INTEGRATING EVALUATIVE INQUIRY INTO THE ORGANIZATIONAL CULTURE: A REVIEW AND SYNTHESIS OF THE KNOWLEDGE BASE

J. Bradley Cousins
Swee C. Goh
Shannon Clark
University of Ottawa
Ottawa, Ontario

Linda E. Lee
Proactive Information Service Inc.
Winnipeg, Manitoba

Abstract: The purpose of this article is to explore, through an extensive review and integration of recent scholarly literature, the conceptual interconnections and linkages among developments in the domains of evaluation utilization, evaluation capacity building, and organizational learning. Our goal is to describe and critique the current state of the knowledge base concerning the general problem of integrating evaluation into the organizational culture. We located and reviewed 36 recent empirical studies and used them to elaborate a conceptual framework that was partially based on prior work. Methodologically, our results show that research in this area is underdeveloped. Substantively, they show that organizational readiness for evaluation may be favourably influenced through direct evaluation capacity building (ECB) initiatives and indirectly through doing and using evaluation. We discuss these results in terms of an agenda for ongoing research and implications for practice.

Résumé: Cet article a pour but d'explorer, grâce à un examen poussé et à l'intégration de la littérature académique récente, les interconnexions et liens conceptuels entre les faits nouveaux dans les domaines de l'utilisation de l'évaluation, du renforcement des capacités d'évaluation, et de l'apprentissage organisationnel. L'objectif est de décrire et de critiquer l'état actuel de la base de connaissances en ce qui a trait au problème général de l'intégration de l'évaluation à la culture organisationnelle. Nous avons identifié et examiné 36 études empiriques récentes utilisées pour

Corresponding author: B. Cousins, Faculty of Education, University of Ottawa, 145 Jean Jacques Lussier, Ottawa, ON Canada, K1N 6N5; <bcousins@uottawa.ca>

élaborer un cadre conceptuel partiellement fondé sur des travaux antérieurs. Du point de vue méthodologique, les résultats démontrent que la recherche dans ce domaine est encore sous-développée. Du point de vue «validation», ils démontrent que l'état de préparation de l'organisation à l'évaluation peut être favorablement influencée directement par des initiatives de renforcement des capacités d'évaluation (RCE) et indirectement par l'exécution et l'utilisation d'évaluations. La discussion propose un plan pour la recherche continue et identifie les implications pour la pratique.

███████   The concept of evaluation utilization has been of interest to evaluation scholars and practitioners for quite some time. This particular domain of interest has undergone considerable evolution and development due in no small measure to the pioneering efforts of such noted theorists and researchers as Weiss (1981), Alkin, Daillak, and White (1979), and Patton et al. (1977). Perhaps among the most dramatic early developments in the conceptualization of the utilization construct was the recognition that its use need not be limited to directly observable, instrumental consequences of evaluation findings and recommendations; conceptual development or learning about the object being evaluated was identified as a legitimate if not pervasive consequence of evaluation activity. Indeed, such learning might be conceived to be an instance of "capacity building," a topic of considerable interest within evaluation circles as we begin the new millennium.

In response to a call from the president of the American Evaluation Association (AEA), Milstein and Cotton (2000) opened up the potential for dialogue among professional evaluators about concepts relevant to "evaluation capacity building" (ECB) by providing a thoughtful framework for consideration. They define evaluation capacity as "the ability to conduct an effective evaluation (i.e., one that meets accepted standards of the discipline)" (pp. 1–2). They then proceed to differentiate ECB from other kinds of capacity building such as the ability of individuals, organizations, or communities to achieve broad social or organizational goals. Presumably the kinds of instrumental and conceptual consequences of evaluation mentioned above would fall into this latter category. To the extent that evaluation helps individuals, organizations, or communities understand phenomena of interest or take appropriate actions in order to achieve valued goals, it would be considered a capacity building exercise, but one that is differentiated from ECB per se. In their words, "while undeniably important, this form of capacity is different from

the capacity needed to conduct effective evaluation activities in the first place" (Milstein & Cotton, p. 2).

A recent volume of *New Directions for Evaluation* by Compton, Baizerman, and Stockdill (2002) takes this perspective on ECB one step further. These authors frame ECB not only in terms of the ability to do quality evaluation but also to use it within the organizational context. Specifically, they define the term, albeit unparsimoniously, as:

> A context-dependent intentional action system of guided processes and practices for bringing about and sustaining a state of affairs in which quality program evaluation and its *appropriate uses* are ordinary and ongoing practices within and/or between one or more organizations/programs/sites. (Stockdill, Baizerman, & Compton, 2002, p. 8, emphasis added)

In this conceptual definition the consequences of evaluation, in terms of its appropriate uses, are thrown into the mix. We would argue that such an extension is laudable and that a sharp distinction between ECB and capacity building associated with the consequences (i.e., uses) of evaluation is unwarranted and perhaps counterproductive. Moreover, the integration of evaluation into the culture of organizations has, in our view, as much to do with the consequences of evaluation as it does the development of skills and knowledge of evaluation logic and methods.

From our point of view, there is much to be learned from an integration of what we know about evaluation utilization and ECB. We choose to frame evaluation as an organizational learning system. In other words, evaluative inquiry, as part and parcel of organizational culture and operations, has the potential to serve as a potent means for organizations to develop their organizational learning capacity (OLC). Through doing evaluation and developing the capacity to do it, organizations become more adroit in constructing shared representations of knowledge and structures, predisposed to generate new knowledge, inclined to capture and interpret external information, and apt to question basic assumptions about the organization, its goals, and strategies for achieving them. Such consequences, in the parlance of evaluation utilization, represent group or organization level "process use."

The purpose of this article is to explore, through an extensive review and integration of recent scholarly literature, the conceptual

interconnections and linkages among developments in the domains of evaluation utilization, evaluation capacity building, and organizational learning. We argue that the capacity to do effective evaluation is inextricably linked to the kinds of capacity building outcomes of evaluation that we might generally associate with evaluation utilization. Further, we propose that ECB is tightly connected to organizational development and capacity building to the extent that it contributes to an organization's propensity to learn. A product of our review and integration is a conceptual framework that links evaluation activities and consequences with organizational capacity building. We begin by introducing this framework and then proceed to systematically review and integrate recent empirical literature using it as an organizing mechanism. Finally, we develop a research agenda around the problem of integrating evaluative inquiry into the organizational culture.

CONCEPTUAL FRAMEWORK

Using some of our prior conceptual work (Cousins, 1996, 2003; Goh, 2000; Goh & Richards, 1997) as a platform, we developed the theoretical framework that appears in Figure 1 through reading, critiquing, and synthesizing recent scholarly literature linking evaluation to organizational development and learning. The framework is a conceptual representation of the key variables of interest (spheres) and relationships among them (arrows). In this section we describe the framework, which then serves as an advance organizer for the review of empirical literature to follow.

The spheres of interest are roughly divided into those associated with the organization itself and those associated with evaluation activities and consequences situated within the organizational context. Organizational consequences are shown to depend on an organization's capacity to learn, which in turn is sustained by various organizational support structures. One such support structure would be evaluative inquiry that leads to various consequences including the use of evaluation findings. We considered evaluative consequences to be organizational consequences in this context. Through the use of evaluative inquiry and experiences in using evaluation, the organization develops its evaluation capacity. Such capacity becomes integrated with OLC, manifest in the development of the cultural variable "organizational readiness for evaluation." We now consider the individual components of the framework more directly.

Figure 1
Conceptual Framework of Evaluative Inquiry as an Organizational Learning System
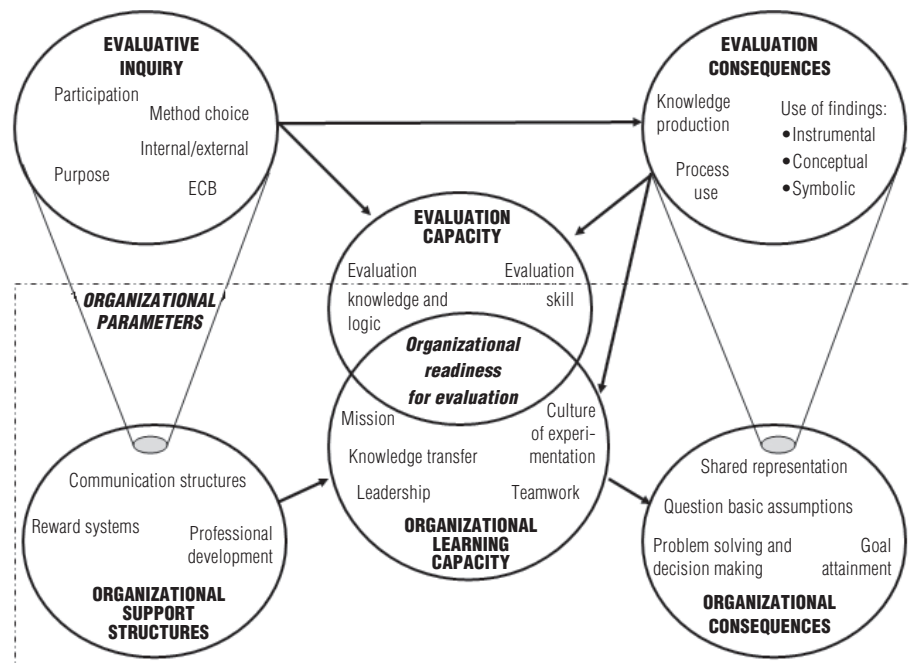
Organizational Dimensions

> Organizational learning is a psychosocial construct that has gar-
> nered considerable attention from organization theorists, educational
> administrators, and other corporate managers over the past decade
> (Cousins, 1996; Fiol & Lyles, 1985; Garvin, 1993; Goh & Richards,
> 1997; Huber, 1991; Leithwood, Aitken, & Jantzi, 2001). It is well
> known that organizations differ quite enormously in their ability to
> learn and adapt to rapidly changing environments (Fiol & Lyles,
> 1985; Senge, 1990), but what is not readily understood is how to
> foster the development of significant and sustainable OLC (Garvin,
> 1993).

Organizational Consequences

> Fundamental to conceptions of organizational learning is the devel-
> opment among organization members of shared mental representa-
> tions or understandings of the organization and how it operates.

Most theorists agree that organizational learning cannot happen in the absence of individual learning by organization members. This multi-dimensional construct ranges from low-level, first-order, or single-loop learning, where change is incremental, to high-level, second-order, or double-loop learning where fundamental assumptions about the organization and its operation are brought to the surface, questioned, and ultimately altered (Fiol & Lyles, 1985; Huber, 1991; Lant & Mezias, 1992; Lundberg, 1989).

Organizational Learning Capacity

The allure of understanding organizations as learning systems is powerful, but as Garvin (1993) suggests, it is imperative that we move on to clearer guidelines for practice and operational advice as to how to build a learning organization. To this end Goh (2000) synthesized and integrated the theoretical literature and argues that learning organizations have the following core strategic building blocks:

- *Mission and vision:* clarity and employee support of the mission, strategy, and espoused values of the organization;
- *Leadership:* leadership that is perceived as empowering employees, encouraging an experimenting culture, and showing strong commitment to the organization;
- *Experimentation:* a strong culture of experimentation that is rewarded and supported at all levels in the organization;
- *Transfer of knowledge:* the ability of an organization to transfer knowledge within and from outside the organization and to learn from failures;
- *Teamwork and cooperation:* an emphasis on teamwork and group problem-solving as the mode of operation and for developing innovative ideas.

Although they are presented as separate dimensions, Goh (2000) proposed that these building blocks are interdependent and mutually supportive conditions in a learning organization; in essence they define an organization's learning capacity. Such capacity determines the extent to which organizational consequences occur.

Organizational Support Structures

OLC building blocks are also said to depend on organizational structures and supports such as low job formalization and the acquisition of appropriate knowledge and skills by organization members

(Goh, 2000). Reward systems, in the form of formal and informal incentive mechanisms, also represent structural support for OLC, as do various communication structures within the organization, which serve to foster the horizontal and vertical flow of knowledge and information. Professional development activities, formal and informal, represent yet another organizational support structure.

Evaluation Dimensions

We define evaluation as systematic inquiry leading to judgements about program (or organization) merit, worth, and significance, and support for program (or organizational) decision making. Judgement implies making comparisons between systematically gathered data (observation) and some standard or basis for comparison (another program, performance at a previous point in time, an external benchmark). We include in our definition such activities as "needs assessment," systematic inquiry into identified needs to which programs (or organizations) are intended to respond, and "monitoring" systematic observation that does not necessarily imply judgement but nevertheless provides support for decision making.

Evaluative Inquiry

Evaluation purposes range from hard-nosed judgement-oriented summative exercises to formative, improvement-oriented ones intended solely to provide support for decision making. These evaluations can be conducted by members or groups external to the organization, by those internal to it, or by a blend of the two. It varies in the extent to which it involves non-evaluator stakeholders (i.e., those not trained as evaluators), ranging from exclusive use as sources of data and information to deep and full partnerships between evaluators and non-evaluator stakeholders in carrying out the evaluation (Cousins & Whitmore, 1998). Multiple methods can be employed in evaluation, ranging from quantitative to qualitative and often involving a mix of the two. In some cases, method choice implies deeper epistemological considerations, but it is important to distinguish between method and methodology (Kushner, 2002).

We consider ECB to be a special case of evaluative inquiry. We concur with the definition of ECB put forward by Stockdill et al. (2002) (see above), but are of the view that actions associated with developing infrastructure, processes, and procedures that work to develop evaluation capacity represent interventions that have a distinct

evaluative component (Preskill & Torres, 1999). And so, evaluative inquiry can lead to various forms of utilization or more directly and intentionally to building evaluation capacity.

Evaluative inquiry, within our framework, is conceived to be an organizational support structure — not unlike communication structures, professional development structures, and the like — with the potential to foster organizational learning. It is enhanced and distinguished in Figure 1 because it provides the central focus for our paper. We are interested to know how evaluative inquiry can be integrated into the organizational culture. It is important to note that evaluative inquiry is a structure that can in and of itself be supported by other organizational support structures, such as training and professional development, resources, and infrastructure.

Evaluation Consequences

Ultimately, evaluation leads to knowledge production, the validity, credibility, sophistication, timeliness, and relevance of which depend on the evaluation processes in place. It also leads to forms of the use of the knowledge produced. Conventional forms are captured under the heading "use of findings," including instrumental uses (support for discrete decisions), conceptual uses (enlightenment, learning), and symbolic uses (persuasion, compliance) (Shulha & Cousins, 1997). But "process use," a term coined by Patton, is an additional discernable consequence of stakeholder participation in, or knowledge of, evaluation activities (Cousins, 1996; Patton, 1997, 1998; Preskill & Caracelli, 1997; Shulha & Cousins, 1997). By virtue of their proximity to the evaluation, stakeholders may develop in ways that are quite independent of the findings or substantive knowledge emerging from the inquiry. For example, participation has the potential to lead to the development of research skills and the capacity for self-critique, self-determination, and systematic inquiry at the level of individual stakeholders. There may also be collective effects at the group, team, or organizational level. Process use of this sort is of central interest to us in this article. We see it as a more indirect mode of evaluative inquiry leading to the development of evaluation capacity.

As pointed out above, we consider evaluation consequences to be a special case of organizational consequences. Instrumental, conceptual, and symbolic uses of evaluative findings can be framed as elements of organizational decision making, problem solving, and learning.

Evaluation Capacity

Following Stockdill et al. (2002), evaluation capacity is represented by the organizational processes and practices that are in place to make a quality evaluation and its uses routine. To the extent to which evaluation capacity is present within the organization, members of the organization develop their knowledge of evaluation logic and methods. They also develop their skills in actually doing evaluation, including those associated with planning; instrument development; data collection, processing, and analysis; interpretation; and reporting and follow-up. In essence, through sustained evaluative inquiry and, in particular, continued and routine use of evaluation findings and processes, evaluation capacity becomes integrated within the organizational culture. It becomes what might be labeled "organizational readiness for evaluation," a term originally coined by Mayer (1975, cited by Seiden, 2000) in relation to Davis and Salasin's (1975/1983) A-VICTORY model of organizational development and change.

Many interventions have been proposed as ways to enhance OLC within organizations, including strategic planning, employee training initiatives, focused organizational development and school improvement projects, and the installation of management information systems. Evaluative inquiry is an intervention that holds much promise in this regard. The literature suggests that organizational readiness for evaluation is a cultural dimension that becomes part of OLC. In theory, if evaluation becomes integrated into the ongoing activities within an organization, it may become a learning system that fosters the development of shared values and understanding among organization members. Several evaluation theorists have recognized the potential for evaluation to foster such collective developmental effects (Cousins, 1996; Cousins & Earl, 1992; Forss, Cracknell, & Samset, 1994; Owen & Lambert, 1995; Patton, 1999; Preskill, 1994; Preskill & Torres, 1999; Torres, Preskill, & Piontek, 1996).

REVIEW AND INTEGRATION OF CURRENT EMPIRICAL LITERATURE

In essence this article represents a form of meta-analysis of empirical research in the area. Described below are the sampling criteria, characteristics of the sample, and the method used to analyze findings across studies. We then turn to an integration of the findings.

Sampling Criteria

Despite early interest in organizational readiness for evaluation, the construct appears to have moved off the radar screen of evaluation scholars until recently (Seiden, 2000). With burgeoning interest in evaluation's relationship to organizational learning, the term has once again come into favour among researchers. For this reason, we sampled available empirical research from within the past five years only. We implemented computer searches of standard data bases — Educational Resources Information Clearinghouse (ERIC), Psychlit, Sociofile — to identify recent published work. Keywords included "evaluative inquiry," "organizational learning," "organizational readiness for evaluation," and "evaluation utilization." We sampled only published research subject to peer review or dissertation/thesis committee review. From an initial sample, we tracked down other works through bibliographic follow-up. We also scanned recent issues of well-known evaluation journals (e.g., *American Journal of Evaluation*, *Canadian Journal of Program Evaluation*, *Evaluation*, and *Evaluation and Program Planning*).

The resulting sample is therefore multidisciplinary with empirical studies reported in many different evaluation practice and organizational contexts. To be included, studies had to be empirical, that is, based on systematic observation of evaluation or organizational practice. Organizational learning studies that were located had to include an evaluative inquiry element in order to be retained.

Sample Descriptive Characteristics

Ultimately we identified a sample of 36 studies, which are described in Table 1. Fifteen of the studies are what we might loosely call empirical (e.g., Boaz, 2002; King, 2002; VanderPlaat, Samson, & Raven, 2001). That is to say, they are based on reflective accounts of the author (usually an evaluator) with regard to one or more specific case examples. These narratives are based on observation and interpretation of lived experiences with evaluation, yet authors do not specify methods for capturing their observations nor other relevant sources of evidence supporting the case. While rich in detail and insight, it is not possible for the analyst to judge the credibility or trustworthiness of the account on the basis of conventional canons for social inquiry. These papers are narratives implicitly in accordance with constructivist principles of inquiry and attendant relativist epistemological assumptions.

Table 1
Descriptive Characteristics of Empirical Studies

Study	Sample	Context	Methods	Purpose / Research Question(s)	Main Findings
Balthasar & Rieder (2000)	Evaluators, program directors and federal office of energy decision makers from 4 out of 40 evaluations	Second phase of a larger three-study project that involved the evaluation of the Swiss Energy 2000 program	Meta-evaluation: 8 interviews plus document review in each of 4 case studies	-The forms of learning through evaluation -The conditions of learning through evaluation	Different kinds of learning in each case study: direct—where feedback info led to the recommendations of the evaluations being used, indirect—new appreciation of the situation is observed and arguments can influence the communication of political issues. It is necessary to distinguish the types of learning and effect of evaluations at the levels of operation, structure, and strategy.
Biott & Cook (2000)	Not applicable	Evaluation of the national Early Years Excellence Centres Pilot Programme (EYECPP) in the UK	Reflective case study: written from perspective of evaluator	To explore cross cutting evaluation issues	As the EYECPP project progressed there were increased tensions between performance management and participatory evaluation. Difficult to do locally given limited resources and pressure; key to a worthwhile role for local evaluators is in their contribution to learning.
Boaz (2002)	Not applicable	The evaluation of the Better Government for Older People (BGOP) in the UK	Reflective case study: Written from point of view of the evaluator	To examine how evidence-based policy making presents challenges for researchers	Researchers can be more effective in maximizing utilization for policy and practice if they go beyond the traditional roles of delivering reports and presenting findings to funders. The likelihood of use for policy dependent on the form in which its findings are made available.
Brandon & Higa (2004)	41 school administrators and teachers in 17 schools	School-based training initiative to enhance evaluation knowledge and skills in the USA	Longitudinal field study: questionnaires and retrospective interviews.	Change in attitudes toward evaluation; change in confidence to do evaluation; perceptions of evaluation capabilities	No improvement in attitudes. Self-confidence as evaluators improved. Assessment of capabilities as evaluators favourable.

Table 1 (continued)

Study	Sample	Context	Methods	Purpose / Research Question(s)	Main Findings
Chacon-Moscoso et al. (2002)	The Homeless Child Care Services Organization in Seville, Spain	Formative evaluation with homeless child care centers in Seville, Spain	Reflective case study: written from point of view of evaluators	To describe example of mutual catalytic model of formative evaluation	Model is implemented successfully; evaluation would not have been as sensitive without the active cooperation of local practitioners and representatives at the centre; evaluator key to process.
Compton et al. (2001)	Not applicable	Implementation of Collaborative Evaluation Fellows Project (CEFP), involving American Cancer Society, graduate students, and faculty in the USA	Reflective case study: written from the perspectives of those involved in the design and implementation of the CEFP	To describe CEFP and identify lessons learned after first 3 years of implementation	Lessons setting up CEFP: administrative and tech support; annual training, opportunities for consultation, outside funding, evaluation expertise, clarity re: evaluation. Implementation lessons: common framework; training, program evaluation expertise, student entry skills, tensions between report and thesis requirements, time intensive, incentives.
Compton et al. (2002)	Not applicable	ECB efforts within the American Cancer Society since 1995; nationwide, community-based, volunteer organization in the USA	Reflective case study: written from the perspectives of 4 ECB practitioners involved from the beginning in 1995	To tell the story of ECB at the ACS through looking at context, history of ECB, practice principles and their operationalization, assessment of ECB, and lessons learned	ECB triggered by changes in the larger context, especially reform in program development. Various strategies to enhance ECB, but most effective is evaluation fellowship collaborative. Several principles guide ECB, among them is evaluate evaluation unit; mostly informal to date but some formal aspects; training staff in use of evaluation processes and products.
Connor & Tanjasiri (1999)	2 project directors in Colorado Healthy Communities Initiative (CHCI)	Two out of 29 communities that are implementing the Colorado Healthy Communities Initiative (CHCI) in the USA	Case studies: participant observations of evaluators; discussions with project directors	To examine cross-cutting evaluation issues	Citizen involvement fostered ownership of the project and suggests that actual use will be more likely. Importance of a broad-based advisory council that kept the project in touch with its community. New partnerships were formed. Evaluation capacity in the community developed.

Gibbs et al. (2002)	8 community-based organizations, 9 health departments, 28 technical assistance providers	8 sites in which Centers for Disease Control funded HIV prevention programs; usually 5 or fewer full-time employees in the USA	Qualitative field study: content analysis of audiotaped interview transcripts	To describe beliefs and attitudes toward evaluation; to identify factors influencing evaluation capacity	Major influence of funding agency expectations, availability of evaluation resources, leadership, and staff, and availability of evaluation tools and technology. Stage model of EC: compliance, investment, advancement. Must tailor ECB to stage.
Gilliam et al. (2003)	58 funded health department personnel	Description and analysis of evaluation training and technical assistance initiative in the context of HIV prevention programs in the USA	Case study: content analysis of technical assistance requests that came in during one-year period	To assess the efficacy of evaluation capacity-building initiative	Continual involvement of grantees and program and evaluation experts important to making guidance and technical assistance systems more user-friendly and cost-effective.
Jacobs (2000)	Not applicable	Evaluation of Learning in English for Academic Purposes (LEAP) in South Africa	Reflective case study: written from perspective of evaluator	To describe the application of the 10-stage model of educational innovation within the context of transformation	The model provided a conceptual framework for the entire evaluation process. Ensured close scrutiny of the contextual factors; brought the needs of the evaluation into focus. Balanced evaluation produced that could speak to the needs of disparate audiences.
Katz et al. (2002)	6 principals from alumni Manitoba School Improvement Program (MSIP) schools	Middle and secondary schools involved in MSIP in Canada	Exploratory case study: semi-structured interviews	To provide evidence of how the processes and practices surrounding the use of data for wise decision-making become intrinsically derived	Evaluation requirement was initially perceived as burdensome. Many alumni schools sustained the program's evaluation legacy evidenced through data-driven goal-related inquiry, systematic reflection and planning, and the presence of an evaluation-minded culture.
King (2002)	Not applicable	Large urban school district in Minnesota that is committed to enhancing evaluation function and capacity. Author hired as internal	Reflective case study: retrospective observations of author, a university professor hired for 2 years as an internal evaluator	To describe experience and explore issues arising in district's effort to enhance evaluation capacity; how the work was originally conceived and how and	Progress reviewed on 4 goals for 2 years of ECB work: (1) develop staff commitment and skills in program evaluation and its use—partial; (2) build an infrastructure for data collection, analysis, and presentation—partial; (3) facilitate existing school improvement process—some; (4) create network

Table 1 (continued)

Study	Sample	Context	Methods	Purpose / Research Question(s)	Main Findings
		evaluator for 2-year period in the USA		why it changed	of people to routinize inquiry—limited. Lessons learned: each project an instructional opportunity; danger of creating more demand than can handle; participatory approach effective.
Kruse (2003)	Administrators, teachers, and students in the US city school district	Understanding remembering as an organizational learning construct in schools in the USA	Case study: narrative analysis by author using interviews, observations, and documents	The development and application of organizational memory as an analytic construct for understanding case data; the utility of memory as a construct helpful in understanding organizational outcomes and performances	Knowledge management and organizational memory are recursive social systems. Soft knowledge along with hard test score data allowed organizational learning to occur. Importance of wisdom in learning.
Lee (1999)	Evaluator reflections on evaluation role in multi-year, multi-site Canadian school improvement initiative	Evaluation support for funded school improvement program in Manitoba, Canada	Case study: participant observation; excerpts from prior data collection/reports	To describe evolution of the role of evaluation in capacity building for the Manitoba School Improvement Program	Evaluation evolved from participatory approach with dependency on evaluator, to empowerment self-directed evaluation, to capacity-building focus driven by teacher empowerment. Internal and external accountability are possible.
Mackay (2002)	Not applicable	World Bank experiences in developing ECB in diverse and difficult governmental contexts over a 20-year period	Reflective case study: developed from the perspective of a bank ECB coordinator	To tell the stories of ECB lessons learned in the Bbnk's experience with governments in diverse and difficult environments, and how ECB is becoming recognized and valued within the bank itself	Demand is the main prerequisite; synergies with other capacity-building work is essential; tailor ECB according to country circumstances; conduct baseline diagnosis. Stages of development described: bank is increasingly committed to ECB as a worthwhile activity and it is being accorded a reasonably high priority.

Matlay (2000)	6000 organizations from yellow pages; 600 small business managers; 60 matched case studies	A report of a study of organizational learning in the small business section of the UK economy	Field study: multi-method staged study involving telephone survey, interviews, case studies	To find out the type and incidences of OL and the locus of OL in their firms; to examine more closely the OL in some firms; to compare and contrast companies that are similar in size in terms of OL	Mostly informal learning taking place. OL occurs more with owner/managers. Work-based learning was incidental and any knowledge gained was only used in the short term. Firms that employed intentional learning strategies were slower and more deliberate, but were more effective at influencing the direction and delivery of a firm's programs.
Milstein et al. (2002)	Not applicable	Mammoth ECB effort in large national USA public health prevention and control organization CDPC over multi-year period. Many partner organizations	Reflective case study: report is integration of internal and external viewpoints of respective authors/participants	Why strengthening evaluation capacity became an explicit priority; principles and procedures; effects of capacity building; lessons learned	Impetus came from senior administration. Six principles guided the multi-year, multi-faceted ECB initiative. Successful ECB but not investigated systematically; no effects on organization specified. Admin role in creating trusting environment pivotal; must deal with defences of resources, righteousness, and resistance.
Morabito (2002)	Not applicable	A USA school that provides instruction to physically and medically disabled children from K to 12	Reflective case study: written from perspective of evaluator	What specific evaluator roles, philosophies, and interpersonal dynamics hold the potential to foster process influence?	The workgroup met its objective and much more. The school is now in the process of setting additional curriculum goals, and data are now being collected on the indicators that were formed into case conference team checklists.
O'Sullivan & D'Agostino (2002)	Program practitioners participating in evaluation networking activities	Evaluation Voices (EV) approach was implemented in the context of early childhood education program evaluation in the USA	Longitudinal case study: reports from evaluation fair presentations given by evaluation contractors in 2 consecutive years	Can collaborative evaluation and networking improve the quality of self-evaluation?	Collaborative evaluation can result in improved quality of evaluations. Structural requirements were important factors.
Porteous et al. (1999)	Not applicable	An evaluation capacity-building initiative in the Ontario, Canada public health system	Reflective case study: written from the point of view of evaluators and ECB providers	To describe a Toolkit initiative designed to help public health managers in Ontario improve their knowledge and skills in logical model development and program evaluation	Toolkit appears to be a successful method for teaching managers to develop logic models and implement evaluation techniques; some suggestions for improvement.

Table 1 (continued)

Study	Sample	Context	Methods	Purpose / Research Question(s)	Main Findings
Preskill et al. (2003)	16 interviews with advisory group members and senior administrators of American Cancer Society (ACS)	Yearly evaluations of the Tell a Friend program of the ACS in the USA	Case study of two evaluations; interviews	What did advisory group members learn from involvement in evaluation? How did they learn? What factors supported or hindered learning?	Five main categories of influence on process use: facilitation of evaluation processes; management support; advisory group member characteristics; communication attributes; organization characteristics. Examples of process use identified: learning about evaluation; about the program; overall evaluation experience.
Robinson & Cousins (2004)	9 Canadian key informant national program practitioners	Canadian national training program reform initiative	Longitudinal case study; participant observation; interviews, focus group	To investigate the organizational consequences of internal participatory evaluation in a Canadian national training organization	Significant organizational consequences observed, especially capacity building. Conditions for participatory evaluation were favourable.
Russ-Eft et al. (2002)	23 sales people, 10 implementation specialists, and 9 consultants; 40 representatives of client companies	Evaluation of a sales program within a business services organization (XYZ Corporation) in the USA	Descriptive case study: satisfaction interviews, surveys	Use and non-use of the evaluation results, with particular emphasis on process use; factors contributing to use and non-use of this private sector evaluation	Program terminated on the basis of non-evaluative information; some development of shared understanding, engagement, self-determination. Environmental factors affecting use: rumours of a merger with another company had profound effect.
Scribner et al. (1999)	3 rural middle schools: school admin and leadership team members and other teachers, $N = 35$	USA district-wide school improvement initiative	Collective case study of three schools: constructivist orientation, school-based observation, conferences and	How do school improvement programs foster the development of professional communities? What organizational factors support and/or impede the	Wider school communities must be involved in discussions of SIP from commitment through implementation. Factors that influence the development of professional communities: leadership style of principal, past events, politics of allocating resources, organization of schools, and

Source	Sample	Description	Method	Purpose	Findings
			projects, document and artifact collection	development of professional communities?	teacher work.
Silins et al. (2002)	Teachers and students from 96 secondary schools from South Australia and Tasmania, $N =$ 2503 (teachers) and $N =$ 3500 (students)	A path model is used to test the relationships between school-level factors and school outcome measures in terms of students' participation in and engagement with school in Australia	Field study, 2 phases: (a) survey data was collected using two questionnaires; (b) path analysis examined the nature and strength of interrelationships between variables	To examine the nature of organizational learning and the leadership practices and processes that foster organizational learning in Australian high schools	The conditions for organizational learning are the same as the conditions that are associated with the 3 school leadership variables: principal transformational leadership, actively involved administrative teams, and distributed leadership. Organizational learning is also a significant mediator of principal and leadership team effects on teachers' work and on student outcomes.
Smith & Freeman (2002)	Not applicable	Implementation of continuous system-level assessment (CSLA) in context of school-level staff capacity building in the USA	Reflective case study: written from perspective of model developers/ implementers	To introduce and describe the application of CSLA implementation through evaluation of school-wide discipline plan	CSLA builds program and evaluation expertise by creating a framework for professional development that can be used in a flexible manner; helps identify resources; complements external evaluation by building local capacity; uses data to encourage staff to contemplate how their values hinder or contribute to a program's success
Stevenson (2001)	One secondary school in Australia; 1/3 of teaching staff	School with established structures for shared decision-making and collective leadership; characteristics of learning organization in Australia	Case study: observed meetings, formal interviews, informal conversations, collection and analysis of relevant documents	What kinds of issues are addressed and decisions made in a school that has strived to democratize decision-making and become a learning organization?	Curriculum forum: 8 issues addressed but most were administrative; little learning took place. Social justice & equity forum: inquiry and action were integral, feedback solicited from students. Teachers' views on diversity and discrimination were willingly questioned but not views about curriculum.
Stevenson et al. (2002)	14 community-based agencies responsible for evaluating their	State-funded evaluation support unit commissioned to	Longitudinal case study: evaluation proposals, confidence ratings re: multiple	To document successful increase in agency capacity as a consequence of ECB initiative; to identify	Needs assessment useful for focusing tech support, informing reform of state proposal guidelines. Participant confidence varied across four skill categories; more progress on prelimi-

Table 1 (continued)

Study	Sample	Context	Methods	Purpose / Research Question(s)	Main Findings
	own programs	increase evaluation capacity building: needs assessment, modelling, tech assistance, training in the USA	tasks, checklist of evaluation steps completed	main areas of resistance	nary steps and design than on analysis and utilization. Technical assistance essential.
Sutherland (2004)	46 US school administrators, teachers, staff, parents, and students plus 12 district administrators in Edison Project school	Edison Project school, major USA reform initiative funded by foundation designed to create innovative schools that operate within current public expenditures	Longitudinal case study: interview and observation	How do the processes and practices surrounding the use of data for decision-making contribute to a culture of continuous improvement?	External reform viewed as extrinsic motivator, teachers become intrinsically motivated. Major role for leadership, structure for capacity building, continuous flow of information and communication. Culture of continuous data use over time. Elements of reform environment important.
Tang et al. (2002)	Not applicable	Large state-wide Tobacco Control Program, sponsored by the California State Dept. of Health in USA	Reflective case study: written from point of view of evaluators	To illustrate the utility and feasibility of empowerment evaluation in a comprehensive state-funded program	An empowerment evaluation with its emphasis on self-assessment, facilitation, and capacity building can meet the challenges of diversity yet guarantee that the evaluation will strengthen the individual local programs.
Terziovski et al. (2000)	5 case corporations	Australian corporations implementing TQM in context of Malcolm Baldridge National Quality Award (MBNQA)	Multiple case study: semi-structured interviews	What are the mutual dependencies between Senge's learning organization change strategies and the MBQNA framework?	Mutual dependence between shared vision and leadership; between HR development, team learning, and personal mastery; between mutual models/information and analysis/process improvement; between systems thinking/customer focus/strategic planning; organizational performance.

Author (Year)	Sample	Context	Method	Research Question	Findings
Vakola, M. (2000)	3 case corporations	Organizations in construction industry in 3 European countries that underwent a business process re-engineering exercise (BPR)	Descriptive case studies: the data from semi-structured interviews follow a pattern of analysis	Is there a link between evaluation activities that are part of a BPR exercise and organizational learning and innovation?	Evaluation of BPR had the following benefits: facilitated the change process and led to better OL, led to process innovation such as the introduction of improved technological capabilities in the companies.
Valovirta (2002)	27 agency and supervising department personnel	Government agencies in Finland in diverse policy sectors	Meta-evaluation: semi-structured interviews	To compare how agency evaluations conducted, objectives, and role evaluation played in overall framework of management development	Use takes place in social context; 4 types of arguments about evaluative results; many examples of post hoc legitimizing actions; evaluation incites argument and can result in reasoned debate; written reports offer common platform for dialogue and argumentation.
VanderPlaat et al. (2001)	Not applicable	Experiences with the Atlantic Community Action Program for Children (CAPC) regional evaluation in Canada	Reflective case study: written from perspective of evaluators	How the concepts of participation, empowerment, and social change are constructed	Employing a more inclusive notion of participants changes the understanding of the term "participatory research." Program participants, staff, government funders, community partners, managers, and evaluators all have a capacity to empower and be empowered. The success of a social intervention is better judged by what happens in the overall environment than by what happens to individuals and organizations in the collective.

The remaining studies were relatively evenly distributed over four design types. Five were qualitative case studies (e.g., Balthasar & Reider, 2000; Katz, Sutherland & Earl, 2002), and an additional five were comparative case studies (e.g., Connor & Tanjasiri, 1999; Scribner, Cockrell, Cockrell, & Valentine, 1999) Six studies were based on longitudinal data collection (e.g., O'Sullivan & D'Agostino, 2002; Robinson & Cousins, 2004; Sutherland, 2003) or multi-method field studies (e.g., Matlay, 2000). Many of the studies are North American, taking place in the U.S.A. (17) and Canada (6), but several are from other jurisdictions including the U.K. (Biott & Cook, 2000; Boaz, 2002; Matlay, 2000), Australia (Silins, Mulford, & Zarins, 2002; Stevenson, 2001; Terziovski, Howell, Sohal, & Morrison, 2000), Spain (Chacon-Moscoso, Anguera-Argilaga, Perez-Gil, & Holgado-Tello, 2002), Finland (Valovirta, 2002), Switzerland (Balthasar & Rieder, 2000), and South Africa (Jacobs, 2000). Two studies (Mackay, 2002; Vakola, 2000) involved data collection from multiple countries.

Method

Having identified and obtained the articles in our sample, we divided them among the authors for analysis and summarization. Each analyst produced a one-page summary for each of the allotted articles. The summary was structured by the column headings appearing in Table 1. In summarizing the articles, analysts looked for evidence and verbatim quotations that illustrated relationships among the variables of interest. Verbatim quotations were captured in bold with appropriate page numbers. All summaries were compiled into a single electronic file ordered alphabetically by author. The file was then converted to Folio Views (4.3) and coded according to the main variables of interest by one of the analysts. In working through the files, the analyst developed the conceptual framework showing relationships among the variables (see Figure 1). This figure and descriptive text were then circulated among the other authors for input and verification. The analyst made some modifications to the framework and proceeded to integrate the findings as presented below. Available to the analyst were copies of the original studies as well as the coded summaries.

All in all, we judge the knowledge base linking evaluative inquiry to organizational learning and development to be at a relatively early stage of maturity, as evidenced by a heavy reliance on qualitative methods and limited variety in modes of inquiry. Given this obser-

vation, our confidence in the veracity of findings, trends, and generalizations is necessarily limited. With these caveats in mind, we now turn to a review and integration of findings across studies.

Integration of Findings

The review was guided by the framework appearing in Figure 1, but the studies could be grouped into three main categories: studies of (a) evaluation capacity building (ECB) initiatives (direct and indirect), (b) incidental capacity building through evaluation implementation and consequences, and (c) studies of organizational variables in our framework with some link to evaluative inquiry. We now present the findings associated with each of these categories.

Evaluation Capacity Building Studies

Direct ECB studies: Ten studies were reports on the implementation and consequences of direct ECB initiatives that may or may not have commented on actual evaluation implementation. Four of these were published in a single volume of *New Directions for Program Evaluation*. Studies in this category looked at ECB initiatives in the U.S.A. domestic contexts of nation-wide organizations/agencies (Compton, Baizerman, Preskill, Rieker, & Miner, 2001; Compton, Glover-Kudon, Smith, & Avery, 2002; Gilliam et al., 2003; Milstein, Chapel, Wetterhall, & Cotton, 2002), large school districts (Brandon & Higa, 2004; King, 2002), a statewide initiative to support community-based organizations (Stevenson, Florin, Mills, & Andrade, 2002), and internationally through evaluation capacity development efforts of the World Bank (Mackay, 2002). These studies described ECB initiatives and processes and discussed tensions, successes, and frustrations. All met with some degree of success in developing evaluation capacity but within limits. Only two came to this conclusion on the basis of systematically collected evidence (Brandon & Higa, 2004; Stevenson et al., 2002). According to Milstein et al. (2002, p. 41), "One problem is that the evaluation profession as a whole still lacks a well-developed theory and associated indicators for understanding evaluation capacity at an organizational level, particularly its inherent change over time and 'ongoingness'."

Compton et al. (2002) found that developing capacity for use met with less attention than capacity to do evaluation: "[it] means work-

ing with individuals to use evaluation processes, such as creating logic models to make their program theory explicit, developing measurable objectives, and translating evaluation reports into action" (p. 58). This finding was corroborated by Stevenson et al. (2002), who showed that participant confidence was stronger for preliminary steps and evaluation design tasks than for analysis and utilization. King (2002) experienced some success in developing capacity within the school district. She was hired for a two-year term to assist in this way and described several tensions and challenges along the way to developing commitment and building an infrastructure that would support evaluation. King reported that an effective strategy was for her to participate directly on multiple district-level committees to help develop networks and foster communication. Ultimately, in her words,

> the members of certain committees accepted responsibility for working on evaluations of their activities, and over three hundred district staff, students, and parents had helped to manage or participate actively in one or more evaluations, learning about the process along the route. (p. 71)

Working with the American Cancer Society (ACS), Milstein et al. (2002) found that expectations for ECB arising from experience in doing evaluation were relatively benign. An effective strategy in this volunteer sector context was the development of a collaborative evaluation fellowship program (CEFP) with universities (see also Compton et al., 2001). A major impetus for ECB within the ACS context came from senior administration who lamented how evaluation activities were highly inconsistent across programs and how decisions about what, how, and when to evaluate were very haphazard and questionable. Administrative support was also identified as being key by Compton et al. (2002) and King (2002). Mackay (2002) reported on a 20-year history at the World Bank of ECB with diverse and difficult governments within the context of international development. Mackay reported not only on the World Bank's successes in developing monitoring and evaluation capacity, but also on those responsible for ECB legitimizing it among senior bank officials. According to him, "it has often been easier to persuade developing country governments about the need for ECB than it has been to persuade Bank staff" (p. 93). Regarding ECB in developing countries, Mackay highlighted the need to synchronize ECB activities with other ongoing forms of capacity building and to be culture-sen-

sitive. He observed that demand is a main prerequisite for ECB. King (2002) agrees, but cautions about the danger of creating demand that might exceed evaluation capabilities.

Other factors key to sustaining ECB initiatives were reported by Compton et al. (2001), who identified lessons learned for setting up the CEFP mentioned above as well as for implementation. Influencing factors include the existence and availability of evaluation expertise and training support and common understandings about evaluation among stakeholders. Because CEFP involved students basing thesis work on evaluation activities, various tensions between the practical and academic milieus surfaced. Gibbs, Napp, Jolly, Westover, and Uhl's (2002) findings corroborated many of the factors identified by Compton et al. Gibbs et al. observed that ECB sustainability for agencies working in HIV/AIDS prevention is a function of external pressure from funders, leadership, and availability of resources for evaluation. They proposed a stage model of evaluation readiness to which ECB initiatives should be tailored.

Four other studies reported on ECB initiatives. Gilliam et al. (2003) described and evaluated a framework developed by the Center for Disease Control in the U.S. for the purpose of developing evaluation capacity among HIV prevention program personnel. Requests for technical assistance were content-analyzed over a one-year period. The authors developed a set of lessons for organizations interested in developing evaluation capacity; the continual involvement of stakeholders including program and evaluation experts was portrayed as being key to adapting and changing guidance and technical assistance systems to be more user-friendly and cost-effective. Porteous, Sheldrick, and Stewart (1999) described an initiative designed to help public health managers in Ontario improve their knowledge and skills in program evaluation. This initiative involved the development of the Program Evaluation Toolkit (an aid to conducting local-level evaluations) and a workshop on its use. The Toolkit was viewed as being credible in large part due to the diverse nature of the group producing it, including those with practical public health experience. According to the authors, feedback from the 500 people who attended the various workshops was very positive. The success was likely due to: different formats to appeal to different learning styles, memory aids, examples, humour, and activities to encourage participation. Smith and Freeman (2002) examined an evidence-based intervention called Continuous Systems-Level Assessment (CSLA), a continuous, comprehensive, integrated process

within the school system. The initiative required staff to integrate evidence into program decision making and organizational improvement. Smith and Freeman concluded that the initiative was successful but that further study was needed:

> [CSLA] contributes to the ways school and other program staff and consumers can utilize data in a continuous, collaborative fashion to make data-based decisions about the efficacy of programs, value of specific interventions and the staff capacities needed to sustain those programs and interventions deemed valuable. Although CSLA is a promising approach for assisting schools as they implement new programs, further evaluation research is needed to assess its effectiveness over time and across settings. (2002, p. 317)

Finally, Brandon and Higa (2004) reported a study of change in school personnels' attitude and perceptions about evaluation as a result of training and consultation. Training was provided to administrators and teachers in 17 Hawai'ian schools; all were expected to carry out summative evaluations of school projects. Brandon and Higa showed that while attitudes toward evaluation did not change appreciably, participants' confidence in themselves as evaluators and their self-assessment of evaluation capabilities both improved.

Indirect ECB studies: A second category of studies of ECB-related initiatives were reports about external organizational interventions that had significant evaluative requirements built into them, thereby *indirectly* contributing to capacity building through process use. Two of the studies (Katz et al., 2002; Lee, 1999) were associated with the Manitoba School Improvement Program (MSIP), a secondary and middle school initiative that was initially funded by a private foundation. The MSIP initiative required that schools evaluate their own school improvement initiatives and in that sense provided an external accountability demand. Lee (1999) reported her reflections from the point of view of an evaluation consultant who assisted schools with evaluation and provided training where possible and feasible. Although Lee identified several tensions, she provided some positive observations about cultural change within the schools:

> It has been possible to move teachers from a dependence on the evaluation consultant … to a relationship

> where the evaluation consultant is a resource and support for the school's ongoing inquiry and reflection. (p. 174)

> The concept is that if people in schools can see the relevance of data collection in relation to their own goals and outcomes, they will begin to value and use the evaluation process … [and] have an array of information to support demands for external accountability. (pp. 175–176)

Similar observations were provided by Katz et al. (2002). In their interview study with school principals from MSIP alumni schools (i.e., project and funding completed), they found that the external evaluation requirement was initially perceived as burdensome, but schools sustained the MSIP's evaluation legacy even after funding had run its course. This was evidenced through data-driven goal-related inquiry, systematic reflection and planning, and the presence of an "evaluation-minded culture." A third study, by Sutherland (2004), also related to a foundation-funded school-improvement initiative with externally mandated evaluation requirements. In an Edison Project school — a Californian initiative designed to create innovative schools that operate at current public school spending levels — Sutherland collected interview data over time from a variety of members of the school community. Her data revealed a trend toward moving from the external mandate for evaluation as an extrinsic motivator to one that became intrinsic within the school. According to Sutherland, "a culture of continuous data use for continual improvement can occur when individuals stop thinking about data as something that is done *to* the school and start thinking about data as something that is done *by* the school and *for* the school" (p. 289, emphasis in original).

Finally, O'Sullivan and D'Agostino (2002) invoked an approach called Evaluation Voices, in which external evaluators worked with the contractors of an early childhood education initiative to provide them with the means to conduct evaluations. The external evaluators helped the contractors set up data collection systems, design instruments, analyze data, and the like. These evaluation activities built upon each other over a three-year period. The main conclusion from this study was that collaborative evaluation can result in improved quality of evaluations; clearly stated reporting requirements, cluster networking, and an annual presentation requirement may have encouraged this effect.

Summary: Many ECB studies examine the contribution to the development of evaluation capacity directly through interventions involving personnel, training, resources, and conceptual models. Other studies observed external initiatives that carry with them accountability requirements; they demonstrate through process use the development of evaluation skills, knowledge, and "habits of mind." None of these studies commented on or showed direct links to the integration of evaluation into the organizational learning capacity of the respective organizations, nor were organizational consequences of evaluation capacity discussed. However, it would be fair to say that links to the development of a culture of evaluative inquiry or organizational readiness for evaluation were evident in the studies. We now turn to studies of evaluation use that do not have direct or explicit connections to ECB.

Evaluation Use Studies

Fourteen studies focused on actual evaluations and observed consequences for organizations and systems. Two such studies were meta-evaluations that systematically examined evaluation implementation and effects from samples of evaluations. Valovirta (2002) used semi-structured interviews to capture perspectives from 27 agency and supervising personnel in the context of Finnish government sector policy evaluation. He characterized evaluation use in terms of argumentation within social context. Evaluation was observed to incite argument and can result in reasoned debate; written reports, for example, offer a common platform for dialogue and argumentation. Here we see evaluation contributing to systemic capacity building and fostering systemic (policy) outcomes of decision making, problem solving, and shared representation of agency performance and effects. Balthasar and Rieder (2000) also engaged in a meta-evaluation of sorts, although they sampled only 4 out of 40 evaluations of the Swiss Energy 2000 program. The study provides important linkages to the learning function of evaluation in capacity building. Balthasar and Rieder found that it is necessary to distinguish the types of learning and effects of evaluations at the levels of operation, structure, and strategy. At the operational level, the fine-tuning of the program occurs. At the structural level, it includes changes in the organizational structure that supports an action or measure. At the strategic level, the focus is on the overall conception and aims of a program. Different types of learning, direct and indirect, were observed in each of the four selected sites.

The vast majority of use-oriented studies in our sample link collaborative, participatory, and empowerment forms of evaluation use to evaluation capacity building, and organizational learning and development. In total, 10 studies had something to offer in this regard. Robinson and Cousins (2004) report a longitudinal study of a participatory evaluation of a national Canadian training program. Their data show that the evaluation, which involved members of partnering organizations as evaluation team members, led to a significant rethinking of the training program and the organizational structure through which it was offered. Preskill, Zuckerman, and Matthews (2003) conducted an exploratory case study of process use looking at advisory committee involvement in the American Cancer Society evaluation of the Tell a Friend program for the early detection of breast cancer. They found that advisory group learning was a function of the evaluation process and that key factors were facilitation, management support, communication, advisory committee member characteristics, and characteristics associated with the organization. Biott and Cook (2000) examined participatory evaluation in the national early years educational context in the U.K. In their study, they observed tensions arising between national level performance management demands and local level participatory evaluation. However, the authors conclude with arguments favouring enhanced articulation between local and national levels:

> At a national level, local evaluators should be able to influence decisions about what value should be placed on what findings. They will turn national indicators into well-grounded questions, and gradually help to transform frames of reference at a local level. (Biott & Cook, 2000, p. 412)

The theme of stepping out of traditional research-oriented roles so that evaluation may have an impact on policy deliberation and making was highlighted by Boaz (2002) in the context of a case examination of an evaluation of the Better Government for Older People policy initiative in the U.K. In this evaluation, local project staff valued the opportunity to talk with someone from the national program and Boaz observed that the involvement of intended program beneficiaries — older people — was a key factor in bringing about changes in attitudes and services for their generation. The benefits of expanding the role of the researcher were evident in the provision of support to the pilot sites. However, the activities drew upon skills (particularly in management and facilitation) that research-

ers did not necessarily possess. Boaz elevates the values of "the ability to put oneself, as a researcher, in the client's or potential client's shoes in order to understand the needs of the policy maker practitioner and realize that no one is obliged to use the research. As in partnerships, everything is up for negotiation!" (p. 451).

Looking at a regional evaluation context in the Canadian Maritimes, VanderPlaat et al. (2001) challenged some assumptions about empowerment evaluation of a multi-site community action program for children. Who are the participants and what does participation mean? Who empowers? Who is empowered? They made the case that employing a more inclusive notion of participants changes the understanding of the term participatory research. Traditionally, they argued, it is the project participants who are considered to be in need of empowerment; program designers, managers and evaluators do not see this need. However, as program participants learn to give voice to their realities, program managers must learn to listen. Therefore, they admonished, program participants, staff, government funders, community partners, managers, and evaluators all have a capacity to empower and be empowered. Tang et al. (2002) also looked at an empowerment evaluation model of a tobacco control program at the state-wide evaluation level in California. The authors reported details of a three-phase evolution of evaluation activity for this major multi-year initiative and made the claim that "empowerment evaluation with its emphasis on self-assessment, facilitation, and capacity building can meet the challenges of this diversity yet guarantee that the evaluation will strengthen the individual local programs" (p. 54). Yet another multi-site evaluation context that came under study was the Colorado Healthy Communities Initiative. Connor and Tanjasiri (1999) focused on 2 of 29 communities involved with the initiative and learned a great deal in their case study about the importance of citizen involvement fostering a sense of ownership and of having a broad-based advisory council in place. In their words, "Across both projects, it is clear that citizen involvement was key in developing and using indicator data" (p. 133). They observed and commented on the organizational impact of such involvement as well. "Community-based indicators not only provide a means to empower individual citizens but they also can empower communities as a whole, as they provide a roadmap and benchmarks along the way" (p. 134). The link between process use and capacity building is illustrated by these data.

Two of the participatory studies took place at the local level. Chacon-Moscoso et al. (2002) elevate the role of local practitioners in gener-

ating consensus on program objectives, evaluation goals, and strategies for evaluation results in a report on the implementation of what they call the "mutual catalytic model" of formative evaluation. The context was the evaluation of centres for homeless children in Spain. The implementation of the model began with the organization developing a shared version of the program theory in order to clarify program goals, evaluation objectives, and program re-planning. This shared version of the program theory was then discussed at each site-specific group and the theory was modified to suit each site, leading to revisions in evaluation and monitoring instrumentation. The organizational benefits of participatory evaluation were reported by Morabito (2002) in a study of evaluation in a U.S. school for special needs students. In this context, the evaluation was found to provide a forum for organizational deliberation and thereby contributed to the examination of basic assumptions and rethinking structures in place. Morabito described, as an example, how an extraneous issue arose incidentally in discussions about data collection activities:

> These discussions related to the future possibility of re-structuring the school to be consistent with [an external learning standard structure] … Problem solving discussions of this nature are artifacts of the learning culture, which the evaluator encouraged even at the expense of immediate progress on the task at hand. (p. 323)

Here we see the evaluator opportunistically adopting an organizational development facilitator role even though that was not his original intention. The evaluation provided a forum for reform deliberation that might not have otherwise taken place. As mentioned above, evaluator role expansion issues were also discussed by Boaz (2002).

Other studies looked at the application of systemic models in the context of program or operations management. Linkages to organizational and institutional capacity building were evident. Jacobs (2000) provided an example in the South African higher education setting. The focus of the research was the application of a model that simultaneously promotes institutional self-evaluation and external evaluation to an academic literacy course. Jacobs characterized the links to organizational capacity building in the following way:

> While the evaluation model proposed … places an oner-
> ous and immense responsibility on the evaluator(s) … it
> also places the institutional context and its governing
> policies under the scrutiny of evaluation. In this way an
> innovation is not evaluated and understood in isolation,
> but rather as an integral part of the context which so
> powerfully shapes its very chances of success or failure.
> (p. 279)

The importance of situating the evaluation within context is identi-
fied as a primary force influencing the impact of the evaluation, a
theme echoed by Vakola (2000) in a completely different organiza-
tional context. Vakola's study focused on business process re-engi-
neering and organizational learning in the corporate Australian
sector. Within the process, evaluation plays a key role as represented
by Vakola:

> [A]nother scope of evaluation is knowledge construction
> and capacity building … Evaluation facilitates the proc-
> ess of knowledge transfer to similar situations … les-
> sons are transformed into knowledge when they are
> analyzed, systematized, disseminated and internalized
> within an organization through evaluative processes. (pp.
> 814–815)

According to Vakola (2000), evaluation of the re-engineering proc-
ess facilitated the change process and led to better organizational
learning of the business process and the skills required (e.g., infor-
mation communication). Evaluation also helped the implementation
process and identified the major enablers allowing learning to take
place. This led to process innovation such as the introduction of im-
proved technological capabilities in the companies. Another outcome
of the evaluation process was the realization by the companies in-
volved that people management and organizational characteristics
play a major role in the re-engineering change process. Vakola's study
speaks in a very direct way to the power of evaluation to influence
OLC and ultimately lead to significant organizational outcomes.

Russ-Eft, Atwood, and Egherman (2002) carried out an evaluation
utilization case study within the American corporate sector. They
inquired into use and non-use of evaluation results and process and
factors contributing to such patterns in the context of an evaluation
of a sales program within a business services organization. Through

satisfaction interviews and surveys they concluded that ultimately the fate of the program was determined by forces outside the evaluation (e.g., rumours of merger). However, despite this apparent non-use of findings, Russ-Eft et al. identified benefits of the evaluation associated more with its process. They claim that the evaluation was useful in the following ways: supporting and reinforcing the program intervention; increasing engagement, self-determination, ownership, and program and organizational development; and enhancing shared understandings.

Summary: In comparison to ECB studies, research on evaluation utilization identified in our sample provides evidence of the effects of evaluation on organization capacity building manifesting as consequences such as organizational reform, the development of shared understanding by members of the program, and organizational phenomena, problem solving, and decision making. Many of these studies showed that collaborative and participatory evaluation approaches can have profound effects on the use of evaluation findings as well as processes. We now turn to a final category of studies that focuses more directly on organizational processes with evaluation playing a more incidental role.

Organizational Learning Studies

Five studies in our sample examine organizational learning processes and outcomes and in doing so highlight the role of organizational support structures, including, in some instances, systematic inquiry. Two of these studies are in the corporate sector, while three are located in the context of schools. First, Matlay (2000) carried out a large-scale survey of organizational learning within the U.K. small-business sector, examining the types of learning that were taking place and the extent to which learning was intentional and formalized (i.e., systematically recorded and monitored). Micro and small businesses' work-based learning was incidental, occurred sporadically, and was often unnoticed or unrecorded. Usually, this incidental knowledge was used in the short term. Time constraints and administrative pressures reduced the owners' or managers' ability to evaluate initial learning-based strategies. Matlay found that incidental learning was less time-consuming and more cost-effective than intentional learning, and that firms that employed intentional learning strategies were slower and more deliberate in their efforts to meet demands. However, incidental learning did not contribute to sustaining competitiveness and organizational growth,

whereas intentional learning influenced considerably the direction, extent, and delivery of a firm's proactive human resource development strategy. Intentional learning involved systematic data collection and review. An Australian multiple case study reported by Terziovski et al. (2000) examined in five organizations the correspondence and mutual dependencies between Total Quality Management (TQM) and Senge's (1990) systemic conception of organizational learning. The authors concluded that TQM principles and concepts underpin the evolution of the learning organization. In some of the cases, inquiry strategies such as the use of management information systems and direct surveys of customers for feedback were evident.

Finally, organizational learning studies in schools reveal some interesting findings regarding organizational culture. Stevenson's (2001) Australian case study looked at the "substance or content of the dialogue and inquiry that occurred within a school that appeared to exemplify the [organizational learning] conditions and processes advocated in the literature" (p. 103). He selected the school for analysis on the basis of its reputation for being a learning school. Close examination revealed that more systematic inquiry was occurring in relation to equity issues within the school than with curriculum. Teachers' views on diversity and discrimination were willingly questioned, but not their views about teaching and learning. Stevenson postulated this to be the case because there existed no common shared value or focus concerning teaching and learning that clearly permeated school goals, policies, and practices in the way that equity and caring were emphasized. Second, the structure of meeting times and obsession with consensus tend to be less conducive to considering issues such as educational values and teaching approaches. Finally, school leadership appeared to be more focused on political empowerment (addressing issues of unequal power relations) than epistemological empowerment (testing validity of knowledge claims).

Scribner et al. (1999) inquired into the role of school improvement processes in creating professional communities within three rural American middle schools. They reported that the extent to which professional communities developed over two years varied across schools and identified promises and pitfalls of school improvement process in this regard. Influential factors were: the leadership style of the principal (most potent); past events and occurrences remembered and passed on through stories and myths; politics of allocating scarce resources (competition); and persistent bureaucratic

organization of schools and organization of teacher work (e.g., grade level collaboration). The authors maintain that it is difficult to sort out whether cultural change is attributable to school improvement processes or pre-existing conditions. Nevertheless, school improvement provided the organizational architecture supporting professional communities and, potentially, double-loop learning. Kruse (2003) provided case study data supporting organizational memory as an analytic construct for understanding organizational outcomes and performances. Her data support a view of knowledge management and organizational memory systems as recursive social systems with the acquisition, retention, and retrieval of data and information being integral. Kruse made the case, however, that organizations should rely more on less formal forms of knowledge such as belief structures, judgement, and consistent intuition, although use of harder evidence factors into organizational learning.

Summary: Studies of organizational learning in our sample referred in only incidental ways to the role of systematic inquiry in developing an organizational culture of learning. These studies do, however, underscore the importance of organizational support structures in developing cultures of learning and also clarify some organizational consequences of the learning culture.

DISCUSSION AND CONCLUSIONS

Our sample of studies for the present review and integration, though comprehensive, is not exhaustive. Nevertheless, it provides a very good estimate of the current state of research on evaluative inquiry as it relates to the development of cultures of organizational learning. The conceptual framework in Figure 1 is a product of the findings of the current review integrated with prior conceptual understandings. The extent to which we may have confidence in it as a way to conceptualize evaluation as an organizational learning system remains an empirical question at this juncture.

Methodologically, we conclude that research on evaluation as an organizational learning system is highly underdeveloped and that confidence in knowledge in this domain depends quite significantly on the extent to which more, and different, research occurs. Many of the studies that we located, some would argue, are not empirical studies at all. We refer here to the reflective case study, a mode of inquiry that is a popular form of inquiry into other evaluation-related questions. These studies are undeniably valuable sources of knowledge and

understanding of rather complex processes, particularly with regard to such abstract cultural constructs as organizational learning capacity and organizational readiness for evaluation. Often written from a constructivist epistemological orientation, these studies help to illuminate in fine detail some of the more interesting issues for consideration. They are necessary and important. Yet they are limited. These narratives are based on real lived experiences with actual case examples, but in the end they represent one person's or group's interpretation of events. They leave to the reader the challenge of deciding veracity and generalizability and provide no clear means for doing so. Although they appear from a constructivist perspective, the reflective case studies tend to provide only one side of the story, from one perspective. The reader is left pondering the contestability of the narrative. Would organizational and program practitioners agree with the evaluator's take on events, influences, and consequences? In what ways would they disagree? Why?

Even in the studies that embraced principles of systematic inquiry more directly and provided bases for evaluating their trustworthiness and validity of conclusions, we did not observe a great deal of variety in method or sophistication. Only a handful of studies employed longitudinal designs, a long-recognized powerful means of studying complex phenomena such as evaluation use (Weiss, 1981). Several did use multiple methods of inquiry, thereby invoking principles of triangulation. Nevertheless, most studies were in essence qualitative. They provide rich detail and vivid snapshots of the variables of interest and relationships among them. But they tend to generate far more questions than they answer. Our view is that the field is in need of sophisticated methodologies designed to test relationships and assumptions and to confirm understanding about important ones being put forward. Methods such as concept mapping could prove to be extremely valuable in helping to capture the essence of complex cultural phenomena and provide signposts to their formal measurement. Quantitative studies that capitalize on sophisticated estimations of latent structures, path analyses, and hierarchical explanations would serve to move the research agenda forward in confirmatory ways. To the extent that such studies could be longitudinal, the development of the knowledge base would benefit.

Substantively, our review identified three types of study: ECB studies, which inquire into direct and indirect capacity building interventions; evaluation use studies, which capitalize on sensible re-conceptions of use in process terms and links to capacity building

and organizational and program effects; and organizational learning studies, which tend to speak incidentally to the role and influence of evaluative inquiry. Each of these types of studies have offered keen and valued insight into the variables and relationships of interest. They report on widely varying contexts ranging from the government policy arena to public sector agencies and organizations to corporate organizational cultures. They report on processes, influences, and consequences in divergent national cultures from North America, Europe, Australia, and elsewhere. However, they do not address the complex array of variables and relationships in Figure 1 in a comprehensive way. Needed are studies guided by research questions that capture more of the variables represented in the figure and examinations of their interrelationships. To what extent can ECB initiatives affect OLC and lead to the development of effective learning organizations? Under what conditions will process use effectively contribute to sustained capacity building? How do organizations come to develop effective and sustainable cultures of continuous inquiry?

Notably absent from the foregoing review is an explicit discussion about micro-politics and the impact of power relations on the extent to which evaluation gets integrated into organizations. Wildavsky (1972) identified long ago the prominent influence of internal organizational coalitions among key actors and how such configurations might act to shape the culture of the organization. What are the implications for organizational self-evaluation? How will negative evaluation findings be handled? Despite the focus on evaluation for learning, evaluation by definition is judgemental, and so may carry with it summative connotations. Further research in this vein is required to sharpen understanding of how organizational readiness for evaluation may be enhanced.

From a practical point of view, several forces that influence the integration of evaluation into the organizational culture were identified. These have implications for ECB and organizational practice, keeping in mind caveats mentioned above concerning confidence in the knowledge base. Facilitative leadership and modelling; ongoing training and technical support; the existence of prior knowledge, skill, and facility with evaluation logic; the availability of resources for evaluation; and exigencies for evidence about program and organizational performance and results are among the most salient of these forces. Preliminary needs assessment and tailoring ECB initiatives to the stage of development are suggested to be promising

strategies (see, e.g., Gibbs et al., 2002). Regardless, it seems clear that organizations will need to experience success in using evaluation before it is integrated as a cultural norm. Such a norm would be characterized by minimal data anxiety, data literacy, and inquiry mindedness (Earl & Katz, 2002).

It is our hope that the present review and the resulting framework will serve to stimulate a great deal of research activity within this domain. It is clear that the integration of evaluation into existing organizational cultures is a challenge of enormous proportions, but one, we believe, that is undeniably worthy of figuring out. While suggested practical implications and strategies may seem worthwhile, they would stand to benefit from the fruits of a more thoroughly developed knowledge base. As Torres and Preskill (2001) put it, although organizations have moved beyond just being interested to actually beginning the transition to an organizational learning approach to evaluation, there is a need for continued understanding and dialogue about how existing tools, methods, and frameworks for learning-centred evaluation operate in practice. Targeted research programs that capitalize on diversity of method, context, and perspective, we would argue, are likely to be a powerful mode for bringing about such understanding.

ACKNOWLEDGEMENT

REFERENCES

Alkin, M.C., Daillak, R., & White, B. (1979). *Using evaluations*. Beverly Hills, CA: Sage.

Balthasar, A., & Rieder, S. (2000). Learning from evaluations: Effects of the evaluation of the Swiss Energy 2000 programme. *Evaluation, 6*, 245–260.

Biott, C., & Cook, T. (2000). Local evaluation in a national early years excellence centres pilot programme. *Evaluation*, 6, 399–413.

Boaz, A. (2002). Pro-active evaluators: Enabling research to be useful, usable and used. *Evaluation, 8*, 440–453.

Brandon, P.R., & Higa, T.A.F. (2004). An empirical study of building the evaluation capacity of K–12 site-managed project personnel. *Canadian Journal of Program Evaluation, 19*(1), 125–142.

Chacon-Moscoso, S., Anguera-Argilaga, M.T., Perez-Gil, J.A., & Holgado-Tello, F.P. (2002). A mutual catalytic model of formative evaluation: The interdependent roles of evaluators and local programme practionners. *Evaluation, 8*, 413–432.

Compton, D., Baizerman, M., Preskill, H., Rieker, P., & Miner, K. (2001). Developing evaluation capacity while improving evaluation training in public health: The American Cancer Society's Collaborative Evaluation Fellow's Project. *Evaluation and Program Planning, 24*, 33–40.

Compton, D.W., Baizerman, M., & Stockdill, S.H. (Eds.). (2002). *The art, craft and science of evaluation capacity building*. *New Directions for Evaluation, No. 93*. San Francisco: Jossey-Bass.

Compton, D., Glover-Kudon, R., Smith, I.E., & Avery, M.E. (2002). Ongoing capacity building in the American Cancer Society (ACS) 1995–2001. In D.W. Compton, M. Baizerman, & S.H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building*. *New Directions for Evaluation, No. 93* (pp. 47–61). San Francisco: Jossey-Bass.

Connor, R.F., & Tanjasiri, S.P. (1999). Communities evaluating community-level interventions: The development of community-based indicators in the Colorado Healthy Communities Initiative. *Canadian Journal of Program Evaluation*, *Special Issue,* 115–136.

Cousins, J.B. (1996). Understanding organizational learning for educational leadership and school reform. In K.A. Leithwood (Ed.), *International handbook of educational leadership and administration* (pp. 575–640). Boston: Kluwer.

Cousins, J.B. (2003). Utilization effects of participatory evaluation. In T. Kellaghan & D. Stufflebeam (Eds.), *International handbook of educational evaluation* (pp. 245–266). Boston: Kluwer.

Cousins, J.B., & Earl, L.M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis, 14*, 397–418.

Cousins, J.B., & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation. New Directions for Evaluation, No. 80* (pp. 3–23). San Francisco: Jossey-Bass.

Davis, H., & Salasin, S. (1983). The utilization of evaluation. In E. Streuning & M.G. Brewer (Eds.), *The university edition of the handbook of evaluation research* (pp. 381–427). Beverly Hills, CA: Sage. (Original work published in 1975).

Earl, L., & Katz, S. (2002). Leading schools in a data-rich world. In K. Leithwood & P. Hallinger (Eds.), *Second international handbook of educational leadership and administration* (pp. 1003–1022). Boston: Kluwer.

Fiol, C.M., & Lyles, M.A. (1985). Organizational learning. *Academy of Management Review, 10*, 803–813.

Forss, K., Cracknell, B., & Samset, K. (1994). Can evaluation help an organization to learn? *Evaluation Review, 18*, 574–591.

Garvin, D.A. (1993, July-August). Building a learning organization. *Harvard Business Review,* 78–91.

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

Gilliam, A., Barrington, T., Davis, D., Lacson, R., Uhl, G., & Phoenix, U. (2003). Building evaluation capacity for HIV prevention programs. *Evaluation and Program Planning, 26*, 133–142.

Goh, S. (2000). Towards a learning organization: The strategic building blocks. *SAM Advanced Management Journal, 63*(2), 15–22.

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

Huber, G.P. (1991). Organizational learning: The contributing processes and the literature. *Organizational Science, 2*, 88–115.

Jacobs, C. (2000). The evaluation of educational innovation. *Evaluation, 6*(3), 261–280.

Katz, S., Sutherland, S., & Earl, L. (2002). Developing an evaluation habit of mind. *Canadian Journal of Program Evaluation, 17*(2), 103–119.

King, J.A. (2002). Building the evaluation capacity of a school district. In D.W. Compton, M. Baizerman, & S.H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, No. 93* (pp. 63–80). San Francisco: Jossey-Bass.

Kruse, S.D. (2003). Remembering as organizational memory. *Journal of Educational Administration, 41*, 332–347.

Kushner, S. (2002). I'll take mine neat: Multiple methods but a single methodology. *Evaluation, 8*, 249–258.

Lant, T.K., & Mezias, S.J. (1992). An organizational learning model of convergence and reorientation. *Organization Science, 3*(1), 47–71.

Lee, L. (1999). Building capacity for school improvement through evaluation: Experiences of the Manitoba School Improvement Program Inc. *Canadian Journal of Program Evaluation, 14*(2), 155–178.

Leithwood, K., Aitken, R., & Jantzi, D. (2001). *Making schools smarter: A system for monitoring school and district progress* (2nd ed.). Thousand Oaks, CA: Corwin.

Lundberg, C.C. (1989). On organizational learning: Implications and opportunities for expanding organizational development. *Research in Organizational Change and Development, 3*, 61–82.

Mackay, K. (2002). The World Bank's ECB experience. In D.W. Compton, M. Baizerman, & S.H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, No. 93* (pp. 81–107). San Francisco: Jossey-Bass.

Matlay, H. (2000). Organisational learning in small organisations: An empirical overview. *Education and Training, 42*(4–5), 202–210.

Milstein, B., Chapel, T.J., Wetterhall, S.F., & Cotton, D.A. (2002). Building capacity for program evaluation at the Centers for Disease Control and Prevention. In D.W. Compton, M. Baizerman, & S.H.

Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, No. 93* (pp. 27–46). San Francisco: Jossey-Bass.

Milstein, B., & Cotton, D. (2000). Thoughts for presidential strand on building evaluation capacity. EVALTALK listserve <EVALTALK@ BAMA.UA.EDU>, March 28, 12:15 p.m.

Morabito, S. (2002). Evaluator roles and strategies for expanding evaluation process influence. *American Journal of Evaluation, 23*, 321–330.

O'Sullivan, R.G., & D'Agostino, A. (2002). Promoting evaluation through collaboration: Findings from community-based programs for young children and their families. *Evaluation, 8*, 372–387.

Owen, J.M., & Lambert, F.C. (1995). Roles for evaluation in learning organizations. *Evaluation, 1*, 237–250.

Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Newbury Park, CA: Sage.

Patton, M.Q. (1998). Discovering process use. *Evaluation, 4*, 225–233.

Patton, M.Q. (1999). Organizational development and evaluation. *Canadian Journal of Program Evaluation, Special Issue,* 93–114.

Patton, M.Q., Grimes, P.S., Gutherie, K.M., Brennan, N.J., French, B.D., & Blyth, D.A. (1977). In search of impact: An analysis of the utilization of federal health evaluation research. In C.H. Weiss (Ed.), *Using social research in public policy making* (pp. 141–163). New York: D.C. Heath.

Porteous, N.L., Sheldrick, B.J., & Stewart, P.J. (1999). Enhancing managers' evaluation capacity: A case study from Ontario public health. *Canadian Journal of Program Evaluation, 14*(2), 137–154.

Preskill, H. (1994). Evaluation's role in enhancing organizational learning. *Evaluation and Program Planning, 17*, 291–297.

Preskill, H., & Caracelli, V. (1997). Current and developing conceptions of use: Evaluation Use TIG survey results. *Evaluation Practice, 18,* 209–225.

Preskill, H., & Torres, R. (1999). *Evaluative inquiry for organizational learning*. Twin Oaks, CA: Sage.

Preskill, H., Zuckerman, B., & Matthews, B. (2003). An exploratory study of process use: Findings and implications for future research. *American Journal of Evaluation, 24*, 423–442.

Robinson, T., & Cousins, J.B. (2004). Internal participatory evaluation as an organizational learning system: A longitudinal case study. *Studies in Educational Evaluation, 30*, 1–22.

Russ-Eft, D., Atwood, R., & Egherman, T. (2002). Use and non-use of evaluation results: Case study of environmental influences in the private sector. *American Journal of Evaluation, 23*, 19–31.

Scribner, J., Cockrell, K., Cockrell, D., & Valentine, J. (1999). Creating professional communities in schools through organizational learning: An evaluation of a school improvement process. *Educational Administration Quarterly, 35*, 130–160

Seiden, K. (2000). *Development and validation of the "organizational readiness for evaluation" survey instrument*. Unpublished doctoral dissertation, University of Minnesota, Minneapolis.

Senge, P.M. (1990). *The fifth discipline: The art and practice of organizational learning*. New York: Doubleday.

Shulha, L., & Cousins, J.B. (1997). Evaluation use: Theory, research and practice since 1986. *Evaluation Practice*, *18*, 195–208.

Silins, H., Mulford, W., & Zarins, S. (2002). Organizational learning and school change. *Educational Administration Quarterly, 38*, 613–642.

Smith, C.L., & Freeman, R.L. (2002). Using continuous system level assessment to build school capacity. *American Journal of Evaluation, 23*, 307–319.

Stevenson, J.F., Florin, P., Mills, D.S., & Andrade, M. (2002). Building evaluation capacity in human service organizations: A case study. *Evaluation and Program Planning, 25*, 233–243.

Stevenson, R.B. (2001). Shared decision making and core school values: A case study of organizational learning. *International Journal of Educational Management, 15*, 103.

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, No. 93* (pp. 7–25). San Francisco: Jossey-Bass.

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

Tang, D., Cowling, D.W., Koumjian, K., Roeseler, A., Lloyd, J., & Rogers, T. (2002). Building local program evaluation capacity toward a comprehensive evaluation. In R. Mohan, D.J. Berstein, & M.D. Whitsett (Eds.), *Responding to sponsors and stakeholders in complex evaluation environments. New Directions for Evaluation, No. 95* (pp. 39–56). San Francisco: Jossey-Bass.

Terziovski, M., Howell, A., Sohal, A., & Morrison, M. (2000). Establishing mutual dependence between TQM and the learning organization: A multiple case study analysis. *The Learning Organization, 7*, 23–31.

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

Torres, R.T., Preskill, H., & Piontek, M.E. (1996). *Evaluation strategies for communicating and reporting: Enhancing learning in organizations*. Thousand Oaks, CA: Sage.

Vakola, M. (2000). Exploring the relationship between the use of evaluation in business process re-engineering and organizational learning and innovation. *Journal of Management Development, 19*, 812–835.

Valovirta, V. (2002). Evaluation utilization as argumentation. *Evaluation*, 8, 60–80.

VanderPlaat, M., Samson, Y., & Raven, P. (2001). The politics and practice of empowerment evaluation and social interventions: Lessons from the Atlantic community action program for children regional evaluation. *Canadian Journal of Program Evaluation, 16*(1), 79–98.

Weiss, C.H. (1981). Measuring the use of evaluation. In J.A. Ciarlo (Ed.), *Utilizing evaluation: Concepts and measurement techniques* (pp. 17–33). Beverly Hills, CA: Sage.

Wildavsky, A. (1972). The self-evaluating organization. *Public Administration Review, 32*, 515–520.

J. Bradley Cousins, Ph.D., is professor of educational administration at the Faculty of Education, University of Ottawa, with a specialization in program evaluation theory and practice. He is the 1999 winner of the Contributions to Evaluation in Canada award of the Canadian Evaluation Society. He has published numerous articles and books and is currently completing *Participatory Evaluation Up Close: An Integration of Research-Based Knowledge*.

Swee C. Goh, Ph.D., is professor of organizational behaviour and human resource management in the School of Management, University of Ottawa. His research interests are in the areas of organizational learning, knowledge transfer, and change management. His research has been published in journals such as *The Learning Organization: An International Journal*, *European Management Journal*, *International Journal of Organizational Theory and Behavior*, and *The Academy of Management Journal*. He was recently awarded the 2004 University of Ottawa Excellence in Education Prize.

Shannon Clark is a doctoral candidate at the Faculty of Education at the University of Ottawa and has worked as a graduate assistant on this project for the past three years. Her interests are in developmental learning models and sports psychology.

Linda E. Lee is a partner in Proactive Information Services Inc., a Canadian company specializing in social research and program evaluation. She works extensively in educational evaluation in Canada and internationally. She is a past national president of the Canadian Evaluation Society and recipient of the 2003 Award for Service to the Canadian Evaluation Society.