

## ORGANIZATIONAL CAPACITY TO DO AND USE EVALUATION: RESULTS OF A PAN-CANADIAN SURVEY OF EVALUATORS

J. Bradley Cousins  
Catherine Elliott  
Faculty of Education, University of Ottawa

Courtney Amo  
Atlantic Canada Opportunities Agency

Isabelle Bourgeois  
National Research Council Canada

Jill Chouinard  
Faculty of Education, University of Ottawa

Swee C. Goh  
TELFER School of Management, University of Ottawa

Robert Lahey  
REL Solutions Inc.

**Abstract:** Despite increasing interest in the integration of evaluative inquiry into organizational functions and culture, the availability of empirical research addressing organizational capacity building to do and use evaluation is limited. This exploratory descriptive survey of internal evaluators in Canada asked about evaluation capacity building in the context of organizational characteristics (learning, support structures), evaluative activity and use, and variables that mediate use. We received a total of 340 usable responses to an online survey. This article provides a descriptive account of the findings with a cursory look at differences across respondent role, organization type, and self-reported perceived level of evaluation knowledge. Results showed a pattern of moderately high ratings of organizational learning and support functions, the extent to which evaluation is being conducted and used, and stakeholder involvement in evaluation. Some differences across respondent roles, organization type, and evaluation knowledge were observed. Results are discussed in terms of an agenda for future inquiry.

**Résumé :** Malgré un intérêt accru pour l'intégration de l'enquête évaluative dans les fonctions organisationnelles et dans la culture organisationnelle, la disponibilité de recherche empirique sur le ren-

forcement des capacités organisationnelles à faire et à utiliser l'évaluation est limitée. Cette étude exploratoire et descriptive auprès d'évaluateurs internes canadiens situe le renforcement des capacités d'évaluation dans le contexte des caractéristiques organisationnelles (apprentissage, structures d'appui), de la réalisation et de l'utilisation des évaluations, et des variables qui interviennent pour médier l'utilisation. Nous avons reçu 340 réponses utilisables à un sondage en ligne. Cet article fournit un compte-rendu des résultats et un aperçu préliminaire des différences selon le rôle des répondants, les types d'organisme, et le niveau perçu de connaissances en évaluation. Les résultats démontrent une tendance vers des scores modérément élevés en ce qui a trait à l'apprentissage organisationnelle et aux fonctions d'appui. Le niveau de mise-en-œuvre et d'utilisation des évaluations ainsi que l'implication des partis prenants dans les processus d'évaluation ont également démontré des scores relativement élevés. Quelques différences selon le rôle des répondants, les types d'organismes, et le niveau de connaissances en évaluation ont été observées. La discussion traite des résultats en tant que pistes pour d'autres recherches dans ce domaine.

## INTRODUCTION

■■■■■■ Evaluation researchers and practitioners are increasingly aware of the importance of understanding the concepts of evaluation capacity and evaluation consequences (Shulha & Cousins, 1997; Stockdill, Baizerman & Compton, 2002;). Despite this interest, there is a lack of empirical research addressing capacity building issues in organizations. These issues include how to build evaluation capacity and the circumstances under which evaluation results are used in decision making. Cousins, Goh, Clark, and Lee (2004) offer some direction in this area with a conceptual framework that attempts to link some of the important concepts in the evaluation literature, especially organizational learning, evaluative inquiry, evaluation consequences, and evaluation capacity. The framework was developed through a review and synthesis of the empirical knowledge base. Through this process, the authors provide some insights on how to build evaluation capacity in organizations and on the drivers for evaluation use. In this conceptualization, evaluation capacity building (ECB) is seen to include both the capacity to *do* evaluation studies as well as the capacity within organizations to *use* evaluation results and processes.

Many public and nonprofit agencies are dedicating increased resources to ECB in an effort to link evaluative inquiry to organizational

learning. Evaluative inquiry—defined as systematic inquiry leading to judgements about program (or organizational) merit, worth, and significance in support of decision making—has been identified as a specific type of organizational learning system (Cousins, 2003; Cousins et al., 2004; Owen & Lambert, 1995; Preskill, 1994). When integrated into the organizational culture, evaluative inquiry becomes integral to the creation of new knowledge and shared understanding within the organization (Cousins et al., 2004; Preskill & Torres, 1999). Indeed, evaluative inquiry, mediated by organizational readiness for evaluation, explains within- and between-organization variation in organizational learning capacity (Cousins, Goh, & Lee, 2003). In a multiple case study, Cousins, Goh, and Clark (2005) furthered understanding of organizational benefits and consequences of evaluation, and of the factors that influence the development of a culture of evaluative inquiry in schools. However, further investigations of the conditions under which evaluation capacity develops and the extent to which such a culture can be brought about by direct ECB initiatives are still needed. In particular, organizations' development of capacity to use evaluation has received less attention from those doing research on ECB than their capacity to do evaluation. The results of the survey reported in this article are meant to inform such understanding.

Following a brief overview of the scope of the study, this article presents a conceptual framework of ECB and its underlying concepts, and then specifies the research questions driving the survey. A description of the methodology and discussion of selected survey results follows. The article ends with some general thoughts and reflections about the challenges and opportunities for integrating evaluation into the organizational culture.

## SCOPE AND PURPOSE OF THE STUDY

This article presents findings from a pan-Canadian survey of internal evaluators that explores issues of evaluation capacity building and utilization across a wide range of organizations. More specifically, the survey sought to describe the perceptions of internal evaluators regarding evaluation capacity in their organizations as well as the factors that influence evaluation capacity building and use of evaluation results and processes for decision making. In addition, we take a preliminary look at patterns of differences across organizational types (government vs. other organizations) and survey participant roles (evaluator vs. manager). With this data set we are also pursuing

more in-depth investigations of differences across groups and factor analyses of item sets; the results of these studies, though, are beyond the scope of the present article.

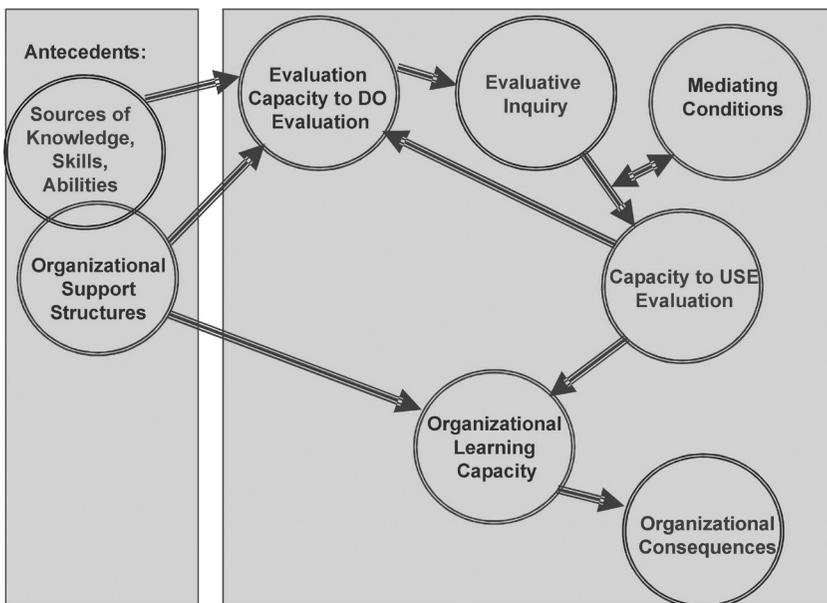
## CONSTRUCTS AND QUESTIONS GUIDING THIS STUDY

Figure 1 depicts the conceptual framework that was developed to guide this research. The framework is grounded in the authors' prior research (Cousins et al., 2004, 2005, 2006). A description of the various constructs that influenced the design of the survey questionnaire follows.

### Capacity for Organizational Learning

In the proposed model, evaluation capacity is integral to the broader construct of organizational learning capacity (OLC). Those organizations with a higher OLC are more likely to demonstrate the characteristics of learning organizations, namely: clarity and support for organizational mission and vision, leadership that supports learning, an experimenting organizational culture, the ability to transfer knowledge effectively, and teamwork and cooperation (Goh, 2000).

**Figure 1**  
**Conceptual Framework of the Capacity to Do and Use Evaluation**



The strength of each of these attributes is measured using the Organizational Learning Survey (Goh & Richards, 1997; Goh, Quon, & Cousins, 2007).

### Organizational Support Structures

OLC depends in part on organizational support structures (OSS). The first dimension of OSS, formalization, refers to the fact that organizational learning is enhanced when an organization has a flat, non-hierarchical structure with few formalized controls over employees' work. The second dimension, training support, refers to the need for organizations to invest in professional development by building team competencies (critical thinking, team problem-solving), rather than focusing only on individual job-based skills (Goh, 2000).

### Capacity to *Do* Evaluation

The capacity to *do* evaluation refers to the knowledge and skills required to carry out evaluation internally or to oversee external evaluation activities. These would include knowledge and skills pertaining to evaluation planning, standards of professional evaluation practice, instrument development, data collection, analysis and interpretation, and reporting and follow-up.

### Sources of Knowledge, Skill, and Ability

Evaluation capacity building often depends upon training and professional development opportunities. These would include indirect ECB such as on-the-job informal learning experiences as well as the more formalized direct ECB opportunities such as workshops, courses, and degree or certificate programs.

### Evaluative Inquiry

Evaluative inquiry refers to a range of activities, from needs assessment, program evaluation, and outcome monitoring to environmental scanning. Our own conceptualization of evaluation emphasizes positive, formative, or learning-oriented aspects. For example, evaluative inquiry can assist senior managers in making informed decisions about an improvement of a program or project (Cousins & Earl, 1995). But evaluative inquiry may also be driven by accountability demands. Questions related to this construct assess the degree to which evaluators see evidence of evaluation activities occurring in their organization.

## Capacity to *Use* Evaluation

In our view, organizations' capacity to use evaluation is essential to making evaluation meaningful. For example, use of findings can provide support for organizational or program decision making, promote learning about the organization or program itself, or fulfill other symbolic or political needs. The capacity to use evaluation is also reflected in process use, the extent to which members of the organization benefit from being close to or participating in the evaluation. Process use, as defined by Patton (2008), "occurs when those involved in the evaluation learn from the evaluation process itself or make program change based on the evaluation process rather than just the evaluation's findings. [It] includes cognitive, attitudinal, and behavior changes in individuals, and program or organizational changes resulting, either directly or indirectly, from engagement in the evaluation process and learning to think evaluatively" (p. 156). Process use might, for example, include increased understanding of evaluation logic or enhanced research skills (Amo & Cousins, 2007); it may also promote critical thinking, or foster an "evaluative habit of mind" (Katz, Sutherland, & Earl, 2002).

## Mediating Conditions

One of our recent studies using concept mapping techniques yielded interesting findings about the factors or conditions that mediate the use of evaluation in the Canadian federal government (Cousins et al., 2006). Some of these include evaluation quality, evaluator credibility, involvement of users, and timeliness. Several of these factors are consistent with variables that have been previously reported in the evaluation use literature (see, e.g., Cousins & Leithwood, 1986). We now turn to a description of the methods we used for the survey and then to a descriptive presentation of findings.

## METHODS

### Sampling Procedures

The study team approached a number of public and not-for-profit organizations to obtain mailing lists of possible survey participants, specifically the Canadian Evaluation Society (CES) ( $n \sim 1,600$  to 1,800 individuals), Treasury Board of Canada Secretariat (TBS) ( $n \sim 75$  individuals) and United Way (Greater Toronto and Ottawa;  $\sim 220$

organizations, number of individuals unknown). The goal was to recruit individuals who do evaluations as part of their role as well as organization members who are responsible for oversight of evaluation work that is contracted out. As such, individuals working at private consulting firms or universities were outside the scope of our sample unless they had internal evaluation responsibilities.

Although the invitation to participate was sent electronically to all list members with a two-week timeline to complete the online survey, the deadline was subsequently extended by two weeks to increase the number of survey responses. Participants could complete the survey in one or multiple sittings, and a progress bar informed them about completion levels.

### Instrument Development

The purpose of the survey was to gather information about factors related to evaluation practices within organizations in order to better understand how organizations build evaluation capacity. The questionnaire, which was developed and pilot tested by Cousins and associates, was a hybrid tool that drew items from our own prior research (Cousins, Donohue, & Bloom, 1996; Cousins et al., 2006; Goh, Cousins, & Elliott, 2006; Goh & Richards, 1997), other instruments (Preskill, Torres, & Martinez-Papponi, 1999; Seiden, 1999; Turnbull, 1999), and syntheses of recent research (Amo & Cousins, 2007; Cousins et al., 2004). The online questionnaire asked respondents to rate their level of agreement or estimated frequency of occurrence of various items associated with categories in our conceptual framework. We developed scale scores associated with nine categories of variables from our conceptual framework, each meeting acceptable thresholds of internal consistency (i.e., Cronbach's alpha exceeding 8.0). The scales were:

1. Capacity for organizational learning
2. Organizational support structures (which has two subscales: one for formalization and the other for training support)
3. Capacity to do evaluation;
4. Evaluation capacity building activities
5. Specific types of evaluation activities
6. Stakeholder participation in evaluation
7. Use of evaluation findings
8. Use of evaluation process
9. Conditions mediating evaluation use.

Reliability coefficients are reported in the ensuing tables; for all scale variables except for one—OSS-formalization—Cronbach's alpha met or exceeded the threshold of .80. The coefficient for OSS-formalization was marginally acceptable at .68; results should be treated with caution. Respondents were also asked to provide basic background information on their organization and on themselves.

## Respondents

In total, we received 340 usable responses. Because the population of persons who do internal evaluation is unknown in the organizations through which we recruited participants, we were unable to estimate the response rate for the survey. We suspect that the demographic characteristics of our sample do not depart dramatically from those of CES members who serve in internal evaluation roles (excluding, e.g., CES members who are in the private and university sectors). It is likely that fewer members in our sample are associated with government than is the case with CES, due to our efforts to recruit participants in the voluntary sector through United Way organizations.

As can be seen in Table 1, about 55% of respondents came from the government sector, while the remaining participants worked for other organizations in not-for-profit, non-governmental organization (NGO), university, or private sectors. Almost 60% of the respondents worked as evaluators (including directors of evaluation units) or performance measurement and research specialists, while just over 40% either were program managers or held senior administrative positions.

Respondents reported that approximately half of their time, on average, was devoted to evaluation activities; in terms of experience, the mean number of years in the field was calculated at 7.1 years. The self-reported level of knowledge of evaluation theory and practice was deemed to be very high, with close to 85% of respondents indicating an excellent, very good, or good level of knowledge. About 65% of the respondents were female, and close to 75% of the sample was over the age of 40. Close to 80% of respondents possessed a graduate or post-graduate degree. We observed that those working in evaluation/research roles were more likely to be associated with government departments and agencies; most of those in management positions had evaluation responsibilities but worked at not-for-profit, NGO, university, or private sector organizations:  $\chi^2(1, N = 318) = 33.76, p < .001$ . No other differences in demographic variables were associated with organizational affiliation.

## RESULTS

Descriptive statistics were prepared using SPSS 18.0. In addition, bivariate analyses (independent *t*-tests and one-way ANOVAs) were undertaken in order to examine differences in survey results by characteristics of the sample (i.e., respondent role, organization type, overall level of evaluation knowledge). Tables 2 through 9 present descriptive results for the top three and bottom three item scores within each of the main construct areas, as well as the overall scale scores (based on averages of all items in the category). In addition, these tables present the results of bivariate analyses on scale vari-

**Table 1**  
**Organizational Affiliation by Current Position, Gender, Age Category and Level of Education (*N* = 340)**

	Organizational affiliation ( <i>n</i> = 337) (%)		Total (%)
	Government	Other Sectors*	
Current position ( <i>n</i> = 318)			
Evaluator (includes performance measurement specialist)	74.0	41.8	59.7
Manager (includes all management positions)	26.0	58.2	40.3
No response			
Gender ( <i>n</i> = 328)			
Male	38.1	34.2	36.3
Female	61.9	55.8	63.7
No response			
Age category ( <i>n</i> = 323)			
Less than 30	7.6	5.3	6.5
30-39	24.0	19.7	22.0
40-49	33.3	29.6	31.6
50-59	31.6	37.5	34.4
60+	3.5	7.9	5.6
No response			
Highest diploma/degree attained ( <i>n</i> = 332)			
High school	1.7	0.7	1.2
College	0	3.3	1.5
University - undergraduate	17.2	18.4	17.8
University – graduate	67.2	53.3	60.8
University – doctoral	13.9	24.3	18.7
No response			
Total (%)	54.0	46.0	100

\* Not-for-profit, NGO, university, private sector

ables using respondent role, organizational affiliation, and overall level of evaluation knowledge as independent variables. Results are presented and discussed according to each construct measured by the survey. Given the exploratory nature of the study, additional analyses are beyond the scope of the present article.

### Organizational Learning Capacity

The first section of the questionnaire employed a revised version of Goh and Richards' (1997) Organizational Learning Survey to assess the OLC of participating schools (see also Goh et al., 2007). It consists of 18 statements to which respondents were asked to indicate their level of agreement on a 7-point Likert scale. These statements reflect the five dimensions attributed to learning organizations: mission and vision, leadership, experimentation, transfer of knowledge, teamwork and cooperation. As can be seen in Table 2, there was a modest tendency for respondents to see their organizations as having a well-developed capacity for organizational learning. Although there was some variability across respondents on this measure (i.e.,  $SD = 1.2$  on a 7-point scale), the average score across all questions was at about the midpoint of the scale. The individual item scores were also quite consistent, ranging from a low mean score of about 4 to a high of well over 5.

Scores related to mission and vision were amongst the top-ranked scores, while leadership scores were clustered near the bottom. Statements about experimentation were fairly evenly split at either extreme. In contrast, scores related to teamwork and cooperation (not included in Table 2) tended toward the middle. In general, responses indicate that there is not one single dimension that is particularly dominant in terms of contribution to OLC. However, upon closer examination of the statements, some common themes are evident. The top three scores are related to the following: clearly articulated organizational mission and values, bringing new ideas into the organization, and employee involvement in assessment. In contrast, the lowest average scores were related to accepting criticism and discussing failures or mistakes.

Bivariate analyses revealed that scale scores were significantly higher for respondents who self-identified as managers compared to those who primarily self-identified as evaluators. This may reflect managers' privileged knowledge of organizational management and strategic planning processes as well as their organizational vantage point. The role of managers in the organizational hierarchy may also

**Table 2**  
**Capacity for Organizational Learning**

Item	N	Response Distribution			Mean	SD	Rank*	95% C.I.		Scale Score
		St Disag (1-2)	(3-4-5)	St Agree (6-7)				Lower	Upper	
The organization's mission statement identifies values to which all staff must conform.	316	8%	38%	54%	5.32	1.65	1	5.14	5.50	Scale Characteristics No. items = 18 N = 306 Mean = 4.48 Std. Dev. = 1.21 Cronbach's $\alpha$ = .96
We can often bring new ideas into the organization.	333	7%	50%	43%	5.10	1.49	2	4.94	5.26	
We have opportunities for self-assessment with respect to goal attainment.	329	11%	51%	38%	4.73	1.60	3	4.55	4.90	
We have a system that allows us to learn successful practices from other organizations.	324	19%	59%	22%	4.12	1.57	16	3.95	4.29	
Managers can accept criticism without becoming overly defensive.	315	21%	57%	22%	4.06	1.57	17	3.89	4.24	Group Mean Diff's Primary Role Evaluator = 4.17 Manager = 4.94 $t(289) = -5.57$ , $p < .001$ Organization Type Government = 4.16 Other = 4.87 $t(303) = -5.36$ , $p < .001$ Eval. Knowledge N.S.
People who are new to the organization are encouraged to question the way things are done.	322	21%	58%	21%	4.03	1.68	18	3.85	4.22	

\*Rank 1 = highest agreement; 18 = lowest agreement

help to explain their more positive responses, as they may have more control over managerial processes and may be closer to the mission and values of the organization. In addition, managers tend to be generalists rather than specialists, and may therefore have a greater understanding than evaluators of the processes that are characteristic of organizational learning.

Scores were also significantly higher for respondents from other (non-government) organizations as compared to those from government organizations. The characteristics of evaluators in not-for-profit, NGO, university, and private sector organizations may be closer to managers in that they also tend to be generalists rather than specialists. The smaller size of these organizations as compared to the typical government organization may also be conducive to a more informal atmosphere or culture where ideas, innovations, and knowledge are more easily shared.

These results suggest that while employees are encouraged to participate in some degree of experimentation, problem-solving, and assessment, there is a tendency for only positive news to be welcomed and shared. There appears to be reluctance in questioning the status quo, viewing criticism as constructive, or taking risks, more so in government as compared to other organizations. In general, these results characterize organizations that are inward-looking and that may consider new ideas proposed by accepted members, but that do not necessarily seek best practices or easily accept constructive criticism.

## Organizational Support Structures

This section of the questionnaire elicited respondents' opinions about two different constructs related to organizational support for OLC: (a) the degree to which the organization supports the respondent in terms of training, and (b) the degree of formalization in the organization. Whereas the former has been shown to be positively correlated to organizational learning capacity, the latter has been shown to be negatively correlated (Goh, 2000). To analyze the results, we separated the statements related to each construct in Tables 3a and 3b; each will be discussed separately.

### Organization Support Structures – Training

Overall, the responses to this section were positive, with an average score of about 5 on a 7-point scale. Respondents gave the highest

marks to statements related to the importance and relevance of training. They also indicated that staff is encouraged to continuously develop work-related skills. However, the actual *provision* of training was rated slightly less favourably. Although training is *encouraged*, it does not necessarily appear to be *supported* or *provided* equally to all levels in the hierarchy. This was evidenced by the two lowest scores. When examining differences between respondent groups, managers provided significantly higher scale scores than evaluators, as shown in Table 3a. This may reflect the role of managers as sponsors or supporters of training and their level of awareness of such activities across the organization. Respondents from government provided significantly lower scores than those in other organizations for the scale score. There were no differences attributable to overall evaluation knowledge.

#### Organization Support Structures – Formalization

Table 3b illustrates the results of the formalization scores. Formalization refers to the presence of articulated, formalized practices such as documented policies and procedures, clearly articulated rules and controls governing organizational behaviour, and typically bureaucratic decision making (multiple levels of approvals required). As mentioned above, Goh (2000) has found that increased levels of formalization are negatively correlated to OLC; therefore, a lower score is desired. An overall average score of just less than 4.5 indicates a moderate or neutral level of agreement (mid-point). Worthy of note is the comparatively high average score for using the “proper channels” for “information and decision making” (almost 5.5; 80% agree to strongly agree). Such a response suggests that organizational decision making is quite clearly defined, but could possibly run the risk of being inflexible and cumbersome. The next highest scores were related to adhering to “formal rules and procedures” (overall average score was about 5 on the 7-point scale); and work being “monitored and inspected by management” (overall average score was over 4). These results could be a sign that some respondents feel constrained by “micro-management” or an overly rigid set of rules (it is important to note, however, that fewer than 10% selected a score of 7). Not surprisingly, the scores of respondents from other organizations are significantly lower than those from government, indicating a lower perceived level of formalization in such organizations. Many of these organizations are generally smaller, more mission-driven, and more informal than government organizations, which tend to be unionized, rules-driven, and bureaucratic. Respondents who identify as managers and those

**Table 3a**  
**Organizational Support Structures — Training**

Item	N	Response Distribution			Mean	SD	Rank*	95% C.I.		Scale Score
		St Disag (1-2)	(3-4-5)	St Agree (6-7)				Lower	Upper	
Staff training is relevant to our work.	329	3%	41%	56%	5.44	1.36	1	5.30	5.59	<i>Scale Characteristics</i> No. items = 8 N = 314 Mean = 5.10 Std. Dev. = 1.21 Cronbach's $\alpha$ = .93
The skills training that we receive can be applied to improve our work.	324	4%	46%	50%	5.31	1.34	2	5.16	5.45	
Learning that improves the work skills and knowledge of staff is encouraged.	331	4%	44%	52%	5.29	1.33	3	5.14	5.43	
Training is done in teams where appropriate.	309	11%	67%	45%	5.02	1.60	6	4.84	5.20	
This organization supports the development of skills such as leadership, coaching, and teambuilding among staff.	330	7%	55%	38%	4.79	1.53	7	4.62	4.95	<i>Group Mean Diffis</i> Primary Role Evaluator = 4.89 Manager = 5.40 $t(298) = -3.63$ , $p < .001$ Organization Type Government = 4.94 Other = 5.30 $t(311) = -2.69$ $p = .008$ Eval. Knowledge N.S.
Staff training is emphasized equally at all levels	308	17%	45%	38%	4.54	1.78	8	4.34	4.74	

\*Rank 1 = highest agreement; 8 = lowest agreement

**Table 3b**  
**Organizational Support Structures — Formalization**

Item	N	Response Distribution			Mean	SD	Rank*	95% C.I.		Scale Score
		St Disag (1-2)	(3-4-5)	St Agree (6-7)				Lower	Upper	
Information and decision-making must always go through proper channels.	329	6%	36%	58%	5.40	1.50	1	5.24	5.57	Scale Characteristics No. items = 7 N = 320 Mean = 4.35 Std. Dev. = 1.04 Cronbach's $\alpha$ = .68
Most of our work must adhere to formal rules and procedures.	331	8%	48%	44%	4.95	1.61	2	4.77	5.12	Group Mean Diffis Primary Role N.S.
Our work is usually closely monitored and inspected by management.	329	18%	55%	27%	4.27	1.66	3	4.09	4.45	Organization Type Government = 4.54 Other = 4.15 $t(317) = 3.50$ $p = .001$
We require approval in writing for the introduction of new work activities	314	22%	54%	24%	4.13	1.74	4	3.94	4.32	Eval. Knowledge N.S.
In my opinion, our organization has too many levels of hierarchy	294	31%	41%	28%	3.95	1.97	5	3.73	4.18	
There is very little overlap in work between different units in the organization	329	35%	36%	29%	3.88	2.02	6	3.66	4.10	

\*Rank: 1= highest agreement; 6= lowest agreement

who identify as evaluators held very similar perceptions of the degree of formalization within their organizations. Likewise, level of evaluation knowledge was not related to perceived formalization.

### Capacity to Do Evaluation

This section of the questionnaire elicited respondents' opinions about their organization's capacity to *do* evaluation. It includes questions related to resources for conducting evaluations (human resources, financial resources, time, tools) and organizational support (rewards, sponsorship, planning processes, and follow-up). Participants were presented with 12 statements to which they indicated their level of agreement on a 7-point scale. On average, respondents reported a relatively high level of agreement with these statements, with a mean scale score of almost 5. As can be seen in Table 4, the highest ratings were given to statements about accountability and performance reporting or monitoring. Also highly rated were statements about the provision of basic tools and resources to support evaluation. It is interesting to note, however, that employees do not believe that they are necessarily rewarded for *using* performance information despite the fact that collecting this information seems to be a very high priority in organizations. Also interesting is the fact that employees, on average, are not given sufficient time to reflect on organizational successes or failures, nor are they confident that they have the long-term, dedicated fiscal support for their evaluation activities. Taken in its totality, these scores paint a picture of organizations that could be more concerned with meeting immediate requirements for accountability and reporting than with longer-term continuous improvement and learning from results.

Comparisons across respondent groups revealed that perceptions about organizational capacity to do evaluation did not differ by organizational role or type of organization. The extent to which respondents self-reported levels of evaluation knowledge did make a difference, however. Respondents with higher levels of evaluation knowledge rated their organization as having a higher capacity to conduct evaluation. It is likely that many of these respondents participate directly in such work.

### Capacity Building Activities

To obtain an overall reading of capacity building activities, respondents were asked to indicate their level of agreement with a general

**Table 4**  
**Capacity to do Evaluation**

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.		Scale Score
		St Disag (1-2)	(3-5)	St Agree (6-7)				Lower	Upper	
We have formal requirements to report on performance.	321	7%	32%	61%	5.53	1.63	1	5.35	5.71	<i>Scale Characteristics</i> No. items = 12 N = 314 Mean = 4.81 Std. Dev. = 1.22 Cronbach's $\alpha$ = .91  <i>Group Mean Diff</i> Primary Role N.S. Organization Type N.S. Eval. Knowledge Low = 4.59 Mod = 4.86 High = 5.14 $F(2,310) = 4.45,$ $p = .012$
Our organization provides us with the basic tools/resources to support evaluation (e.g. computers, software, copying, administrative support).	333	7%	31%	62%	5.49	1.61	2	5.31	5.66	
Performance measurement is integral to our organizational accountability framework.	325	9%	35%	56%	5.32	1.71	3	5.13	5.50	
We are rewarded for using performance information.	297	21%	48%	31%	4.27	1.83	10	4.06	4.47	
Our organization has long-term, dedicated fiscal support for our evaluation activities.	320	27%	40%	33%	4.19	2.04	11	3.96	4.41	
Employees are given sufficient time to reflect on organizational successes or failures.	319	20%	61%	19%	3.98	1.61	12	3.81	4.16	

\*Rank: 1= highest agreement; 12= lowest agreement

statement: "Our organization supports evaluation capacity building activities." The average response was moderately favourable, at 4.87 ( $SD = 1.71$ ), which is similar to the overall average from the previous scale, "Capacity to do evaluation." This score was explained by both primary role and organization type. Managers were more likely to rate their organization as being committed to evaluation capacity building ( $M = 5.17$ ,  $SD = 1.2$ ) than were evaluators ( $M = 4.72$ ,  $SD = 1.72$ ),  $t(310) = -2.31$ ,  $p = .022$ . It is likely that evaluators were responding to perceived resource constraints, compared with their manager counterparts. Similarly, government respondents were less likely to report that their organization supports evaluation capacity building activities ( $M = 4.62$ ,  $SD = 1.72$ ) than did those from other organizations ( $M = 5.20$ ,  $SD = 1.61$ ),  $t(325) = -3.16$ ,  $p = .002$ , which may be reflective of a comparative (and more recent) push from those outside of government organizations and program funders to provide evaluation capacity building activities in response to calls for greater accountability and reporting on results. Such organizations often require performance measurement to be undertaken by program staff, and evaluation functions sometimes follow suit. No differences were observed in relation to overall knowledge about evaluation.

### Specific Types of Evaluation Activities

To gain a greater understanding of the degree to which different evaluation activities are pursued, respondents were presented with 17 items, to which they responded on a 7-point scale, ranging from *Never* (1) to *Always* (7). As can be seen in Table 5, in keeping with the results concerning the capacity to do evaluation, the top two most frequently reported activities were "produce reports for Boards of Directors and/or senior management" (over 5.5) and "produce reports about program activities" (just under 5.5). These reports could be produced, at least in part, to meet the performance reporting and accountability requirements identified earlier. At the other extreme, the least frequently performed activities were using certain types of research designs: comparative group designs (under 3.0), single-case mixed-method designs, program theoretical designs, and conducting first-hand observation of program activities (all of the last three with means at about 4.5). In the middle were items such as preparing logic models, monitoring program outcomes, monitoring program implementation, and monitoring program effectiveness, which are not included in Table 5. Respondents who identify primarily as managers report a significantly higher frequency in terms of the production of reports for Boards of Directors and/or senior managers, as well as

**Table 5**  
**Specific Types of Evaluation Activities**

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.		Scale Score
		Never (1-2)	(3-5)	Always (6-7)				Lower	Upper	
Produce reports for Boards of Directors and/or senior management.	321	8%	26%	66%	5.63	1.60	1	5.45	5.80	Scale Characteristics No. items = 17 N = 303 Mean = 4.85 Std. Dev. = 1.17 Cronbach's $\alpha$ = .92
Produce reports about program activities	327	8%	34%	58%	5.42	1.61	2	5.24	5.59	Group Mean Diff's Primary Role N.S. Organization Type N.S. Eval. Knowledge Low = 4.64 Mod = 4.91 High = 5.08 $F(2,300) = 3.27$ , $p = .04$
Review program documentation (e.g., participant records, case notes)	310	8%	37%	55%	5.37	1.63	3	5.19	5.55	
Use program theoretical designs (i.e., link program components to outcomes)	311	20%	41%	39%	4.54	1.94	15	4.32	4.76	
Employ single-case mixed-method designs (i.e., qualitative and quantitative methods) management	310	22%	41%	37%	4.47	1.92	16	4.26	4.68	
Use comparative group designs (e.g., randomized control trials, quasi experimental)	307	61%	23%	11%	2.70	1.87	17	2.49	2.91	

\*Rank: 1 = highest agreement; 17 = lowest agreement

reports about program activities, than do respondents who identify as evaluators. Although we do not have data to confirm this, it is possible that managers are using a broader definition of evaluation that may include a larger set of analyses and reports being prepared. In addition, reports about program activities were identified more frequently in non-government settings, which may reflect a greater focus on formative issues, rather than the need for information due to policy requirements.

The scale score for types of evaluation activities had a mean of almost 5. No differences were associated with primary role or organization type, but we observed that level of evaluation knowledge was positively associated with the dependent variable. That is, the more knowledgeable the respondent, the more likely they were to report a greater range of evaluative activities in their organization.

### Stakeholder Participation

We were also interested in the degree to which various stakeholder groups were involved in evaluation. This was measured according to two dimensions: level of participation (*low to high* on a 7-point scale; see Table 6a) and frequency of participation (*never to always* on a 7-point scale; see Table 6b). On both scales, stakeholder level of involvement was consistent. The rank order, from highest to lowest participation by stakeholders, was: (1) staff responsible for implementing the program, (2) program managers or directors, (3) program developers, (4) intended beneficiaries of the program, (5) program sponsors or funders, and (6) special interest groups. In terms of the strength of the responses, the first two groups were rated significantly higher than the third (program developers). These results are not surprising given that, in most organizations including government, programs are often longstanding and the actual developers may no longer be involved in the organization or available to participate.

When comparing across organization types, the higher level of participation of stakeholders in organizations outside of government may speak to an environment where stakeholders—including program beneficiaries—are closer to programs in terms of both proximity (local and community-based) and organizational responsibility for doing and using evaluation (these results were statistically marginal). The level of resources for evaluation in this context may also result in program staff and other stakeholders being more directly involved

in evaluation, whereas in government, resources allow for the hiring of external consultants. It is interesting to note that in both types of organizations, those closest to the program show higher and more frequent levels of participation in evaluation. No group differences were found to be associated with primary role, but for overall level of evaluation knowledge a marginal difference was observed in relation to the frequency of stakeholder participation. The scale score for frequency of participation was found to increase as did evaluation knowledge.

## Evaluation Utilization

Two sections of the survey addressed evaluation use. First, evaluators were asked about utilization of evaluation results—the top three and bottom three scored items are presented in Table 7. Second, they were asked to respond to a series of questions about process use—again, the top three and bottom three are presented in Table 8.

### *Use of Evaluation Findings*

Evaluators were asked to respond to 10 statements regarding the extent to which they had observed different consequences of evaluation use. Once again, the top scores related to accountability or reporting “to the board (or equivalent),” both with averages over 5.5. These responses parallel the results of two previously discussed scales associated with the capacity to do evaluation as well as specific types of evaluation activities. It is also notable that these top scores are relatively high, indicating that respondents perceive evaluation results are very frequently used in these ways.

The next group of scores is related more to the formative function of evaluation. For example, the third and fourth ranked items refer to learning about “program functioning” and helping to make “changes to existing programs.” In contrast, the lowest scores indicate that evaluation results are not used as commonly in other activities. These include decision making for staffing or fiscal allocations, performing “outreach and public relations,” acquiring “new funding,” or conducting strategic planning. Certainly within the government sector, many of these functions—staffing, financial—are quite structured and bureaucratic in nature; these processes may not have the adaptability or flexibility to include evaluation results as an additional source of data, even if these data are available in a timely fashion.

**Table 6a**  
**Stakeholder Participation in Evaluation: Level of Participation**

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.		Scale Score
		Low (1-2)	(3-4-5)	High (6-7)				Lower	Upper	
Staff responsible for implementing the program	321	6%	34%	60%	5.48	1.55	1	5.31	5.65	<i>Scale Characteristics</i> No. items = 6 N = 294 Mean = 4.55 Std. Dev. = 1.27 Cronbach's $\alpha$ = .81
Program managers or directors	322	5%	36%	59%	5.45	1.47	2	5.29	5.61	
Program developers	302	14%	42%	44%	4.81	1.81	3	4.61	5.02	
Intended beneficiaries of the program	316	24%	45%	31%	4.21	1.92	4	4.00	4.42	
Program sponsors or funders	282	30%	42%	28%	4.01	1.98	5	3.78	4.24	
Special interest groups	269	46%	34%	20%	3.33	1.95	6	3.09	3.56	<i>Group Mean Diff's</i> Primary Role N.S. Organization Type Government = 4.43 Other = 4.71 $t(291) = -1.89$ $p = .059$ Eval. Knowledge N.S.

\*Rank: 1 = highest agreement; 6 = lowest agreement

**Table 6b**  
**Stakeholder Participation in Evaluation: Frequency of Participation**

Item	N	Response Distribution			Mean	SD	Rank <sup>2</sup>	95% C. I.		Scale Score
		Never (1-2) <sup>1</sup>	(3-4-5)	Always (6-7)				Lower	Upper	
Staff responsible for implementing the program	316	5%	32%	63%	5.54	1.614	1	5.36	5.72	<i>Scale Characteristics</i> No. Items = 6 N = 290 Mean = 4.56 Std. Dev. = 1.26 Cronbach's $\alpha$ = .80  <i>Group Mean Diff's</i> Primary Role N.S. Organization Type Government = 4.45 Other = 4.71 $t(287) = -1.75$ $p = .082$ Eval. Knowledge Low = 4.45 Mod = 4.48 High = 4.89 $F(2, 287) = 2.73,$ $p = .067$
Program managers or directors	318	8%	32%	60%	5.48	1.584	2	5.31	5.66	
Program developers- Frequency of participation	300	16%	43%	41%	4.78	1.815	3	4.58	4.99	
Intended beneficiaries of the program	313	25%	45%	30%	4.18	1.902	4	3.97	4.39	
Program sponsors or funders	277	31%	41%	28%	3.95	2.008	5	3.72	4.19	
Special interest groups	259	48%	35%	17%	3.24	1.969	6	3.00	3.48	

<sup>1</sup>Rank 1 = highest agreement; 6 = lowest agreement

**Table 7**  
**Use of Evaluation Findings**

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.		Scale Score
		Never (1-2)	(3-4-5)	Always (6-7)				Lower	Upper	
Evaluation helps us to meet external accountability requirements.	318	6%	25%	69%	5.71	1.52	1	5.55	5.88	Scale Characteristics No. items = 10 N = 315 Mean = 5.0 Std. Dev. = 1.21 Cronbach's $\alpha$ = .91
Evaluation helps us to report to the board (or equivalent).	320	7%	29%	64%	5.63	1.50	2	5.46	5.79	Group Mean Diff's Primary Role Evaluator = 4.89 Manager = 5.18 $t(300) = -2.10$ , $p = .036$
Evaluation helps us to learn about program functioning.	331	5%	31%	64%	5.61	1.42	3	5.45	5.76	Organization Type Government = 4.74 Other = 5.32 $t(312) = -4.45$ $p < .001$
Evaluation helps us to make decisions about fiscal allocations.	314	13%	50%	37%	4.61	1.72	8	4.42	4.80	Eval. Knowledge N.S.
Evaluation helps us to perform outreach and public relations.	313	20%	46%	34%	4.42	1.82	9	4.22	4.62	
Evaluation helps us to make decisions about staffing.	305	28%	50%	29%	4.05	1.89	10	3.83	4.26	

\*Rank: 1= highest agreement; 10= lowest agreement

**Table 8**  
**Use of Evaluation Process**

Involvement in evaluation activities and processes has helped us to:

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.		Scale Score
		Never (1,2)	(3-4-5)	Always (6-7)				Lower	Upper	
develop a better understanding of the program/policy/intervention being evaluated.	324	8%	29	63	5.48	1.52	1	5.32	5.65	<i>Scale Characteristics</i> No. items = 17 N = 299 Mean = 4.76 Std. Dev. = 1.31 Cronbach's $\alpha$ = .97  <i>Group Mean Diffs</i> Primary Role N.S. Organization Type Government = 4.64 Other = 4.92 $t(296) = -1.84$ $p = .067$ Eval. Knowledge N.S.
develop knowledge about evaluation logic.	318	9%	37	54	5.27	1.55	2	5.10	5.44	
develop knowledge about evaluation methods.	324	9%	39	52	5.21	1.57	3	5.04	5.38	
increase organizational commitment.	306	15%	52	33	4.49	1.64	15	4.30	4.67	
expedite the use of evaluation findings.	315	14%	55	31	4.46	1.63	16	4.28	4.64	
develop professional networks.	312	19%	48	33	4.31	1.77	17	4.12	4.51	

\*Rank: 1 = highest agreement; 17 = lowest agreement

When comparing results across organization types, those respondents working in organizations outside of government perceived a higher level of use of evaluation findings than did their governmental counterparts. This may be reflective of the differing governance structures between these two settings, with the existence of boards or councils being less frequent in government. Differences in the scale score for findings use were also associated with the respondent's primary role; managers perceived higher levels of use than did evaluators. It seems likely that managers would to some extent be involved in the use of evaluation findings or at a minimum be more privy to the use of evaluation findings than would their evaluator counterparts. No differences were associated with overall level of evaluation knowledge.

### *Use of Evaluation Process*

Seventeen (17) scale items were included in the process use scale; respondents were asked to "indicate the extent to which involvement in evaluation activities and processes have had the following consequences for members of your organization." Responses were moderate overall, with an average score of under 5.0 (slightly lower than for use of findings). Respondents gave the highest marks to developing a "better understanding of the program/policy/intervention being evaluated" with an average score of about 5.5. The next three items were related to developing evaluation capacity through process use: developing knowledge about evaluation logic, methods, and technical skills (all means at 5.0 or higher).

At the bottom of the list were more general consequences such as developing professional networks or increasing organizational commitment. There were also statements regarding "expediting" or "expanding" the use of evaluation findings, improving management practices, and statements related to organizational learning. The lower scores on these statements may reflect respondents' lack of observation of a direct link between involvement in—or proximity to—evaluation and changes in attitudes or actions within the organization, whereas changes in knowledge and skills may be more readily apparent. When comparing scores across respondent groups, no differences were noted for primary role or for overall level of evaluation knowledge. However, respondents working in organizations other than government had significantly higher scale scores for process use than did their government counterparts. It seems likely that such respondents would be better positioned to observe such effects

given the smaller size of such organizations and the more widespread involvement of stakeholders in evaluation (see Tables 6a and 6b). As mentioned, these participants also observed higher levels of use of findings (see Table 7).

### Conditions Mediating Evaluation Use

This scale elicited responses to 19 items about the factors and conditions that mediate evaluation use. Respondents were asked to rate the degree to which they have observed these factors in their organization, using a scale of 1 (*never*) to 7 (*always*). Seven of the 19 statements had a mean score of over 5, indicating that these factors are frequently observed. As can be seen in Table 9, the top response was significantly higher than the others, with an average of over 6.0, suggesting that evaluations are readily “accessible to senior management.” The strength of this response implies that evaluation is well-positioned at a high level in the organization and/or the results of the evaluation are made available to senior decision makers. However, it does not say that evaluations are necessarily *used* to make “tough decisions,” nor are they frequently “timely” or “[linked to the decision-cycle of senior decision makers].” Several of the other top results relate to *users*—their involvement in the evaluation process and their understanding of the process. For example, respondents indicated that evaluations are frequently “informed by user input,” that users are involved “throughout the evaluation process,” and that they understand the methodology (all means over 5.0). Other top results dealt with perceived credibility of the evaluation, the objectivity of data, and the fact that evaluations followed pre-existing plans (again, all means over 5.0).

At the other end of the spectrum, respondents indicated that they do not often use peer reviews and that evaluations are often not timely or used to make tough decisions (as mentioned above). Furthermore, they indicate that evaluations are not often compared “against external standards or benchmarks,” nor are they seen to be “properly resourced” (mean scores at about 4.5).

In terms of differences across respondent groups, evaluators scored slightly higher than did managers on the scale score for mediating conditions (marginally statistically significant). It is possible that such individuals generally present the findings of their evaluations to senior managers and are therefore probably more knowledgeable about how these results are made available and may obtain direct

feedback in terms of the credibility of results. Differences across respondents' level of knowledge were also found, with those being more knowledgeable about evaluation having higher scale scores for mediating conditions. No differences were found for organization type.

## DISCUSSION AND CONCLUSION

This article presented and discussed the descriptive results of a pan-Canadian survey of internal evaluators exploring organizational capacity to *do* and to *use* evaluation in both government and non-government settings. The study sits within a larger program of research aimed at increasing our understanding of the conditions under which ECB interventions are most likely to be effective. We are unaware of comprehensive ECB intervention studies that seek to further understand not only how to do evaluative inquiry but also how to develop a culture of evaluation within organizations. To that end, this interdisciplinary research has great potential to fill identified gaps in the academic knowledge base. The survey results presented and discussed in this article provide some interesting glimpses into the organizational practices and uses of evaluation in Canada. Further analyses and discussion of differences between government and non-governmental settings are presented in Cousins, Goh, Elliott, and Aubry (2008).

Some limitations need to be taken into account in interpreting the results of this survey. First, although the survey was directed at internal evaluators, close to half of the respondents (42.1%) identified their current role as being one related to program management or senior management within the organization. When examining the relationship between current role and time spent on evaluation, we found a significant negative correlation indicating that respondents who identify as managers are spending significantly less time on evaluation than those identifying as evaluators ( $r [308] = -.570, p < .001$ ). We also found significant negative correlations between role and level of knowledge of evaluation *theory* and *practice* ( $r (321) = -.178, p < .0001$  and  $r (321) = -.330, p < .001$ , respectively). It is important to consider that the significant differences found in the scores provided by these two groups of respondents may reflect not only their differing vantage points, but also their differing levels of involvement in and knowledge of evaluation. In addition, we found a significant positive correlation between respondent role and respondent organization type ( $r [317] = .271, p < .001$ ), indicating that managers were more likely to be employed in organizations outside of government than were respondents who identified as evaluators. As such, the interpre-

**Table 9**  
**Conditions Mediating Evaluation Use**

Item	N	Response Distribution			Mean	SD	Rank*	95% C. I.	
		Never (1-2)	(3-4-5)	Always (6-7)				Lower	Upper
Evaluations in our organization are accessible to senior management	325	2%	22%	76%	6.06	1.204	1	5.93	6.19
Evaluations in our organization are produced by evaluators who are perceived as credible	310	4%	43%	53%	5.28	1.437	2	5.12	5.44
Evaluations in our organization are informed by user input	315	6%	44%	50%	5.24	1.450	3	5.08	5.40
Evaluations in our organization are timely (linked to decision-cycle of senior decision makers)	317	14%	58%	28%	4.37	1.581	17	4.20	4.54
Evaluations in our organization help users make tough decisions	306	14%	59%	27%	4.34	1.541	18	4.17	4.51
Evaluations in our organization are supported by quality assurance mechanisms (e.g., peer reviews)	308	21%	48%	31%	4.32	1.634	19	4.11	4.52

\*Rank: 1= highest agreement; 18= lowest agreement

*Scale Score*

*Scale Characteristics*

No. items = 19  
 N = 287  
 Mean = 4.9  
 Std. Dev. = 1.17  
 Cronbach's  $\alpha$  = .96

*Group Mean Diff's*

Primary Role Evaluator = 4.98  
 Manager = 4.73  
 $t(300) = 1.79$ ,  
 $p = .074$

*Organization Type*

N.S.  
 Eval. Knowledge  
 Low = 4.54  
 Mod = 5.00  
 High = 5.18  
 $F(2, 284) = 7.89$ ,  
 $p < .001$

tation of differences between manager and evaluator responses may be less applicable to the government setting where the line between these respective roles is likely to be more concretely delineated.

A second limitation that should be taken into account is the fact that we have only a limited sense of the population from which this sample is drawn and we are unable to credibly estimate a response rate. The generalizability of these specific findings is therefore limited, but will be somewhat mitigated within the broader context of the overall research program that involves multiple lines of evidence. Third, and related, the non-random nature of the sample impedes our ability to generalize the findings with confidence to a population of persons responsible for internal evaluation. Despite this limitation we elected to use inferential statistics on the data and to set our alpha level to .10 (instead of .05), decisions justified by the exploratory nature of the study. Finally, it is important to note that our data are based on self-reported perceptions with no corroborating information and that we surveyed only internal evaluators (supply side) and not organizational members who commission evaluation (demand side). The last two points are perhaps less limitations of the exploratory study than considerations for an agenda for future research.

When taken as a whole, the survey results presented in this article paint a portrait of moderate capacity to do and to use evaluation, with higher capacity to use evaluation in organizations outside of government (significantly higher in some areas). Although the challenges in integrating evaluation into the organizational culture are quite different in government and non-government settings, these results provide insight into the factors and conditions that could support greater evaluation capacity in both. In particular, both settings show low scores on items related to experimentation, openness to new ideas, and encouraging employees to be critical and to ask questions. Higher scores on these items are characteristic of an “evaluation habit of mind” (Katz et al., 2002) or the ability to “think evaluatively” (Patton, 2008), which is at the heart of integrating evaluation into the organizational culture—an integration that is essential to the success of results-based management and accountability regimes.

In addition, there are a few factors characteristic of the non-government setting that may facilitate individuals’ development of evaluative thinking. First, evaluators appear to play multiple roles within the organization; second, the involvement of those closest to the program being evaluated appears to be greater and more frequent.

Given that both government and non-government settings share the ultimate objectives of organizational learning, continuous improvement, and optimal use of resources, it may be valuable for evaluation policymakers and practitioners in the Canadian federal government context to gain a greater understanding of evaluation culture within non-government settings.

Overall, these findings can be of value in both government and non-government organizations in both assessing and increasing organizational capacity to do and to use evaluation. Some would argue that evaluation can serve as a lever for organizational learning and change, yet for this to come about, significant use of evaluation findings and process are required. There is a growing body of literature to show that relationship building between evaluators and program community members (e.g., program managers, high-level decision makers) is essential to the prospects of enhancing use. This is particularly relevant in the federal government context, where participatory approaches to evaluation are sometimes seen to be incompatible with evaluator objectivity and neutrality (actual or perceived).

The results of this survey also raise a number of avenues for further inquiry. First, it would be valuable to examine in more detail the various organization types grouped within the government and non-government categories. Second, given that this survey was administered in 2007, a re-administration in a few years would likely yield quite different results, particularly given the new evaluation policy directions of the Treasury Board of Canada. For instance, it is likely that impending changes to the Evaluation Policy in the Government of Canada will result in changes to both stakeholder involvement (given the strong emphasis on neutrality) and use of evaluation findings (given the need to monitor implementation of evaluation recommendations). Third, this article did not explore the relationships between the various concepts within the conceptual model presented in Figure 1, but the findings presented speak to the need for multivariate analyses. Such analyses would help to increase understanding of the conceptual framework and increase its practical value. Finally, although the need for public accountability continues to grow, the results of this survey show overall moderate levels of evaluation use. It would be valuable to explore the relationship between the drivers for evaluation and various types of evaluation use. In this context, it will be important for future surveys to involve defined evaluation users rather than evaluators speaking to perceived or assumed use within their organizations.

## ACKNOWLEDGEMENTS

The authors gratefully acknowledge support provided by the Social Science and Humanities Research Council of Canada (grant 410-2005-0303). Special thanks to Katherine Rizzo for her the assistance on this project. A previous version of this article was presented at the annual meeting of the Canadian Evaluation Society, Winnipeg, MB, June 2007.

## REFERENCES

- Amo, C., & Cousins, J. B. (2007). Going through the process: An examination of the operationalization of process use in empirical research on evaluation. In J. B. Cousins (Ed.), *Process use in theory, research and practice. New Directions in Evaluation* (pp. 5–26). San Francisco, CA: Jossey Bass.
- Cousins, J. B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan, D. L. Stufflebeam, & L. A. Wingate (Eds.), *International handbook of educational evaluation* (pp. 245–265). Boston, MA: Kluwer.
- Cousins, J. B., Donohue, J. J., & Bloom, G. A. (1996). Collaborative evaluation in North America: Evaluators' self-reported opinions, practices and consequences. *Evaluation Practice*, 17(3), 207–226.
- Cousins, J. B., & Earl, L. M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer Press.
- Cousins, J. B., Goh, S., Aubry, T., Lahey, R., Montague, S., & Elliott, C. (2006, November). *What makes evaluation useful in government? A concept mapping study*. Paper presented at the annual meeting of the American Evaluation Association, Portland, OR.
- Cousins, J. B., Goh, S., & Clark, S. (2005). Data use leads to data valuing: Evaluative inquiry for school decision making. *Leadership and Policy in Schools*, 4, 155–176.
- Cousins, J. B., Goh, S., Clark, S., & Lee, L. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation*, 19(2), 99–141.

- Cousins, J. B., Goh, S., Elliott, 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, Quebec, QC.
- Cousins, J. B., Goh, S., & Lee, L. (2003, July). *The learning capacity of schools as a function of evaluative inquiry: An empirical study*. Paper presented at the Learning Conference, London, University of London, Institute of Education.
- Cousins, J. B., & Leithwood, K. A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–364.
- Goh, S. (2000). Towards a learning organization: The strategic building blocks. *SAM Advanced Management Journal*, 63(2), 15–22.
- Goh, S., Cousins, J. B., & Elliott, C. (2006). Organizational learning capacity, evaluative inquiry, and readiness for change in schools: Views and perceptions of educators. *Journal of Educational Change*, 7, 289–318.
- Goh, S., & Richards, G. (1997). Benchmarking the learning capability of organizations. *European Management Journal*, 15(5), 575–583.
- Goh, S. C., Quon, T. K., & Cousins, J. B. (2007). The organizational learning survey: A re-evaluation of unidimensionality. *Psychological Reports*, 101, 707–721.
- Katz, S., Sutherland, S., & Earl, L. (2002). Developing an evaluation habit of mind. *Canadian Journal of Program Evaluation*, 17(2), 103–119.
- Owen, J., & Lambert, K. (1995). Roles for evaluators in learning organizations. *Evaluation*, 1, 237–250.
- Patton, M.Q. (2008). *Utilization-focused evaluation* (4th ed.). Thousand Oaks, CA: Sage.
- Preskill, H. (1994). Evaluation's role in enhancing organizational learning. *Evaluation and Program Planning*, 17(3), 291–297.
- Preskill, H., & Torres, R. (1998). *Evaluative inquiry for organizational learning*. Twin Oaks, CA: Sage.

- Preskill, H., Torres, R., & Martinez-Papponi, B. (1999, November). *Assessing an organization's readiness for learning from evaluative inquiry*. Paper presented at the annual conference of the American Evaluation Association, Orlando, FL.
- Seiden, K. (1999, November). *Organizational readiness for evaluation survey instrument: A preliminary factor analysis*. Paper presented at the annual meeting of the American Evaluation Association, Orlando, FL.
- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research and practice since 1986. *Evaluation Practice*, 18, 195–208.
- Stockdill, S. H., Baizerman, M., & Compton, D. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. In D. W. Compton, M. Baizerman, & S. H. Stockdill (Eds.), *The art, craft, and science of evaluation capacity building. New Directions for Evaluation*, 93, 7–25.
- Turnbull, B. (1999). The mediating effect of participation efficacy on evaluation use. *Evaluation and Program Planning*, 22(2), 131–140.

**J. Bradley Cousins**, Ph.D., is a Professor in the Faculty of Education, University of Ottawa, where he teaches and conducts research on program evaluation. He is the author of several articles, chapters and books on evaluation use, participatory evaluation, and evaluation capacity building. In 2008 Professor Cousins was in his seventh year as editor-in-chief of the *Canadian Journal of Program Evaluation*. He has since stepped down from that post.

**Catherine J. Elliott**, Ph.D., is a Research Associate at the University of Ottawa, Faculty of Education, in Ottawa, Ontario, Canada. Her research interests focus on improving organizational performance through evaluative inquiry, performance management, and organizational learning. In 2008 Dr. Elliott was a doctoral candidate at the University of Ottawa.

**Courtney Amo**, M.A., is Manager, Evaluation, at the Atlantic Canada Opportunities Agency, in Moncton, New Brunswick, Canada. Her research interests focus on knowledge use and evaluation process

use. In 2008, Ms. Amo was employed by the Social Sciences and Humanities Research Council of Canada and completing her master's degree at the University of Ottawa.

**Isabelle Bourgeois**, Ph.D., is a Senior Evaluation Officer with the National Research Council of Canada. Isabelle has worked as an evaluator for the last 10 years, in the private and public sectors. She holds a doctoral degree in measurement and evaluation from the Faculty of Education, University of Ottawa. Her research interests primarily focus on evaluation capacity building in public service organizations.

**Jill Chouinard**, Ph.D., is a lecturer in the Faculty of Education, University of Ottawa, where she recently completed her doctoral degree in educational studies and evaluation. Her areas of research interest are cross-cultural and participatory approaches to program evaluation and educational administration. In 2008 Dr. Chouinard was working on her doctoral research at the University of Ottawa.

**Swee C. Goh** is Professor of Organizational Behaviour and Interis Research Fellow at the Telfer School of Management, University of Ottawa. Professor Goh's main research interests are in the areas of performance measurement in public sector organizations, organizational learning, knowledge management, and evaluation capacity building.

**Robert Lahey** was the founding head of the TBS (Treasury Board of Canada Secretariat) Centre of Excellence for Evaluation, the Canadian government's Evaluation policy centre. He has headed Evaluation units in five departments and agencies over a 32 year career in the Canadian public service. Since establishing his own consulting firm in 2004, he has advised many organizations in Canada and countries around the world on building M&E capacity appropriate to their circumstances. He is a member of the Canadian Evaluation Society (CES) Credentialing Board and was recognized by the CES in 2004 for his contribution to the theory and practice of evaluation in Canada.