Government and voluntary sector differences in organizational capacity to do and use evaluation

CrossMark

J. Bradley Cousins [a,*], Swee C. Goh [b], Catherine Elliott [b], Tim Aubry [c], Nathalie Gilbert [d]

[a] Faculty of Education, Centre for Research on Educational and Community Services (CRECS), University of Ottawa, Canada
[b] Telfer School of Management, University of Ottawa, Canada
[c] School of Psychology, CRECS, University of Ottawa, Canada
[d] Faculty of Education, University of Ottawa, Canada

ABSTRACT

Research on evaluation capacity is limited although a recent survey article on integrating evaluation into the organizational culture (Cousins, Goh, Clark, & Lee, 2004) revealed that interest in the topic is increasing. While knowledge about building the capacity to do evaluation has developed considerably, less is understood about building the organizational capacity to use evaluation. This article reports on the results of a pan-Canadian survey of evaluators working in organizations (internal evaluators or organization members with evaluation responsibility) conducted in 2007. Reliability across all constructs was high. Responses from government evaluators ($N = 160$) were compared to responses from evaluators who work in the voluntary sector ($N = 89$). The former were found to self-identify more highly as 'evaluators' (specialists) whereas the latter tended to identify as 'managers' (non-specialists). As a result, government evaluators had significantly higher self-reported levels of evaluation knowledge (both theory and practice); and they spent more time performing evaluation functions. However, irrespective of role, voluntary sector respondents rated their organizations more favorably than did their government sector counterparts with respect to the antecedents or conditions supporting evaluation capacity, and the capacity to use evaluation. Results are discussed in terms of their implications for evaluation practice and ongoing research.

© 2013 Elsevier Ltd. All rights reserved.

1. Introduction

Increasingly evaluation scholars and practitioners have expressed a significant interest in understanding organizational capacity for evaluation (Barnette & Sanders, 2003; Cousins, Goh, Clark, & Lee, 2004). Conceptually, this interest lies at the confluence of three distinct, yet related, streams of inquiry. First, several authors have focused on evaluation capacity building (ECB): how to develop within organizations the knowledge and skills required to conduct or commission and receive evaluations that are methodologically defensible and relevant to organizational information needs. The most commonly quoted definition of ECB is the one by Stockdill, Baizerman, and Comptom (2002, p. 8): "the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its

use routine." This stream of inquiry has a significant focus on training and development, including understanding beyond that which is represented by program evaluator competencies (Boyle & Preskill, 2007; Nielsen, Lemire, & Skov, 2011; Preskill & Boyle, 2008; Preskill & Russ-Eft, 2005). A second domain of interest is evaluation utilization, which has captured the interests of evaluation scholars for over forty years (Cousins & Leithwood, 1986; Patton et al., 1977). In recent years investigators have turned their attention toward such questions as 'how can evaluation use and influence be differentiated?' (Henry & Mark, 2003; Kirkhart, 2000) and 'what is the relative impact of use of evaluation findings as opposed to use of evaluation processes?' (Cousins, 2007; Patton, 1997). Finally, Cousins and Earl (1995) have connected the consequences of evaluation to organizational development most notably organizational learning (e.g., Cousins & Earl, 1995; Forss, Cracknell, & Samset, 1994; Owen & Lambert, 1995; Preskill, 1994).

Perhaps the most significant benefit of locating organizational evaluation capacity in these streams of inquiry is that conceptions of ECB have broadened from a focus on training and development to considerations about the organizational capacity to use evaluation (Cousins et al., 2004; Stockdill et al., 2002). Conceptually, integrating evaluation into organizational routines and

* Corresponding author at: Faculty of Education, University of Ottawa, CRECS, 136 Jean-Jacques Lussier, Ottawa, ON, Canada K1N 6N5. Tel.: +1 613 562 5800x4036; fax: +1 613 562 5188.
E-mail addresses: bcousins@uottawa.ca (J.B. Cousins), elliott@telfer.uottawa.ca (C. Elliott).

culture is as much about building the capacity to do evaluation as it is to building the capacity to use it. Some authors have focused their attention on the development of organizational ECB tools that allow organizations to compare themselves against a prescribed set of ECB criteria. Such checklists can be used to optimize the long-term capacity to carry out and use program evaluation (Stufflebeam, 2002) and "for incorporating evaluation routinely into the life of an organization" (Volkov & King, 2007, p. 1).

While conceptual advances have been made, research on organizational evaluation capacity is wanting. Much of what is known rests on findings emerging from post hoc narrative reports on evaluation or evaluation capacity building experiences as opposed to planned empirical inquiry using quantitative, qualitative or mixed-methods (Cousins et al., 2004; Cousins, 2007). In this paper we report findings emerging from a broad-based program of research on evaluation with an explicit focus on organizational evaluation capacity. Specifically, we build on a pan-Canadian survey of internal evaluators (Cousins et al., 2008) to examine differences in self-reported understandings of evaluation capacity of members of government and voluntary sector organizations. We were interested in discovering if and how perspectives between these groups differ with respect to the antecedent conditions underlying evaluation practice in organizations, the organizational capacity to do evaluation and the capacity to use it.

We define evaluation as systematic inquiry used in the formulation of judgments about an evaluand's (e.g., program, strategy, innovation) merit, value or significance, or in support of decision making concerning the evaluand. It has long been understood that evaluation takes place within a political context (e.g., Chelimsky, 1995; Mohan & Sullivan, 2006; Weiss, 1973) and that the extent to which evaluation is embraced by organizations will depend in part on organizational information needs and exigencies. Evaluation can serve a hard-nosed, judgmental, accountability function or it can support learning through inquiry about evaluand strengths and weaknesses in an improvement-oriented manner.

Both government and voluntary sectors in Canada and elsewhere have experienced increasing demand for evaluation in recent years. Evaluation and performance measurement are natural elements of new public management systems with a focus on outcomes for society and results-based management. An increasing focus on outcomes in terms of management and funding is characteristic of funding agencies in the voluntary sector as well as public service management. Moreover, the recent legislation of the Federal Accountability Act in Canada as well as the Government Performance and Results Act of 1993 in the United States has stimulated interest in evaluation in the government sector (Government of Canada, 2007). While government and voluntary sector organizations share these information needs, context considerations are quite distinct, the former being large, heavily bureaucratic and openly subject to changes in government priorities through the political process, and the latter being typically small, front-line and subject to the demands of funding agencies (including government, foundations and other non-governmental funding agencies and networks). Understanding the context within which evaluation takes place is essential to understanding how evaluation is integrated into organizational culture. For this reason we are interested in examining the differences between government and voluntary organizations in terms of their capacity to do and use evaluation.

2. Evaluation in government versus not-for-profit organizations

In this section we briefly provide some background on evaluation in government versus that taking place in the voluntary sector. First, we consider the nature of the evaluation function within public management and how the function is managed by government.

2.1. Government and evaluation

Davies, Newcomer, and Haluk (2006) make the case that government has two primary interests in evaluation, namely the promotion of accountability (informing the public, decision makers, taxpayers, service users and other stakeholders about the worth of government policies, programs and interventions) and improving government management as part of good public management. Evaluation, results-based management, performance-based management, and the like are all seen as tools of new public management. In Canada, principles of new public management have been embraced by federal and provincial governments alike (Segsworth, 2005).

Government demand for evaluation arises principally from legislative and executive mandates (Davies et al., 2005) although evaluation through commission systems is also a longstanding focus. Through the establishment of procedures and policies connected with budgetary processes, government executives typically require evaluation as a means of exerting control over agencies. In Canada, similar to the case in the US as portrayed by Datta (2003), government has played a key role in creating the demand for evaluation by requiring that money be set aside for the evaluation of grants and contributions programs and more recently, by developing a federal management accountability framework, within which evaluation is situated.

Governments use a range of evaluation methods within the context of performance measurement (Davies et al., 2005; Gusman, 2005). According to Datta (2003), not only does government use a range of evaluation methods, but indeed it has served to shape the methods that are used. *Impact evaluation* which involves target setting and monitoring is essentially goals-based evaluation. One dimension of impact evaluation would be value-for-money performance management which uses cost-effectiveness analysis and other econometric tools. Impact evaluation relies on a variety of sources of evidence including experimental and quasi-experimental evaluation, systematic reviews and integration or 'meta-analysis' of existing evidence, statistical and econometric modeling. Another major thrust of evaluation in government is *formative* or *process evaluation* where the intent is to inform program and policy development and improvement on the basis of evidence.

While independent institutions such as universities and other private and public agencies have contributed to evaluation in government (Davies et al., 2005), the function is typically managed by auditing agencies, internal evaluation units of government and contract evaluators hired by governments. Globally, governments are increasingly establishing specific units of evaluation "thus bringing evaluation into the very cradle of government structure" (Davies et al., 2005, p. 173). However, Beere (2005) warns that despite organizations' good intentions of establishing internal evaluation units, they run the risk of themselves being shut down unless there is greater internal demand for quality evaluations. Faced with similar challenges as Canada and the US, Beere (2005) laments that Australian internal evaluation units, particularly in the public sector, are chronically faced with a shortage of staff, resources, and evaluation skills. In Canada, by the mid-1990s, program evaluation and audit were becoming less distinct as administrative functions with over 50% of federal departments placing evaluators and auditors in a combined unit, as compared with about one third in 1993 (McDavid & Huse, 2007).

Recent studies have brought into question the commitment to ECB, quality and relevance of evaluation in government in Canada,