scale enabled the respondents to gauge the applicability of each item to their organization (in

other words, some items do not apply to certain organizations, based on their size, mandate, etc.). This improved the specificity of the score obtained and provided the respondents with a clearer picture of their organization's evaluation capacity by removing extraneous items from the score calculation.

The second phase involved the collection of feedback from a broader audience of potential users through three presentations made to groups of interested evaluators (two of the presentations were made at events organized by the Canadian Evaluation Society). Each presentation provided background information through an overview of the underpinning framework and included a live demonstration of the instrument. In all three cases, audience members were encouraged to provide feedback on the instrument; in one instance, a specific survey tool was developed to capture audience feedback after the presentation. The feedback once again was supportive overall and enabled further minor modifications to the instrument.

The third phase was designed as a pilot of the instrument. Five government organizations participated in the pilot study. Each organization was provided with a copy of the draft instrument as well as a set of questions focusing on ease of use, clarity, who completed and should complete the instrument, and how the information generated by the instrument was used or could be used in the future. The five organizations either provided written feedback for each of these questions or were interviewed by a member of the research team. The data obtained from the pilot participants were collated and analyzed to produce a final set of modifications to the instrument that were mostly formatting and typographical. Some pilot participants representing small organizations reported difficulties in interpreting the meaning of a few items; in these cases, the wording and applicability of the items were reviewed.

RESULTS: DESCRIPTION OF THE INSTRUMENT

With an end objective of developing a truly useable instrument, two design features guided the overall development process: ease of use and flexibility. To accomplish this task, a Microsoft Excel workbook comprising six worksheets was selected in order to use commonly accessible software and make it adaptable to various organizational contexts and structures. Table 1 provides selected examples of the items for each of the dimensions and subdimensions included in the original framework (Bourgeois and Cousins, 2008, 2013).

Table 1
Selected Items² from the Organizational Evaluation Capacity Self-Assessment Instrument

Dimension/Subdimension	Items
<i>Human Resources</i>	
Staffing	Employee turnover is consistent with overall organizational turnover rate Employee retention initiatives and/or policies (e.g., clear career path, internal recognition) have been implemented to facilitate promotion and employee retention
<i>Organizational Resources</i>	
Budget	Budget costs for a specific evaluation are clearly outlined and based on proposed methodology Adequate, stable resources are available to complete evaluation projects
Ongoing data collection	Program managers and evaluators collaborate to ensure that indicators for program results and corresponding data points are clearly identified Ongoing performance measurement data are available, accessible, and complete at the corporate and program levels
<i>Evaluation planning and activities</i>	
Organizational evaluation planning	Evaluation planning includes departmental risk assessment Organizational evaluation plan (formal or informal) sets out clear, reasonable project completion targets and timelines in relation to resources and departmental priorities
Organizational linkages	Evaluation steering or advisory committees are regularly used to guide the direction and key issues for evaluations Evaluation, research, and policy units coordinate efforts to minimize duplication and leverage research
<i>Evaluation literacy</i>	
Results-management orientation	All programs have a logic model, program theory, or results chain Senior executives provide time and resources for performance measurement and evaluation activities
Involvement/participation	Program managers and/or staff understand the purposes of evaluation and how it is used Program managers and/or staff are included in evaluation steering or advisory committees or other working groups related to evaluation
<i>Organizational decision making</i>	
Management processes	Allocation of evaluation unit budget is transparent and reflects departmental concerns and priorities New program financing includes the cost of the corresponding evaluation, and the funds are transferred to the evaluation unit to manage the evaluation
Decision support	Demand for evaluation services originates from all levels of the organization (i.e., from both program-level staff and senior executives) Evaluation findings and recommendations are routinely sought out and considered in budget allocation exercises and other organizational decisions
<i>Learning benefits</i>	
Instrumental/conceptual use	Evaluation recommendations are implemented in a timely manner Evaluation is used to make decisions about fiscal allocation
Process use	Program managers and/or staff recognize the value of systematic inquiry to identify solutions to organizational problems Program management and/or staff report expanded and/or expedited use of evaluation findings due to their involvement in the evaluation process

Rating Scales

As mentioned above, two rating scales are used to calculate the score for each subdimension: The first four-point scale is used to rate the extent to which the user agrees that the item represents his or her organization's situation or practice, from 4 (*strongly agree*) to 1 (*strongly disagree*). The second four-point scale is used to rate the applicability or importance of each item for the user's organization, from 3 (*very important*) to 0 (*not applicable*). This second scale acts as a multiplier in the scoring calculation for each subdimension. A rating of *not applicable* removes the item from the score entirely to avoid artificially lowering the assessment's results. Combined, these two scales enable the user to tailor the instrument to their organizational context by rating current practice and filtering out items that are deemed irrelevant.

The instrument also enables raters to document the assessment process through integrated comment space. For example, users could include details to verify, people to contact, expected changes, or a description of the current organizational context in cases where the self-assessment is expected to be completed on a periodic, recurring basis. This is an important feature of the instrument that is different from the scales published elsewhere, which tend to be researcher-oriented rather than of use to practitioners in the field.

Scoring

The instrument has been developed to calculate scores automatically for each of the six dimensions. The mean score obtained for each of the subdimensions is provided, as well as the overall mean score for the dimension (as shown in Table 2 below). This level of detail provides the user with specific, targeted information about which evaluation capacity building interventions may be most warranted. In the example given in Table 2, subdimension 4.2 (Results-Management Orientation) shows a lower level of capacity than subdimension 4.1 (Involvement in Evaluation); this information could therefore lead to a greater focus on the first subdimension in terms of ECB activities.

Interpretation

The scores obtained for each of the subdimensions are matched automatically to the corresponding capacity level (low, developing,

intermediate, and established) as defined in the conceptual framework. A qualitative description of organizational evaluation capacity is provided for interpretative purposes, as shown in Table 3 below. When all of the statements included in a subdimension are rated as

Table 2
Example of Scoring Results

4. Evaluation literacy		
<i>Level</i>	<i>4.1 Involvement in evaluation</i>	<i>4.2 Results-management orientation</i>
Mean score by subdimension	2.83	1.67
Average score for dimension	2.25	

Table 3
Example of Interpretation of Results for Evaluation Literacy Dimension

4. Evaluation literacy			
<i>Subdimension</i>	<i>Score</i>	<i>Capacity level</i>	<i>Organizational characteristics</i>
4.1 Involvement in evaluation	2.8	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
4.2 Results-management orientation	1.7	Developing capacity	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 are not involved in the development or implementation of performance measurement frameworks and other performance management activities; evaluators conduct these processes with little input from programs
Average by dimension	2.3	Intermediate capacity	

being not applicable to an organization, the instrument returns a result of *Not Applicable* and does not provide a qualitative description of organizational evaluation capacity for this subdimension. For example, a small organization may not maintain an in-house evaluation team; in this case, the Human Resources subdimension is not applicable to this particular organization and does not appear in the final assessment results.

DISCUSSION: INTENDED USERS AND USES OF THE ORGANIZATIONAL SELF-ASSESSMENT INSTRUMENT

The instrument presented here is designed to be flexible; as such, in theory it can be completed by a third party rather than internal evaluators (e.g., an external evaluator or other contractor could be brought in to complete the instrument through interviews with evaluation stakeholders such as evaluation management /staff, program/policy/initiative management/staff, and senior executives). However, one of the greatest benefits of the self-assessment approach raised by pilot participants is the introspection and consideration of the evaluation function over time that are fostered by the instrument. If completed accurately, the instrument provides a reflection of the capacity of an organization to conduct and use high-quality evaluation projects. The self-assessment instrument, in fact, provides the mechanism through which meaningful discussions can take place within the evaluation unit and focus on continued improvement and learning. Given this, the most likely person responsible for the implementation and completion of the instrument is the Head of Evaluation. This position likely has access to all stakeholders from whom input may be required, and is uniquely responsible for increasing the visibility and importance of the evaluation function within the organization.

The self-assessment component focusing on evaluation use can also enable the evaluation unit to consider how its evaluation processes, outreach, and education may be working. In some cases, the self-reflection process on the capacity to use evaluations may need to be expanded and could include consultations with potential users of evaluation within the department/agency, such as program management/staff and/or department/agency senior management (including Departmental Evaluation Committees).

In terms of process, pilot participants completed the instrument in a variety of ways: some completed it on their own, based on their

organizational experience and knowledge, while others completed it through a group discussion with colleagues; still others consulted organizational members, such as program managers, or senior managers, as part of the exercise. One of the key findings of the pilot is that, in some cases, it would have been beneficial to complete the instrument in a group setting, but finding the time and opportunity to bring various stakeholders together proved to be challenging. Although the instrument was designed to be used by organizations without additional support, it appears that there may be a benefit to developing a workshop session to facilitate a group discussion and complete the instrument through a consensus rather than individual ratings.

Ultimately, who should complete the tool and what process is used is dependent upon how the evaluation function is organized and what works best for each organization. For example, completion of the instrument could depend on the size of the organization, the role of the Head of Evaluation in managing the function, and the structure of the evaluation unit within the broader organization.

Some of the uses identified by pilot study participants include providing data as part of the Management Accountability Framework's Quality and Use of Evaluation section, which reports on the capacity to do and use evaluations in making decisions about the continuation of programs or reallocation of expenditures, as well as in the development of Memoranda to Cabinet and Treasury Board Submissions when designing programs or seeking renewal. In addition, the self-assessment results could also provide input into the annual update to the Departmental Evaluation Plan (DEP) as well as individual evaluation project workplans. The expected use of the instrument, however, lies in the generation of systematic and planned evaluation capacity building exercises. As stated elsewhere (Bourgeois et al., 2008), successful ECB initiatives are planned, thoughtful activities rather than opportunity-driven. The instrument presented here is meant to constitute the first step in articulating an organizational ECB strategy, based on an overall diagnostic exercise.

NEXT STEPS: ADAPTATION TO OTHER ORGANIZATIONAL CONTEXTS

The focus of the self-assessment instrument on the Canadian federal government context enabled the development of a targeted tool that accurately represents the context of a specific set of organizations.

The advantage to such specificity lies in its ease of use for these organizations and in the relevance and appropriateness of the results generated by the instrument. However, the self-assessment instrument cannot readily be used by other types of organizations (e.g., provincial government, nonprofits, etc.). Rather, significant adaptations will be required to both the framework of organizational evaluation capacity and the self-assessment instrument to better meet the needs of other organizations also interested in measuring their existing level of organizational evaluation capacity.

CONCLUSION

The diagnosis of an organization's evaluation capacity, enabled by the use of the self-assessment instrument discussed in this article, constitutes a first step in a longer-term effort toward evaluation capacity building. By targeting specific areas of an organization's evaluation capacity, internal evaluators will be better able to devise high-impact interventions, and in this way support their organization in meeting its reporting requirements and fostering evaluation use for organizational learning and improvement. It is hoped that the self-assessment instrument described in this article will, when used by organizations, provide a clear picture of organizational evaluation capacity and support the development of ECB strategies for federal government organizations through its results.

ACKNOWLEDGEMENTS

The views represented in this paper are those of its authors and do not necessarily reflect those of their host organizations.

The authors wish to recognize the important contributions of the three expert reviewers as well as those of the organizations participating in the pilot-testing phase of this study. Our appreciation and thanks go out to all of you.

NOTES

- 1 The "exemplary" label was modified in the instrument to better reflect the current state of organizations. The instrument identifies "established" capacity.

- 2 Because of space considerations, not all subdimensions are included at this time.

REFERENCES

- Adams, J., & Dickinson, P. (2010). Evaluation training to build capability in the community and public health workforce. *American Journal of Evaluation*, 31(3), 421–433. <http://dx.doi.org/10.1177/1098214010366586>
- Bourgeois, I., Chouinard, J. A., & Cousins, J. B. (2008). Understanding organization capacity for evaluation: Synthesis and integration. *Canadian Journal of Program Evaluation*, 23(3), 225–237.
- 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(3), 127–146.
- Bourgeois, I., & Cousins, J. B. (2013). Understanding dimensions of organizational evaluation capacity. *American Journal of Evaluation*, 34(3), 299–319. <http://dx.doi.org/10.1177/1098214013477235>
- Bourgeois, I., Hart, R. E., Townsend, S. H., & Gagné, M. (2011). Using hybrid models to support the development of organizational evaluation capacity: A case narrative. *Evaluation and Program Planning*, 34(3), 228–235. <http://dx.doi.org/10.1016/j.evalprogplan.2011.03.003>. Medline:21555046
- 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 Publishers.
- Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Goh, S., & Lahey, R. (2008). Organizational capacity to do and use evaluation: Results of a pan-Canadian survey of evaluators. *Canadian Journal of Program Evaluation*, 23(3), 1–35.
- 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(2), 99–141.
- Díaz-Puente, J. M., Yagüe, 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(5), 478–506. <http://dx.doi.org/10.1177/0193841X08319015>. Medline:18509064
- Government of Canada, Treasury Board Secretariat. (2009). *Policy on evaluation*. Retrieved from <http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?section=text&id=15024>
- Government of Canada, Treasury Board Secretariat. (2011). *Annual report on the health of the evaluation function*. Retrieved from <http://www.tbs-sct.gc.ca/report/orp/2012/arhef-raefe02-eng.asp#context>
- Karlsson, P.-Å., Beijer, E., Eriksson, B., & Leissner, T. (2008). Evaluation workshops for capacity building in welfare work: Some Swedish examples. *Evaluation*, 14(4), 483–498. <http://dx.doi.org/10.1177/1356389008095489>
- Labin, S. N., Duffy, J. L., Meyers, D. C., Wandersman, A., & Lesesne, C. A. (2012). A research synthesis of the evaluation capacity building literature. *American Journal of Evaluation*, 33(3), 307–338. <http://dx.doi.org/10.1177/1098214011434608>
- Naccarella, L., Pirkis, J., Kohn, F., Morley, B., Burgess, P., & Blashki, G. (2007, Aug). Building evaluation capacity: Definitional and practical implications from an Australian case study. *Evaluation and Program Planning*, 30(3), 231–236. <http://dx.doi.org/10.1016/j.evalprogplan.2007.05.001>. Medline:17689328
- Nielsen, S. B., Lemire, S., & Skov, M. (2011). Measuring evaluation capacity: Results and implications of a Danish study. *American Journal of Evaluation*, 32(3), 324–344. <http://dx.doi.org/10.1177/1098214010396075>
- Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation*, 29(4), 443–459. <http://dx.doi.org/10.1177/1098214008324182>
- Sanders, J. R. (2002). Mainstreaming evaluation. In J. J. Barnette & J. R. Sanders (Eds.), *The mainstreaming of evaluation. New Directions for Evaluation*, No. 99 (pp. 3–6). San Francisco: Jossey-Bass.
- Segsworth, R. V. (2005). Program evaluation in the Government of Canada: Plus ça change.... *Canadian Journal of Program Evaluation*, 20(3), 175–197.
- Shepherd, R. P. (2012). In search of a balanced Canadian federal evaluation function: Getting to relevance. *Canadian Journal of Program Evaluation*, 26(2), 1–45.

- Stockdill, S. H., Baizerman, M., & Compton, D. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. In M. Baizerman, D. W. Compton, & S. H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, No. 93* (pp. 7–25). San Francisco: Jossey-Bass.
- Taylor-Powell, E., & Boyd, H. H. (2008). Evaluation capacity building in complex organizations. *New Directions for Evaluation, 120*, 55–69. <http://dx.doi.org/10.1002/ev.276>
- Taylor-Ritzler, T., Suarez-Balcazar, Y., Garcia-Iriarte, E., Henry, D. B., & Balcazar, F. E. (2013). Understanding and measuring evaluation capacity: A model and instrument validation study. *American Journal of Evaluation, 34*(2), 190–206. <http://dx.doi.org/10.1177/1098214012471421>
- Tseng, S.-H. (2011). An assessment scale of ECB: Perspectives from the faculty and staff of technical colleges in Taiwan. *Evaluation and Program Planning, 34*(2), 154–159. <http://dx.doi.org/10.1016/j.eval-progplan.2010.10.001>. Medline:21106244

Isabelle Bourgeois, PhD, is a Professor of Program Evaluation at l'École nationale d'administration publique (Université du Québec). Her main research interests focus on evaluation utilization and organizational evaluation capacity in public and nonprofit organizations. Prior to joining l'ENAP, Dr. Bourgeois worked as an evaluation practitioner in the private and public sectors and specialized in the evaluation of science-based federal programs.

Eleanor Toews is a Senior Evaluator at Canada Border Services Agency. Combining private and public sector experience, she leads the development and conduct of evaluations focused on service delivery government programs and the corresponding reporting to senior management. She brings strong qualitative and quantitative analytical skills to assessing the 2009 Evaluation Policy's five core issues in relation to corporate priorities. The article was authored while Eleanor was a student in Carleton University's Diploma of Program Evaluation.

Jane Whynot is a consultant specializing in performance, with a background that spans approximately 15 years in independent management for both the public and private sectors. She is the co-founder of both local and national initiatives highlighting capacity building

initiatives for the functional community, including the Lunch and Learn for Evaluators (L&Ls) in Ottawa and Evaluation Mentoring Canada (EMC). Jane was a member of the inaugural class of the graduate diploma program in Program Evaluation at Carleton University, and graduated in December 2010 from the University of Melbourne in Australia with a Master's in Assessment and Evaluation.

Mary Kay Lamarche, CE, is currently the Vice President of the Canadian Evaluation Society. She has been involved with evaluation, performance measurement, and strategic/results-based planning in both the public and private sectors for over 13 years. In October 2008, Mary Kay was part of the first cohort graduating from the Graduate Certificate in Program Evaluation at the University of Ottawa. She is presently working toward a Master of Arts in Public Ethics at Saint Paul University. In recent years, she was also a board member for the CES-National Capital chapter (CES-NCC), where she served as the chapter's representative on the CES National Council, and also sat as Chair of the Governance and Process Committee on CES National. In 2003, 2004, 2008, and 2009, Mary Kay assisted with the organization of the annual CES Student Evaluation Case Competition.