

STUDIES ARE NOT ENOUGH: THE NECESSARY TRANSFORMATION OF EVALUATION

John Mayne
Advisor, Public Sector Performance
Ottawa, Ontario

Ray C. Rist
Advisor, Performance Management, Performance Measurement
Washington, DC, USA

Abstract: The authors contend that developments in the public and not-for-profit sectors over the last decade or so have profound implications for the profession of evaluation, implications that are not being adequately addressed. They argue that evaluation needs to transform itself if it wishes to play a significant role in the management of organizations in these sectors. Beyond traditional evaluation studies, evaluators working in public and not-for-profit organizations need to (a) lead the development of results-based management systems, (b) using this and all available evaluative information, strengthen organizational learning and knowledge management, and (c) create analytic streams of evaluative knowledge. Failing to grasp these challenges will result in a marginalized and diminished role for evaluation in public and not-for-profit sector management.

Résumé: Les auteurs soutiennent que les développements survenus dans les secteurs public et à but non lucratif au cours de la dernière dizaine d'années ont provoqué de profonds bouleversements pour la profession de l'évaluation auxquels cette dernière ne s'est toujours pas ajustée. Ils soutiennent que l'évaluation doit s'autotransformer si elle souhaite jouer un rôle important dans la gestion des organisations de ces secteurs. Abstraction faite des études évaluatives traditionnelles réalisées, les évaluateurs internes d'organismes publics et à but non lucratif doivent (a) diriger l'élaboration de systèmes de gestion axés sur les résultats; (b) utiliser ces systèmes ainsi que toute information provenant de l'évaluation pour renforcer l'apprentissage organisationnel et la gestion du savoir; et (c) créer des volets analytiques de la connaissance issue de l'évaluation. Négliger de profiter de l'opportunité de relever ces défis se soldera par une marginalisation et un affaiblissement de la fonction d'évaluation dans la gestion des secteurs public et à but non lucratif.

Corresponding author: John Mayne, 654 Sherbourne Rd., Ottawa, ON, K2A 3H3;
<john.mayne@rogers.com>

INTRODUCTION

The series of articles collected in this issue of the *Canadian Journal of Program Evaluation* express concern over what is perceived as the diminishing of the traditional roles and responsibilities of evaluation — and essentially argue for evaluation to work hard to regain its former stature and lustre. One arena in which evaluators can now demonstrate their relevance (and utility) is seen to be in the broad domain of results-based management (RBM¹). McDavid and Huse (2006) and Young (2006) suggest that evaluation and RBM are mutually compatible and should be complementary, but that this is for the most part not happening. It is as if each has decided to avoid the other's neighbourhood. Were evaluation to engage with RBM and its need for evaluative thinking, it is argued, it would allow evaluation to again demonstrate its focus on being credible, useful, and relevant — the conventional aspirations of 20th-century evaluation.

This article takes a sympathetic view toward the admonitions of McDavid and Huse (2006) and Young (2006), but we believe that there is a need for a broader and more challenging view of the roles and responsibilities that evaluators might assume in this early phase of the 21st century. Succinctly, this article will assert the need for evaluators to become actively involved in leading organizations in developing evaluative services in three crucial areas:

- the development, analysis, and interpretation of results-based management and evidence-based information;
- using this (and other) information to strengthen the organizational learning and knowledge management; and
- creating and building analytic streams of evaluative knowledge.

None of these three areas, from our vantage, appear deeply embedded in the culture of the evaluation communities — whether nationally or internationally. Indeed, we are concerned as to whether they are often even recognized as emerging challenges for evaluators and their communities of practice. We see that all three of these areas are emerging domains for the time and attention of evaluators, merit consideration as areas of work, and will influence the contours and frameworks of future evaluation practice. Stated differently, if evaluation is to be a credible partner in public and not-for-profit management, be it in developed or developing countries, there will need to be a movement away from a singular focus on producing more and more one-off evalu-

ations and toward a synthesis and integration of evaluative knowledge into management practices and policies. It is the challenge of moving from believing the fundamental role for evaluation in organizations is adversarial (throwing bigger and bigger stones) to evaluation as facilitating organizational learning and improvement.

The underlying motif of the articles here suggests a potential (perhaps even real) decline in the utility and perceived relevance of evaluation and evaluators in the public and not-for-profit sectors. McDavid and Huse (2006) note a number of factors contributing to this decline:

- the “paradigm wars” of the 1970s and 1980s led to a fragmentation of the profession,
- evaluation was often oversold in terms of what it could deliver,
- evaluation practice failed to keep up with the transformation that was occurring in the public sector with its focus on timely results information, and
- the growth of de facto evaluation practice by other professions, in particular auditors and accountants.

Echoing these observations, Young (2006) notes

- the lack of timely information delivered by evaluations,
- evaluation not living up to expectations,
- in the Canadian federal government, the linking of evaluation with internal audit,
- the lack of response, too often, by evaluation units to the growing demand for results information.

We agree with these observations and focus on six contributors to the perceived decline in the position of evaluation in many public and not-for-profit organizations:

- Evaluators have often not taken on roles and responsibilities in helping governments at all levels to clarify and strengthen their result-based and performance measurement systems. Many evaluators seemingly want to wash their hands of taking the time to do the dirty work of actual systems construction — building monitoring and evaluation (M&E) systems;
- Evaluators do not see their role as doing the data analysis and reporting on the evidence that comes from M&E systems;

- Evaluators have had little experience in framing and building learning systems for organizations — or at least the lack of evaluation literature in this area would suggest so;
- Evaluators have had little experience in fitting evaluation knowledge into organizational learning systems — as again the absence of the literature indicates;
- Evaluators see their roles and responsibilities as separate from those of knowledge management, witness the absence of any such discussion on knowledge management in three key evaluation textbooks (Bamberger, Rugh, & Mabry, 2006; Fitzpatrick, Sanders, & Worthen, 2004; Rossi & Freeman, 1993) with the notable exception to this omission being Patton (2002);
- Evaluators judge their own and their peers on the quality and sophistication of individual evaluation reports/studies, not on building streams of evaluative knowledge.

Our thesis here is that the potential of evaluation as a significant contributor to public and not-for-profit sector performance depends on reversing these six limitations. Our concern is that the profession is under threat from these emerging realities. For the profession to continue down the path it is now on appears to be a prescription for (sooner or later) marginalization, irrelevance, and diminution.

Three other asides should be made here to help set the context for the comments to follow. First, evaluators are being marginalized in that non-evaluators are now actively involved in undertaking all six of the roles and responsibilities noted above. All six of these areas, we would argue, are absolutely appropriate for evaluators and, in fact, could and should be done by them as the persons with the relevant skills to do so. But it does not appear to be happening in other than infrequent, isolated, and episodic ways. Evaluators for the most part are on the sidelines in three key areas of major public sector and non-profit organizational growth — RBM, knowledge management, and organizational learning. The positions are going to others and evaluators are spectators, watching positions and opportunities for contribution slipping away. (This is all the more ironic as academic positions for evaluators are also diminishing.) Binnendijk (2000) refers to evaluation being crowded out in favour of RBM efforts, using USAID as an example.

Second, and related to the first, evaluators are remaining persistent in putting all their eggs in one antiquated basket — that of producing still more and more of their one-off studies. We contend that

only conducting studies is an increasingly ineffective way of generating, mediating, and disseminating evaluative information. But it is seemingly what evaluators only know how to do and so persevere in doing.

Third, a note on our scope. Evaluation can play a number of roles in a jurisdiction: as a management tool in organizations, as a research tool for understanding interventions in society, and as an accountability tool (Mayne, Divorski, & Lemaire, 1999). Evaluations and experimental designs are likely to continue to be the major measurement tools in research for assessing experimental-type programs. Evaluation can also play a role in accountability, sometimes located in central agencies or in legislatures, as well as in organizations, although here too, Mayne (in press) argues the need for evaluation to refocus if it is to play an effective role in accountability. Our focus in this article is primarily on the management role — on the role evaluation can play in assisting the management of organizations. This is perhaps the largest role evaluation attempts to play around the world and where, in our view, it is most under threat.

What follows now is both a diagnosis and an encouragement — a diagnosis of why evaluation finds itself in the circumstances of its own potential marginalization and an encouragement for a change in behaviour of both evaluators and those responsible for evaluation in organizations so as to stop the downward spiral and begin the slow ascent necessary to restructure and reorient the profession.

A VISION FOR EVALUATION

In our view, evaluation and evaluators should be playing key roles in all aspects of evaluative information in an organization: in building RBM capacity, in managing evaluative knowledge systems, and in creating evaluative information and knowledge, including through the conduct of evaluation studies. We see evaluation as the provider of evaluative services in an organization, with roles and mandates that go far beyond the conduct of individual studies. We use the term “evaluative information” and its counterparts to include empirical information on and about the results of a service, program, or policy. It is meant to cover information from evaluation studies, results-monitoring systems, other one-off results-focused research, and performance studies (Mayne, 2006), as well as evaluative information derived from these various information sources through analysis, aggregation, and synthesis (Rist & Stame, 2006.)

We first present this vision below. We also say that much of it is not all that “new,” in the sense that earlier authors have suggested many of the elements of an expanded evaluation role before. What may be new is pulling all the elements together and doing so in light of the several other reform elements dealing with evaluative information that mature organizations are developing.

The Office of Evaluative Services

In the 21st century, high-performing organizations see themselves and want to be seen as knowledge organizations, using evaluative thinking and inquiry to continuously learn about their environment and how well what they are doing is working. In that light, seeing an evaluation office’s main role as producing a series of one-off, often isolated evaluation studies is an anachronism, the model of a very traditional organization with limited interest in acquiring new knowledge, much less in evaluative thinking. Twenty-first-century organizations are immersed in streams of evaluative information (Rist & Stame, 2006) originating from many sources, both inside and outside the organization. Organizations in the public, private, and non-profit arenas are all trying to manage these evaluative streams in order to be successful.

In this context, an evaluation office within a public or not-for-profit sector organization could see itself as the intellectual and functional leader in evaluative thinking, and as providing a variety of evaluative services, namely:

- *Facilitating evaluative activities.* Assisting managers
 - in designing services, programs, and policies to achieve results,
 - in monitoring and evaluating services, programs, and policies,
 - in reporting performance, and
 - in developing their evaluative capabilities.
- *Facilitating evaluative learning.* Managing processes and events to enhance learning from evaluative information within the organization.
- *Creating evaluative information.* Evaluators should
 - undertake research into useful methodologies and approaches for evaluative activities,
 - synthesize information from evaluations and other relevant sources on how well services, programs, and policies are working,

- oversee evaluative information production, and
- conduct objective evaluation studies.

A number of other authors have spoken to evaluators playing roles beyond the production of good studies. Patton (1997), for example, lists 10 roles evaluators can play depending on the particular situation: judge, auditor, researcher, consultant, facilitator, team member, collaborator, empowerment facilitator, supporter of a cause, and synthesizer. Scheirer (2000) argues for evaluators to play a larger role in performance measurement: “Evaluators could be the ‘navigators’ to help agency managers and administrators get more ‘bang’ for their investments in performance measures ‘bucks’” (p. 147). Wholey (2001) argues similarly. Sonnichsen (2000) discusses the roles of internal evaluators as facilitators in organizations. Mathison (1991) discusses role conflicts for internal evaluators and how to manage the conflicts.

Smith (2001), in introducing the *American Journal of Evaluation* theme issue on the future of evaluation, points to evaluator skills that several authors in that issue thought were needed, including

- strategies for coping with the information revolution (Love, 2001),
- skills for promoting organizational learning (Torres & Pre-skill, 2001), and
- skills for training of organization members in such areas as strategic planning and development of strategic goals (Wholey, 2001).

Here we have tried to pull together a comprehensive list of the evaluative activities those who see themselves as evaluators and who work in evaluation offices might aspire to in light of the transformations underway in public sector and not-for-profit organizations.

FACILITATING EVALUATIVE ACTIVITIES

Public sector and not-for-profit program managers need to engage in a wide variety of evaluative activities in the planning and managing of their services and programs, especially in an environment of results-based management. In particular, managers need to

- articulate the results they expect their programs to achieve (goals/objectives/ planned results),

- understand and explain how they expect to achieve the results (the theory of change of the program),
- develop performance measures and monitor the results their programs are achieving,
- report in a credible manner the performance they are achieving, and
- assess/evaluate their programs.

These are all evaluative activities at which evaluators should be skilled and experienced, while providing value-added assistance to managers. Managers often struggle with undertaking these activities and the benefits of getting professional help would be significant. Newcomer (2001, p. 339) notes that “[t]ranslating performance measurement exercises into organizational learning about program theory requires that managers think like evaluators, and such thinking needs to be groomed and supported.” Not incidentally, the better these activities are carried out by managers, the better the evaluative data and information available on the programs should be, providing a good basis for any future decisions.

Our concern is that evaluators often shun this work, while other professions (such as management consultants, organizational psychologists, operations researchers, auditors, and accountants) seek to build up this line of business. McDavid and Huse (2006) in this thematic segment discuss the reasons for this apparent lack of interest in results-based management evaluative activities by evaluators. Worthen (2001) predicts that evaluators will not embrace strategic planning, to their detriment.

We would see evaluators inside public sector and not-for-profit evaluation units (as either staff or consultants) advising and assisting managers as follows:

- developing results frameworks for programs, that is, setting out expected results, results chains, and results measurement strategies,
- implementing results measurement monitoring and evaluation (M&E) strategies, that is, gathering performance data, analyzing and interpreting that data, and
- reporting on the performance of the organization’s services, programs, and policies.

To be clear, we see the conducting of these tasks as first and foremost the ultimate responsibility of program managers, not the evaluators.

Evaluators are providing professional help. This help could be directly assisting the managers, and/or helping the managers select consultants and assess the quality of the consultants' work in the areas of data collection, analysis, and reporting.

In assisting managers to carry out these activities, evaluators can play a significant role in the organization in developing the managers' evaluative capabilities. Through working with evaluators, learning how to think through the theories of programs, and using the tools of logic models/results chains, managers will learn how to critically assess the designs of their programs, and how to monitor and interpret key results (Julian, 1997; Kusek & Rist, 2004; McDavid & Hawthorn, 2006; Porteous, Sheldrick, & Stewart, 2002; Scheirer, 2000; Wholey, 2001). They should also learn how to present credible performance stories about their programs (Mayne, 2004; McLaughlin & Jordan, 1999).

We are arguing that in assisting in the design and implementation of monitoring systems, evaluators can help others in the organization build up a capacity for thinking about performance, program theory, and performance measures. Young (2006) in this volume reports this happening in the case of Transport Canada. To be effective, this learning process should be an intended goal from the outset. That is, a significant aim of assistance in developing monitoring systems and in conducting evaluations would be to build up the evaluative capacities of the organization. We will speak to resource pressures on evaluation units who undertake this type of facilitating work later. A clear aim is to develop the evaluative skills of managers and staff so that evaluative knowledge is more integrated into the culture and processes of public sector organizations, enabling managers to more accurately and completely answer questions concerning the performance of their organizations to the citizens.

Challenges

Evaluators face a number of challenges in effectively facilitating evaluative activities in a public or not-for-profit sector organization. First of all, they need to be seen as value-added sources of credible assistance. That is, evaluators may not be seen by some managers as a good resource. Managers may feel that the evaluation unit does not have the needed skills and experience, may not be available when needed, or may not be helpful people to work with. In our own experience, we have known public sector managers who see evaluators as naive, unable to understand the political and institutional ramifica-

tions of their evaluation reports and recommendations. Taylor and Balloch (2005) and McDavid and Hawthorn (2006) have addressed these concerns in their writings.

Evaluators might consider adopting a “just in time” mentality. That is, much of the work in facilitating evaluative activity will need to be timely, done when needed by managers, and done on time. For some, this may be quite a different work environment from that which they are accustomed to when conducting a six-month or longer evaluation study.

Young (2006) in this volume discusses several of these challenges and ways to deal with them. (See also Torres [1991] and Sonnichsen [2000].) Evaluators, as with any other group of professionals seeking to have their voice heard in the political and bureaucratic arenas, might begin thinking about how to develop a deliberate strategy to get accepted and demonstrate their added value, perhaps taking on the evaluation of pilot projects. The unpacking of theories of change and logic models of programs is an obvious time to ensure that managers are coached in evaluative thinking and the benefits of deploying these and other evaluation tools.

A different kind of challenge here is that of who becomes responsible for the program design and monitoring system when evaluators have been involved, and the related issue of the objectivity of future evaluation work. In the roles being discussed here, evaluators are assisting managers as advisors and facilitators. Managers are responsible for their own programs, for the monitoring systems designed to help them manage, and for reports they prepare on performance. They remain responsible whether or not they have had help from the evaluation unit or have hired outside consultants to help them.

The question of objectivity is more serious. Evaluators and the organization may be concerned that if the evaluators have had a major hand in designing aspects of the program and its monitoring system, a subsequent evaluation of the program might not be adequately independent. To some extent evaluators may be seen as evaluating their own work and finding ways to defend their own efforts. There are several points to note here:

- A monitoring system should produce data on various aspects of a program’s performance over time. The managers would be in charge of actual data collection and initial analysis. Part of any evaluation is to assess the robustness of existing

data and information. If certain performance measures turn out to be of less interest or if the collection or analysis of the data is suspect, these facts can be taken into account in the evaluation.

- A good monitoring system will evolve over time as its usefulness is assessed by managers; it may be somewhat different from that initially put in place with the assistance of evaluators.
- The design of the evaluation should not in any event be in the hands of the evaluators alone. The organization and its senior management would normally be involved in setting out the specific terms of reference, including the evaluation questions to be addressed, for the evaluation to reflect the context, concerns, and realities of the time.
- The ethical culture in the organization may be such that professional evaluators do value their objectivity and would want to behave in such a manner.
- Ex-post evaluation studies generally take place several years after an intervention has been completed. If objectivity is seen as an issue, the evaluation unit can ensure different evaluators are involved in the ex-post evaluation, or could hire external evaluators to carry it out.
- As noted by Young (2006), evaluators can make it clear in an organization that there are different clients for the different services they provide.

In supporting managers with credible and trustworthy evaluative information, evaluators do need to keep this objectivity issue in mind. It has the potential of undermining the credibility of current as well as future evaluations, especially since it may be the appearance of objectivity that is at issue, rather than any real concern about objectivity. Evaluators should be aware of previous work they have done with the program and assess how best to counter any bias or appearance of bias that might arise.

Ideally, managers should also have the same concerns. Part of the aim of developing managers' evaluative thinking skills is to enable them to see the value of unbiased, objective evaluative knowledge.

FACILITATING EVALUATIVE LEARNING

The issue of the use of evaluation or evaluative information has a large literature (see, for example, Cousins, Goh, Clark, & Lee, 2004;

Feinstein, 2002; Leviton, 2003; Mark & Henry, 2004; Sridharan, 2003; Weiss, 1998). How to ensure better use of such information remains a key question, as does what one means by “use” (Caracelli & Preskill, 2000; Marra, 2004; Valovirta, 2002; Vedung, 1997). A related and growing literature is looking at the use of research more generally in the policy domain, and in particular in the evidence-based policy agenda (Kusek & Rist, 2004; Learmont & Harding, 2006; Nutley, Walter, & Davies, 2003; Pawson, 2002a, 2002b; Sanderson, 2002; Walter, Nutley, & Davies, 2005).

What has been clear for some time is that simply making evaluative information available is unlikely to ensure its use. Distributing evaluation reports or, say, monthly performance monitoring reports to decision makers is not nearly enough. Forss and Rebien (2006, p. 246) note that “the evaluation reports as such played rather marginal roles. The important learning events occur in preparatory meetings, in procedures of data collection, through joint analysis of achievement, and in the search for data from other sources, benchmarking, and comparative analysis, etc.” We will not get into the whole utilization debate here, but rather look at ways an evaluation unit can facilitate learning from evaluative information.

Levin-Rozalis and Rosenstein (2005) discuss how evaluators can play useful roles in the process of organizational learning. They argue for a participatory form of evaluation and, importantly, that for evaluators to engage in organizational learning, changes are required in the normal evaluation process. The organization has to set aside the time and resources to engage in learning, and the evaluators need different skill sets emphasizing being a “knowledge facilitator” responsible for the “procedures of learning” (p. 95). Torres and Preskill (2001) discuss the challenges facing evaluators in developing an organizational learning approach to evaluation, noting that the more “actively-involved change agent” role it implies may be uncomfortable for many evaluators. They present an incremental approach to adopting such a role in an organization. Cousins et al. (2004) empirically review efforts to integrate evaluation into organizational learning. In a book edited by Leeuw, Rist, and Sonnichsen (1994), comparative perspectives on evaluation and organizational learning are discussed.

So the idea of linking evaluation and organizational learning is not new, but would seem to be an effective way to enhance the use of evaluative information in organizations. We discuss two aspects of this role: holding learning events; and managing evaluative informa-

tion, by working to ensure that evaluative knowledge is part of the policy cycle.

Learning Events

Moynihan (2005) discusses the prevalence of a lack of organizational learning from existing information and suggests the need for structured “learning events.” The idea is that an organized event is scheduled for discussion on a programming issue of interest and the relevant evaluative information is made available in a digestible form. A key point here is to create a scheduled and planned event that has a good chance to encourage learning: learning either in the sense of resulting in decisions on issues or in enhancing knowledge and understanding about an issue. Barrados and Mayne (2003) similarly suggest the need for organizational learning mechanisms to foster results-based learning in organizations. They argue, “There is a need to institutionalise learning processes within a public sector organization” (p. 88).

One example of a learning event is the Annual Learning Forum held at the International Development Research Centre (IDRC) (2006). The evaluation unit organizes the event and brings together recent evaluative information from evaluations and project reports pertinent to the issue being discussed. “The ALF is not a time for planning or decision-making.... The function of the forum ... is to find out what is already known but has yet to be shared” (p. 2). At IDRC, learning from evaluations is seen as a key aim of the organization (Love, 1996). The evaluation unit at the World Bank holds twice yearly learning events on key evaluations that have been completed.

The need for learning events in a public or not-for-profit sector organization would seem compelling. It would also seem clear that the evaluation unit is in a good position to lead and manage such events, given its presumed ability to synthesize, summarize, and effectively display evaluative knowledge from a variety of sources.

Managing Evaluative Information

Many organizations now have some form of a knowledge management regime (Birkinshaw, 2001; Nonaka & Konno, 1998; Patton, 2001; Saussois, 2003). Part of building a knowledge management base in a high-performing public sector organization would entail including evaluative information and knowledge.

In order for evaluative information to be used, evaluative knowledge in an organization needs to be considered at each stage of the knowledge cycle: from creation to mediation (i.e., aggregating and synthesizing evaluative knowledge) to dissemination. There is also a need to link evaluative knowledge to the policy cycle — agenda setting, formulation, implementation, and accountability (Rist, 2000). The knowledge cycle and the policy cycle come together at the point of evaluation utilization.

For the evaluation unit, the issue is how to best manage evaluative knowledge so that it is disseminated to appropriate policy and decision makers at the right time and with the right format in order to address the right questions. To do this, the evaluation unit has to (a) know the appropriate information needs in the respective phases of the policy cycle; (b) know about policy and programming work planned and underway; (c) know which sets of program managers, stakeholders, and policy makers require what types of evaluative information at each phase of the cycle; and (d) be able to facilitate the movement of knowledge and evaluative lessons.

Many key questions asked in the policy cycle are evaluative in nature, and hence ought to be questions that evaluators can assist in answering. The following list is illustrative of the kinds of evaluative questions one might find applicable to the policy cycle. In policy formulation, evaluative questions, for example, are

- What do we know about the problem? (What has changed over time? What is the presumed cause? What aspects have changed? Can it be measured?)
- What has taken place previously? (What projects/programs? How long did they last? In what context were they successful? Did they address the same problem?)
- What is known about previous efforts? (What are the time and cost to get outcomes? What are the institutional capacities to deliver?)
- Is the program logic coherent and supported by clear evidence?

In policy implementation, the evaluative questions are

- Is the logic of the policy theory of change model clearly understood at all administrative levels?
- Is the implementation plan clearly understood at all administrative levels?

- Is the problem still as it was when the policy was first formulated?
- Does the institutional capacity exist to deliver?

In terms of policy accountability, the evaluative questions are

- What was accomplished? What did the program cause?
- Did original objectives stay in place throughout implementation?
- Was there proper management?
- How strong was the theory of change for the program?
- Is there evidence of value for money in the accomplishments of the policy?

Clearly, the potential for evaluators contributing to the policy cycle is large. But they do face challenges in realizing that potential.

Challenges

In order to be seen as a facilitator for evaluative learning, the evaluation unit would have to demonstrate that it can do much more than deliver evaluation studies. It would have to demonstrate not only that it is aware of the larger evaluative streams of knowledge at play, but that it is managing, codifying, and synthesizing those streams: that is, that the unit can use the evaluative information available to deliver relevant evaluative knowledge to the organization at learning events or as input to the various phases of the policy cycle.

The key challenge, of course, is getting a seat at the policy table. Evaluators would have to know of the policy initiatives underway in their organization — to know and perhaps be able to influence the policy agenda. They would need to work to gain acceptance to and knowledge of the timing of upcoming policy discussions. Once aware of the policy agenda, the evaluation unit could hold a learning event on a current issue or contribute more directly to policy making. Young (2006) notes in her case study the importance of the evaluators in Transport Canada getting involved in the policy cycle.

Evaluators would have to package relevant evaluative knowledge — from results-monitoring systems, evaluations, and other results studies — and be prepared to discuss its credibility and its implications. They may indeed have to create new evaluative information for the specific questions being asked at the moment (see below). What

would not be useful would be to simply table a completed evaluation study on the general issue, expecting decision makers to read the study, interpret its findings for the issues at hand, understand the strengths and weaknesses of the study, and know how the study fitted into the larger stream of evaluative knowledge the organization likely has on the issue.

CREATING EVALUATIVE INFORMATION

We have left to the end the discussion about evaluation's traditional role of conducting single, often lengthy evaluation studies. We are in no way suggesting that this role is not important. Rather we are arguing that there are many other roles an evaluation should be playing in an organization today. And even here, we introduce three roles over and above undertaking evaluations: undertaking research into improved evaluation and monitoring methods and approaches, synthesizing the findings in existing evaluations and related research studies, and overseeing the production of evaluative information.

Research on Evaluation

Evaluators are continually seeking more effective ways of measuring and analyzing the performance of programs. Often new ideas and approaches come from academics who are researching evaluation methodologies, and evaluators in public and not-for-profit sector organizations can often make good use of these sources. Good practices in expanding and strengthening survey methodologies come to mind. But relying solely on academic sources for better practices may not always be sufficient:

- The issues of concern to practicing evaluators may be quite simple or even mundane from an academic standpoint, and hence may not get much attention.
- The specific issues addressed may have received very little attention in the academic literature.
- On the contrary, the issue may have been the subject of considerable conflicting research and be difficult to use.

In these and other cases, the evaluation unit may be well advised to undertake its own research on the specific issues of concern. Not only could this effort provide for undertaking better evaluations, but would also be useful for assessing approaches recommended by out-

side evaluators. In our experience, explicit research is not often seen as a legitimate task in internal evaluation groups — with the notable exception of the now defunct “Program Evaluation and Methodology Division” (PEMD) of the United States Government Accountability Office (USGAO.)

Synthesizing Evaluative Information

For most public sector agencies and ministries, there have been numerous evaluations and research studies conducted over many years. There is now a large and ever-growing body of evaluative information on government interventions such as can be found at

- the Campbell Collaboration on social, behavioural, and education interventions (<http://www.campbellcollaboration.org/>)
- the Cochrane Collaboration on health care (<http://www.cochrane.org/>)
- the Evidence for Policy and Practice Information and Co-ordinating Centre on health, education, and welfare (<http://eppi.ioe.ac.uk/cms/>).

In addition, most organizations have been conducting evaluations for many years now, and often continue to only undertake still another one-off evaluation study without trying to assess what has been learned about similar interventions in the past. Perrin (2006) discusses the shortcomings of individual evaluation studies and the need for making better use of past evaluations through the accumulation of knowledge. He speaks to “creating knowledge through synthesis” (p. 25) and discusses three approaches to synthesis — cluster evaluation, synthesis of existing studies, and integrative reviews — as well as approaches to identifying good practices from previous evaluative work. Pawson (2002a, 2002b), among many others, also discusses approaches to synthesizing findings from disparate studies.

Overseeing Evaluative Information Production

In an organization, evaluative information is produced not only by the evaluation unit, but also by managers through their results monitoring and their own program-led evaluations. The robustness of this information and its credibility will likely (and should) vary with its intended use. Where the information is intended to play a critical role in important decisions or will be communicated outside

the organization, the organization will likely want to ensure that the information is and is seen as reasonably credible. If the intended use, alternatively, is a standardized management report for internal consumption only, the requirements of methodological rigour might well be less. These types of considerations might, for example, suggest whether an evaluation is program-led for internal management use or is conducted by the evaluation unit, presuming some level of external dissemination.

Where the credibility of the evaluative information is an issue, the evaluation unit can play a useful quality control role, either during the production of the information or by assessing the quality of the evaluative information that has been produced. An example of this role in the Canadian federal government is the requirement that the head of the evaluation unit “sign off” on the results frameworks produced by program managers and destined for external use (Treasury Board Secretariat, 2005). Similarly, in some departments, the quality of program-led evaluations is assessed by the evaluation unit (Health Canada, 2006; IRDC, 2002).

Conducting Evaluation Studies

Yes, evaluation units still need to carry out strategic evaluation studies of key services, programs, and policies of their organizations. This is their traditional role, and such evaluative activity is needed, providing an independent assessment of the performance of organizational activities. These studies would complement other evaluative activities in the organization, often filling in gaps in knowledge and understanding on specific issues of concern. Our concern all along has been seeing such one-off studies as an end in themselves. Perpetuating such a model is arcane and clearly diminishes the potential contributions of the information contained in each of them.

We note that through the process of carrying out evaluation studies, managers will further learn what evaluative thinking is all about and how to make best use of evaluations. As we noted earlier in this article, the idea of learning from the process of evaluation is not new and has been addressed by many. Forss, Rebien, and Carlson (2002), Patton (1997, 1998), Preskill, Zuckerman, and Matthews (2003), Leeuw et al. (1994), and others have argued for the evaluative learning acquired in the process of being involved in the conduct of an evaluation study. The idea is that by deliberately involving program managers and staff in the evaluation process, they will learn about the con-

ceptualization and conduct of evaluation, building their evaluative thinking, skills, and capacities.

Challenges

The evaluation unit should face no new challenges in conducting evaluation studies, since this is what the unit is assumed to do. Challenges will be those well discussed in the literature of acquiring the needed evaluation skills, designing good evaluations, ensuring relevant findings, getting adequate funding for robust evaluations, and so on, and ensuring that quality evaluations are conducted. Indeed, performance and monitoring information should provide a rich source of useful data for evaluation studies. If evaluators have been involved in the development and analysis of performance information, they would then be well aware of the strengths and weaknesses of such data.

In the other three ways of creating evaluative knowledge discussed above, however, there will be a number of challenges:

- convincing the organization it should invest in research and synthesis activities — getting funding for undertaking research on evaluation methodologies or synthesis of existing evaluations and related research studies,
- convincing the organization of the need in some cases for an oversight function on the production of evaluative information,
- acquiring the skills needed to undertake the research and/or synthesis, and
- ensuring that any research on methods or synthesis is seen by the organization as worthwhile.

CONCLUSION

The members of an evaluation unit in a public or not-for-profit sector organization may presume that they will face no new challenges to their mandate to conduct evaluation studies, since this is what they and others will assume the unit is meant to do. In his review of the state of evaluation in the Canadian federal government, Segsworth (2005, p. 193) concludes that “evaluations in the Government of Canada from 1977 to 2005 suggest that very little in the way of fundamental change to the function has taken place.” It is our view

that the challenges outlined in this article will emerge for the public sector evaluation community sooner rather than later, and the mandate to change, if not the requirements of professional survival, will be evident. Essentially, our message to those in public and not-for-profit sector evaluation units is that you can run but you cannot hide from the coming transformations of public and not-for-profit sector organizations. This is, in fact, our message to the larger evaluation community as well. McDavid and Huse (2006) in this issue note that over the past three decades, evaluation has been transformed by its own ranks. We are suggesting significant further transformation is required and that it needs to be and can be — as Young's article (2006) demonstrates — led by evaluators.

POSTSCRIPT

We have deliberately written this article to raise issues that we believe will challenge the evaluation community. We want to challenge the mindset of evaluation studies being the be-all and end-all of evaluation work. We want to challenge the evaluation community and its seeming refusal to engage in professional partnerships and collaborations with those who work in knowledge management or organizational learning and those who build monitoring systems. We want to challenge the contentment of the evaluation community — even as others come in, encroach on the domain of evaluation, take away jobs, and redefine the roles and responsibilities of what can be done by evaluators. (We do not believe the present interest in credentials for evaluators will be enough to stem the tide.) And, finally, we want to challenge the evaluation community to awaken to the fact that the discipline appears to be on the verge of real decline.

No discipline can presume to last just because it has all the accoutrements of a discipline — journals, meetings, training, and so on. Disciplines do decline: witness what has happened to sociology in the past 20 years as chronicled, for example, by Irving Louis Horowitz in his book, *The Decomposition of Sociology* (1994). (Note that he did not use the word “decline” in the title, but a word with more finality: “decomposition.”)

So this article is not written as a manifesto. It is written to open up a conversation with our fellow evaluators on where we are as a discipline and as a profession, what we are doing at present that is either functional or dysfunctional, and where we are going. What do we think the discipline will be like in 10 or 15 years? Are we prepar-

ing ourselves, our students, our universities, and our public sector organizations for this change? (It is not an irrelevant question — 10 years is but 120 months away!)

There are consequences for evaluation as a discipline and evaluators as professionals to stay largely as we are — and as we believe we have accurately described. There are also consequences for embracing change — changing the assumptions and reward structures, changing the work environment, and changing the ways in which we engage (or not) other disciplines. The debate is whether what we portray as the captivity of evaluation to one-off individual studies and all that it implies is accurate — and if it is, what are the implications?

While a conversation in the discipline and professional ranks would need some format and structure, there are several issues beyond those noted in this article that we can briefly allude to here (and have noted implicitly elsewhere in the article) as important to the discourse. One is about resources/branding and the other is about the ontological question of what is an “evaluator.” Let us take these two issues in turn.

First, it is clear from the Young (2006) article in this issue and the experiences of evaluators working in public sector evaluation units, that trying to do most (not even all!) of what is described here would take more resources than are conventionally allocated to the evaluation function. What we are outlining here is an ambitious agenda of expanding the evaluation function — to the point where studies are perhaps even a minor part of what the unit does overall. Part of this expansion is pushing back on other disciplines and units in organizations that are poaching on evaluation — especially audit and internal control offices as well as knowledge management and organizational learning offices.

The other part of this expansion is to take on new ground and new responsibilities to bring evaluation to the policy table with a broader and more policy-sensitive set of analyses, presumably in close to real time. The resources for these actions and activities will not fall out of the sky. They will have to be earned — and if the budget process is zero-sum, then they have to be earned by demonstrating more relevance, more timeliness, and more accuracy than other offices in the organization. Evaluation has given up much ground to others in the public sector. It is going to have to claw its way back if it wants to be a viable player at the table, now and in the future.

In organizations, evaluators are subject to the policies and culture of the organization. We recognize they are not free to do as they please. Seeing the need to transform evaluation and bringing it about are different challenges. Evaluators will have to demonstrate that the changes they want to implement can be made and establish their credibility to do so, sometimes in the face of competing professionals. Here the need for strong and insightful leadership from those managing evaluation units and those setting overall evaluation policy is clear.

Second, it is clear that we believe the traditional definition of an evaluator is outmoded. Producing single study after study is no longer relevant and useful as the end product. All of the possible roles discussed earlier (including the 10 noted by Patton) imply some changed set of expectations and responsibilities from that of the producer of individual studies. We see the evaluator essentially becoming an agent of change — of helping to frame the thinking and activities of an organization most appropriately and in a focused way to bring about success. Evaluators have to think clearly about institutional, political, and behavioural change — how it is defined and achieved within the organizational and political constraints of the agency in question. Coming to the table and saying that the evaluator might have some evidence from some study in two years is just no longer viable. He or she will have to come to the table with evidence of what works elsewhere, what theory of change is plausible in the present circumstances, and what strategy of change is feasible in the organization. It is with a deft hand that the evaluator helps to inform (even lead) the process — not just sit back and later survey the damage and count the wounded.

So, we seek engagement and dialogue. And, might we add, it should not be even primarily or necessarily with the two of us — it needs to be within the profession and among the professionals. We all have a stake in the discipline and in our own professional status.

NOTE

1. A variety of terms are used for management approaches that focus on the results (outputs and outcomes) that are aimed for: results management, performance management, managing for results, managing for outcomes, and (the term we shall use) results-based management (RBM).

REFERENCES

- Bamberger, M., Rugh, J., & Mabry, L. (2006). *Real world evaluation*. Thousand Oaks, CA: Sage.
- Barrados, M., & Mayne, J. (2003). Can public sector organizations learn? *OECD Journal on Budgeting*, 3(3), 87–103.
- Binnendijk, A. (2000). *Results-based management in the development cooperation agencies, a review of the experience: An executive summary*. Paris: OECD-DAC Working Party on Aid Evaluation.
- Birkinshaw, J. (2001). Making sense of knowledge management. *Ivey Business Journal*, 65(4), 32–36.
- Caracelli, V.J., & Preskill, H. (Eds.). (2000). *The expanding scope of evaluation use*. New Directions for Evaluation, No. 88. San Francisco: Jossey-Bass.
- Cousins, 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.
- Feinstein, O.N. (2002). Use of evaluations and the evaluation of their use. *Evaluation*, 8(4), 433–439.
- Fitzpatrick, J.L., Sanders, J.R., & Worthen, B.R. (2004). *Program evaluation: Alternative approaches and practical guidelines*. Boston: Pearson Education.
- Forss, K., & Rebien, C.C. (2006). Evaluation, knowledge management, and learning: Caught between order and disorder. In R.C. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems* (pp. 235–250). New Brunswick, NJ: Transaction.
- Forss, K., Rebien, C.C., & Carlson, J. (2002). Process use of evaluations. *Evaluation*, 8(1), 29–45.
- Health Canada. (2006). Departmental performance measurement and evaluation directorate. Retrieved December 15, 2006, from <http://www.hc-sc.gc.ca/ahc-asc/branch-dirgen/cfob-dgcm/dpmed-dmerm/index_e.html>.

- Horowitz, I.L. (1994). *The decomposition of sociology*. Oxford: Oxford University Press.
- IDRC. (2002). *Quality assessment of IDRC evaluation reports. Evaluation guidelines, No. 4*. Ottawa: Author. Retrieved August 28, 2006, from <<http://www.idrc.ca/uploads/user-S/115644981714guideline-web.pdf>>.
- IDRC. (2006). *A contemplative recess: IDRC's annual learning forum. Evaluation highlights*. Ottawa: Author. Retrieved August 20, 2006, from <http://www.idrc.ca/uploads/user-S/11424537631Highlight_7.pdf>.
- Julian, D.A. (1997). The utilization of the logic model as a system level planning and evaluation device. *Evaluation and Program Planning*, 20(3), 251–257.
- Kusek, J.Z., & Rist, R.C. (2004). *Ten steps to a results-based monitoring and evaluation system*. Washington, DC: World Bank.
- Learmont, M., & Harding, N. (2006). Evidence-based management: The very idea. *Public Administration*, 8(2), 245–266.
- Leeuw, F., Rist, R., & Sonnichsen, R. (Eds.). (1994). *Can governments learn? Comparative perspectives on evaluation and organizational learning*. New Brunswick, NJ: Transaction.
- Levin-Rozalis, M., & Rosenstein, B. (2005). The changing role of the evaluator in the process of organizational learning. *Canadian Journal of Program Evaluation*, 20(1), 81–104.
- Leviton, L.C. (2003). Evaluation use: Advances, challenges and applications. *American Journal of Evaluation*, 24(4), 525–535.
- Love, A. (1996). Knowledge for empowerment: A visit to the evaluation unit of the International Development Research Centre, Ottawa, Canada. *Evaluation*, 2(3), 349–361.
- Love, A.J. (2001). The future of evaluation: Catching rocks with cauldrons. *American Journal of Evaluation*, 22(3), 437–444.
- Mark, M., & Henry, G. (2004). The mechanisms and outcomes of evaluation influence. *Evaluation*, 10(1), 35–57.
- Marra, M. (2004). The contribution of evaluation to socialization and externalization of tacit knowledge: The case of the World Bank. *Evaluation*, 10(3), 263–283.

- Mathison, S. (1991). Role conflicts for internal evaluation. *Evaluation and Program Planning*, 14, 173–179.
- Mayne, J. (2004). Reporting on outcomes: Setting performance expectations and telling performance stories. *Canadian Journal of Program Evaluation*, 19(1), 31–60.
- Mayne, J. (2006). Performance studies: The missing link? *Canadian Journal of Program Evaluation*, 21(2), 201–208.
- Mayne, J. (in press). Evaluation for accountability: Reality or myth? In M.-L. Bemelmans-Videc, J. Lonsdale, & B. Perrin (Eds.), *Making accountability work: Dilemmas for evaluation and for audit*. New Brunswick, NJ: Transaction.
- Mayne, J., Divorski, S. & Lemaire, D. (1999). Locating evaluation: Anchoring evaluation in the executive or the legislature, or both or elsewhere? In R. Boyle & D. Lemaire (Eds.), *Building evaluation capacity: Lessons from practice* (pp. 23–52). New Brunswick, NJ: Transaction.
- McDavid, J.C., & Hawthorn, L.R.L. (2006). *Program evaluation & performance measurement: An introduction to practice*. Thousand Oaks, CA: Sage.
- McDavid, J.C., & Huse, I. (2006). Will evaluation prosper in the future? *Canadian Journal of Program Evaluation*, 21(3), 47–72.
- McLaughlin, J.A., & Jordan, G.B. (1999). Logic models: A tool for telling your program's performance story. *Evaluation and Program Planning*, 22, 65–72.
- Moynihan, D.P. (2005). Goal-based learning and the future of performance management. *Public Administration Review*, 65(2), 203–216.
- Newcomer, K.E. (2001). Tracking and probing program performance: Fruitful path or blind alley for evaluation professionals. *American Journal of Evaluation*, 22(3), 337–341.
- Nonaka, I., & Konno, N. (1998). The concept of “ba”: Building a foundation for knowledge creation. *California Management Review*, 40(3), 40–54.
- Nutley, S., Walter, I., & Davies, H. (2003). From knowing to doing: A framework for understanding the evidence-into-practice agenda. *Evaluation*, 9(2), 125–148.

- Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text*. Thousand Oaks, CA: Sage.
- Patton, M. (1998). Discovering process use. *Evaluation*, 4(2), 225–233.
- Patton, M.Q. (2001). Evaluation, knowledge management, best practices, and high quality lessons learned. *American Journal of Evaluation*, 22(3), 329–336.
- Patton, M.Q. (2002). *Qualitative research and evaluation methods*. Thousand Oaks, CA: Sage.
- Pawson, R. (2002a). Evidence-based policy: In search of a method. *Evaluation*, 8(2), 157–181.
- Pawson, R. (2002b). Evidence-based policy: The promise of “realist synthesis.” *Evaluation*, 8(3), 340–358.
- Perrin, B. (2006). How evaluation can help make knowledge management real. In R.C. Rist & N. Stame (Eds.), *From studies to streams: Managing evaluative systems* (pp. 23–45). New Brunswick, NJ: Transaction.
- Porteous, N.C., Sheldrick, B.J., & Stewart, P.J. (2002). Introducing program teams to logic models: Facilitating the learning process. *Canadian Journal of Program Evaluation*, 17(3), 113–141.
- Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation*, 24(1), 423–442.
- Rist, R. (2000). Influencing the policy process with qualitative research. In N.K. Denzin & Y.S. Lincoln (Eds.), *Handbook of qualitative research* (2nd ed., pp. 1001–1017). Thousand Oaks, CA: Sage.
- Rist, R., & Stame, N. (Eds.). (2006). *From studies to streams: Managing evaluative systems*. New Brunswick, NJ: Transaction.
- Rossi, P.H., & Freeman, H.E. (1993). *Evaluation: A systematic approach*. Thousand Oaks, CA: Sage.
- Sanderson, I. (2002). Evaluation, policy learning and evidence-based policy making. *Public Administration*, 80(1), 1–22.
- Saussois, J.-M. (2003). Knowledge management in government: An idea whose time has come. *OECD Journal on Budgeting*, 3(3), 105–120.

- Scheirer, M.A. (2000). Getting more “bang” for your performance measures “buck.” *American Journal of Evaluation*, 21(2), 139–149.
- Segsworth, R.V. (2005). Program evaluation in the Government of Canada: Plus ça change ... *Canadian Journal of Program Evaluation*, 20(3), 175–197.
- Smith, M.F. (2001). Evaluation: Preview of the Future #2. *American Journal of Evaluation*, 22(3), 281–300.
- Sonnichsen, R. (2000). *High impact internal evaluation*. Thousand Oaks, CA: Sage.
- Sridharan, S. (2003). Introduction to special section on “What is a useful evaluation?” *American Journal of Evaluation*, 24(4), 483–487.
- Taylor, D., & Balloch, S. (Eds.). (2005). *The politics of evaluation: Participation and policy implementation*. Bristol, UK: Policy Press.
- Torres, R.T. (1991). Improving the quality of internal evaluation: The evaluator as consultant-mediator. *Evaluation and Program Planning*, 14, 189–198.
- Torres, R.T., & Preskill, H. (2001). Evaluation and organizational learning: Past, present and future. *American Journal of Evaluation*, 22(3), 387–395.
- Treasury Board Secretariat. (2005). Preparing and using results-based management and accountability frameworks. Ottawa: Author. Retrieved on December 15, 2006, from <http://www.tbs-sct.gc.ca/eval/pubs/RMAF-CGRR/guide/guide_e.pdf>.
- Valovirta, V. (2002). Evaluation utilization as argumentation. *Evaluation*, 8(1), 60–80.
- Vedung, E. (1997). *Public policy and program evaluation*. New Brunswick, NJ: Transaction.
- Walter, I., Nutley, S., & Davies, H. (2005). What works to promote evidence-based practice? A cross-sector review. *Evidence & Policy: A Journal of Research, Debate and Practice*, 1(3), 335–364.
- Weiss, C.H. (1998). Have we learned anything new about the use of evaluation? *American Journal of Evaluation*, 19(1), 21–33.
- Wholey, J. (2001). Managing for results: Roles for evaluators in a new management era. *American Journal of Evaluation*, 22(3), 343–347.

- Worthen, B.R. (2001). Whither evaluation? That all depends. *American Journal of Evaluation*, 22(3), 409–418.
- Young, G. (2006). Evaluation can cross the boundaries: The case of Transport Canada. *Canadian Journal of Program Evaluation*, 21(3), 73–92.

John Mayne is an independent advisor on public sector performance. He has authored numerous articles and reports, and edited five books in the areas of program evaluation, public administration, and performance monitoring. In 1989 and in 1995, he was awarded the Canadian Evaluation Society *Award for Contribution to Evaluation in Canada*. In 2006, he became a Canadian Evaluation Society Fellow.

Ray C. Rist has recently (December 2006) retired as a senior evaluation officer of the World Bank. He is now serving as an advisor on performance management and performance measurement to governments and organizations worldwide. Dr. Rist is also the co-director of the International Program for Development Evaluation Training (IPDET) in Ottawa, Canada, and chair of the 22-nation International Evaluation Research Group (IERG.) Ray Rist has authored or edited 25 books, published more than 135 articles, and lectured in 76 countries.