PERCEPTIONS OF EVALUATION CAPACITY BUILDING IN THE UNITED STATES: A DESCRIPTIVE STUDY OF AMERICAN EVALUATION ASSOCIATION MEMBERS

Dreolin Nesbitt Fleischer
Claremont Graduate University
Claremont, California

Christina A. Christie
University of California, Los Angeles
Los Angeles, California

Katherine B. LaVelle
Five Acres
Altadena, California

Abstract: This article offers a descriptive picture of American Evaluation Association (AEA) members' attitudes and perceptions related to evaluation capacity building (ECB). For this study, we analyzed data that were originally collected in the spring of 2006 from 1,140 AEA members in the United States on evaluation use. The current study is an attempt to add to the ECB knowledge base by describing respondents' views concerning (a) the importance of ECB as an evaluation approach, (b) the role of evaluators in undertaking ECB-related activities, (c) ECB-related factors that influence use, and (d) the extent to which evaluation activities foster organizational learning and change outcomes. Respondents are largely familiar with ECB and agree that building evaluation capacity is a role of the evaluator. Linkages between organizational learning and ECB were supported. Learning-focused organizational outcomes were rated more favourably than change-focused organizational outcomes.

Résumé :   Cet article donne un aperçu descriptif des attitudes et des perceptions des membres de l'American Evaluation Association (AEA) liés au renforcement des capacités d'évaluation (RCÉ). Dans cette étude, nous avons analysé les données sur l'utilisation d'évaluations qui ont été initialement recueillies au printemps 2006 et provenant de 1 140 membres de l'AEA aux États-Unis. L'étude actuelle tente d'ajouter à la base de connaissances sur le RCÉ en décrivant des opinions de participants concernant :

Corresponding author: Dreolin N. Fleischer, 4 Trowbridge Place, Apt. 6C, Cambridge, MA, USA 02138; <dreolin@gmail.com>

(a) l'importance du RCÉ en tant que méthode d'évaluation, (b) le rôle des évaluateurs dans l'exécution d'activités liées au RCÉ, (c) des facteurs liés au RCÉ qui en influencent l'utilisation, et (d) le degré auquel ces activités d'évaluation stimulent l'apprentissage organisationnel et changent les résultats. La plupart des répondants connaissent le RCÉ et acceptent que le renforcement des capacités d'évaluation est un rôle assumé par l'évaluateur. Les liens entre l'apprentissage organisationnel et le RCÉ ont été soutenus. Les résultats organisationnels axés sur l'apprentissage sont notés plus favorablement que les résultats organisationnels axés sur le changement.

INTRODUCTION & BACKGROUND

Evaluation Capacity Building

Over the past decade, interest in the association between evaluative inquiry and organizational learning and change has grown significantly (e.g., Cousins, Goh, Clark & Lee, 2004; Earl & Katz, 2002; Owen & Rogers, 1999; Preskill & Torres, 1999; Preskill, Zuckerman, & Matthews, 2003; Russ-Eft & Preskill, 2001; Shulha, 2000; Taut, 2007). Along with this expanding area of evaluation scholarship has also emerged increased interest in the concept of evaluation capacity building (ECB). To date, there is little empirical research on ECB. A recent review of the evaluation literature by Cousins et al. (2004) found only 10 empirical studies that specifically examined the implementation or influence of direct (e.g., training) and indirect (e.g., incidental learning by experience) ECB strategies in evaluation practice. These existing studies did, however, look at ECB from a variety of different contexts, including schools (e.g., Sutherland, 2004), government (e.g., Boaz, 2002; Mackay, 2002), public health/social programs (e.g., Compton, Glover-Kudon, Smith, & Avery, 2002; Preskill et al., 2003), and private contexts (e.g., Matlay, 2000), to name a few. As interest in ECB continues, emerging definitions and conceptual frameworks help to guide further research. To contribute to the growing knowledge base, this study uses secondary data analysis to present a descriptive profile of a sample of American Evaluation Association (AEA) members in relation to the topics of ECB, organizational learning and change, and evaluation utilization.

Definitions and conceptualizations of evaluation capacity building are discussed at greater length in the other articles included in this issue; therefore we offer the two definitions of ECB that framed our

thinking about the topic and guided this study. First is the commonly cited definition by Stockdill, Baizerman, and Compton (2002): "ECB is the intentional work to continuously create and sustain overall organizational processes that make quality evaluation and its uses routine" (p. 14). Second is a more recent and comprehensive definition put forth by Preskill and Boyle (2008):

> Evaluation capacity building involves the design and implementation of strategies to help individuals, groups, and organizations learn about what constitutes effective, useful, and professional evaluation practice. The ultimate goal of evaluation capacity building is sustainable evaluation practice—where evaluation members continuously ask questions that matter, collect, analyze, and interpret data, and use evaluation findings for decision-making and action. For evaluation practice to be sustained, organization members must be provided leadership support, incentives, resources, and opportunities to transfer their learning about evaluation to their everyday work. Sustainable evaluation practice also requires the development of systems, processes, policies, and plans that help embed evaluation work into the way the organization accomplishes its strategic goals and mission. (pp. 3–4)

This second definition is particularly useful in extending our understanding of ECB because it conveys more than simply what ECB is—it offers examples of how and with whom ECB is accomplished.

Evaluation Capacity Building and Organizational Learning

Within the field, evaluation capacity building is considered a popular and promising approach for fostering organizational learning (e.g., Cousins et al., 2004; Preskill & Boyle, 2008; Preskill & Torres, 1999). The relationship between ECB and organizational learning and change is being increasingly clarified in the literature, largely due to the recent development of conceptual frameworks (Cousins et al., 2004; Cousins, Goh, & Elliot, 2007a; Preskill & Boyle, 2008). These guiding frameworks serve not only to highlight what organizational and evaluative concepts are important to building evaluation capacity, but also elucidate possible ways that these concepts might interact to produce successful evaluation capacity development and sustainment. From their extensive literature review completed in

2004, Cousins et al. developed a conceptual framework that depicts how evaluation capacity overlaps with organizational learning capacity to create what is referred to as "organizational readiness for evaluation." An expanded conceptual model is in development (Cousins et al., 2007a) that proposes a set of complex linkages among multiple concepts related to ECB: (a) sources of knowledge, skills, and abilities; (b) organizational support structures; (c) capacity to *do* evaluation; (d) evaluative inquiry; (e) capacity to *use* evaluation; (f) organizational learning capacity; and (g) organizational consequences. Another multidimensional conceptual framework, offered by Preskill and Boyle (2008), shows how the design, implementation, and results of ECB initiatives may be influenced by important organizational factors that support organizational learning, such as an organization's culture, leadership, and communication systems. Together, these conceptual frameworks suggest that an organization's capacity to learn and develop, as well as its use of evaluation results, plays an important role in ECB efforts and, ultimately, in organizational change (e.g., informed decision-making and program improvement).

Theoretical and conceptual connections between ECB and organizational learning are becoming more prominent in the evaluation literature, and thus it is important to briefly state how we view organizational learning and change. There are several distinct definitions of organizational learning; however, we refer to the description by Senge (1990) of learning organizations as "organizations where people continually expand their capacity to create the results they truly desire, where new and expansive patterns of thinking are nurtured, where collective aspiration is set free, and where people are continually learning how to learn together" (p. 3). Learning in organizations is characterized as occurring at the individual, group, and organizational level, with each level of learning being interdependent (Preskill & Torres, 1999). In addition, the extent to which an organization experiences learning depends, in large part, on factors such as the organization's infrastructure and culture (Preskill & Torres, 1999) and core strategic building blocks (Goh, 2000). An organizational culture that supports evaluative inquiry is essential for organizational learning to occur (e.g., Cousins et al., 2004; Cousins et al., 2007a; Preskill & Boyle, 2008; Preskill & Torres, 1999).

Although ECB is only one of many strategies for enhancing organizational learning (other strategies include, e.g., strategic planning and management information systems), it is an especially promis-

ing approach. Building evaluation capacity is directly associated with the increased knowledge and skills of individuals, as well as sustained evaluative practice within the organization (e.g., Cousins et al., 2004; Preskill & Torres, 1999). Preskill and Boyle (2008) have recently identified 10 ECB activities that increase organizational members' knowledge of evaluation and build specific skills necessary to conduct effective evaluations (e.g., attending training workshops, using technology during the evaluation process, receiving individualized coaching/mentoring). However, compared to knowledge and skill acquisition, a less common (although still desired) result of ECB efforts in organizations is the continuation of evaluative practice. One explanation for this lack of sustained evaluative inquiry in organizations might be that continued evaluative application requires the transfer of learning to practice, something that involves both an appreciation for and valuing of evaluative inquiry as well as a motivation toward action (Preskill & Torres, 1999). Another possibility is that continuous ECB practice might depend on the extent to which an organization utilizes its evaluation results (Cousins et al., 2007a). The present study explored this connection between ECB and aspects of organizational learning and change, as well as the role of evaluation use from the perspective of a sample of AEA members in the United States.

Evaluation Utilization and Capacity Building

The notion that evaluation use can lead to continued evaluation capacity building is evident in Cousins et al.'s (2007a) latest conceptual model and is being empirically tested in two studies: one looking at internal evaluators' perceptions of ECB and evaluation use in their organizations (Cousins, Goh & Elliot, 2007b) and another examining the differences in evaluation capacity building understanding in members of government and voluntary sector organizations (Cousins, Goh, Elliot, & Aubry, 2008). The literature suggests that there is an important connection between evaluation utilization and evaluation capacity building within an organizational context, such that the more an organization uses their evaluation results, the more likely ECB efforts will continue over time. According to Cousins et al. (2004), more research attention should be devoted to understanding the capacity to *use* evaluation and not only the capacity to *do* evaluation: "The integration of evaluation into the culture of organizations has … as much to do with the consequences of evaluation as it does the development of skills and knowledge of evaluation logic and methods" (p. 101).

Building capacity to use evaluation requires that members of an organization understand and feel ownership of the evaluation process and findings, and are sufficiently prepared to use the evaluation results to enact organizational change. Patton's (2008) utilization-focused approach to evaluation (UFE), emphasizing intended users (i.e., people using the evaluation results) and intended uses (i.e., the intended users' specific information needs), can increase an organization's capacity for evaluation utilization. Focusing the evaluation on intended uses by intended users involves the active participation of primary stakeholders in the evaluation process, which, in turn, facilitates rapport-building, enhances stakeholder buy-in, and leads to process use (i.e., changes to individuals' thinking and behaving and organizations' procedures and culture as a result of participating in an evaluation). Patton (2008) refers to this process of facilitating evaluation utilization as "The Psychology of UFE." The UFE approach is thought to lead not only to enhanced use of evaluation results, but also to the promotion of evaluation capacity building. Thus, from the literature reviewed here, it seems appropriate to view ECB as an iterative process such that ECB leads to organizational learning and evaluation use, which then contributes to developing further evaluation capacity and eventual organizational change.

Despite the recent interest in fostering evaluation capacity within organizations, there remains a need for more empirical study of ECB and how it relates to organizational learning and evaluation utilization. Specifically, the field would benefit from additional research looking at the organizational factors that might influence the feasibility or effectiveness of ECB initiatives. Our study is an attempt to add to the ECB knowledge base by presenting a descriptive profile of AEA members in relation to their perception of (a) the importance of ECB as an evaluation approach, (b) the role of evaluators in undertaking ECB-related activities, (c) ECB-related factors that influence use, and (d) the extent to which evaluation activities foster organizational learning and change outcomes. Furthermore, this study offers new insights on the relationship between evaluation use and ECB by examining the views of current evaluators in terms of the factors that they think influence the use of evaluation results as well as promote ECB. In all, we think that our analysis of the data from the AEA member use survey administered in 2006 reveals some interesting and important findings regarding sampled members' attitudes and perceptions related, both directly and indirectly, to evaluation capacity building.

METHOD

For the purposes of this study, we performed exploratory analyses on an existing dataset. The data were originally collected by two of the authors (Fleischer & Christie) in the spring of 2006 as part of a study looking at evaluation use. At the time of data collection, 10 years had passed since Preskill and Caracelli (1997) conducted their survey on evaluation use with members of the Evaluation Use Topical Interest Group (TIG) within the American Evaluation Association (AEA). The 2006 study was designed to offer an additional perspective as well as be qualitatively compared to the Preskill and Caracelli study; therefore, we based our survey instrument on their original survey instrument. Modifications to the original survey included adding a few additional questions to gather new information, rewording a few items, and removing several questions to decrease the overall length of the survey. Further modifications included administering the survey online and widening the sample to US members of AEA, not just Evaluation Use TIG members.

After receiving both permission from AEA's Executive Board and a list of their membership, we e-mailed 3,824 AEA members in the United States an invitation to participate in the study with a link to the online survey. More than half of the total sample responded within the first week that the survey was open, with a majority of respondents completing the survey within 24 hours of receiving the invitation. Members who did not respond within the first week were contacted up to two additional times by e-mail. The response rate was 30% (1,140/3,824). Although receiving 1,140 responses provided a reasonable sample size, we recognize that a response rate of 30% makes generalizing the results of the study to the entire US membership of AEA not advisable. However, the sample from our study (surveyed in May-June 2006) does resemble the sample from AEA's internal scan of their membership (American Evaluation Association, 2008) administered in September 2007-January 2008 on three key background variables (type of organization they work in, professional identity, and years of experience), thus speaking to the broader representativeness of our sample. An important caveat is that the wording of the questions and the exact answer categories are not identical between the two surveys, so comparisons should be considered approximate and tentative.

The internal scan yielded a response rate of 49% (2,657/5,460). Of the members who responded to the internal scan and responded to

the question regarding where they are contracted to do evaluation work, 29% worked at a college/university (compared to 35% in our sample); 19% worked in a research, evaluation, or consulting firm (compared to 22% in our sample); 7% worked in a nonprofit organization (compared to 20% in our sample); and 12% worked at either the local, state, or federal government level (compared to 11% in our sample). Forty-nine percent of members who responded to the internal scan indicated "evaluator" as their primary professional identity, compared to 51% in our sample. In terms of years of experience in the field of evaluation, 29% of our sample reported less than 5 years, while 33% of the internal scan sample reported the same experience level. Twenty-four percent of the internal scan sample reported 6–10 years of experience, compared to 28% in our sample. The internal scan data showed that 16% of the sample reported 11–15 years of experience and 27% reported 16+ years of experience. On our survey, we also asked respondents to report experience level according to 11–20 years (24%) and 20+ years (19%).

Instrument

Our survey consisted mostly of Likert-type ordinal response scale items, with the exception of a ranking item, a forced choice item, an open-ended item, and several background items with nominal scales. The background questions related to professional identity, experience and knowledge level with evaluation, whether they predominantly work as internal or external evaluators, the type of organization they predominantly work in, the average number of evaluations they conduct each year, the size of evaluations they typically conduct, the length of time they have been conducting evaluations, and whether they are Evaluation Use TIG members.

Respondents were asked to rank the top three purposes of evaluation according to their perceptions of their actual importance and ideally how important they would be (1 question; 2 items). Respondents were given a list of eight approaches to conducting evaluation and asked to rate each on level of importance (1 question; 8 items); they were then asked to select the most important approach and least important approach from the list of eight (1 question; 2 items). Next, respondents were asked the extent to which they agree or disagree with 10 statements on various roles evaluators might play in an evaluation (1 question; 10 items).

Consistent with the 1996 survey, a definition for use of evaluation findings was offered to provide context for the remainder of survey

items, which all focused on evaluation use. Seven statements focused on the nature and extent of stakeholder involvement in an evaluation. Respondents were then asked to rate the influence of 15 activities that might have an impact on evaluation use. Ten statements focused on the extent to which evaluation can foster organizational outcomes. Finally, respondents were asked to report the extent to which they believe "nonuse of evaluation results," "intentional misuse of evaluation results," and "unintentional misuse of evaluation results" were a problem as well as rate the frequency that six types of misuse occur, including seeing clients "dismiss undesirable results," "selectively report results," and "exaggerate results."

Internal Consistency Reliability

Responses to all of the Likert-type scale items on the survey were factor analyzed using a Principal Components Analysis with a varimax rotation that resulted in a five-factor solution retaining 29 of the 50 items. The optimal solution was selected based on the Kaiser rule, amount of variance explained, and the fact that these five factors made theoretical sense. Cronbach's Alpha was conducted for the five factors; each factor had "good" to "high" reliability. Factor 1 included 10 items (α = .880) and was named "organizational outcomes" because it included each of the 10 items grouped together on the survey about the ability of evaluation activities to foster organizational outcomes. Factor 2 included 6 items (α = .757) and was named "stakeholder involvement" because it consisted of items from three separate portions of the survey that asked about how stakeholders should be included in the evaluation. Factor 3 included 4 items (α = .733) and was named "planning for use" because it consisted of items that asked respondents to rate the extent to which certain factors can increase evaluation use. Factor 4 included 5 items (α = .807) and was named "user-oriented approaches" because it included five of the user-oriented approach items from the series of items that asked respondents to rate the importance of a variety of different evaluation approaches. Finally, factor 5 included 4 items (α = .638) and was named "evaluator responsibilities" because it consisted of items that asked respondents to indicate whether they agree or disagree with various roles the evaluator can take within an evaluation (Fleischer & Christie, 2009).

Respondents

The majority of respondents (96%) reported that evaluation was either their primary or secondary professional identity. Novices were in the minority, as 87% of respondents classified themselves as hav-

ing either an intermediate or an advanced level of knowledge and experience. Half of the respondents reported typically working on a combination of small- and large-scale evaluations, but of those who typically work with one size over another, more respondents worked on small-scale evaluations (34%) than large-scale evaluations (14%). The largest number of respondents reported working within a college or university (35%), followed by respondents working within private business or consulting (22%) and nonprofit organizations (20%). Twenty-two percent of respondents reported that they were members of AEA's Evaluation Use TIG, 55% reported not being a member of the Evaluation Use TIG, and 24% reported "I don't know/don't remember" to the question (this adds up to 101% due to rounding).

The 2006 instrument was developed to collect data from members of AEA on the topic of evaluation use. While ECB was not the primary emphasis of the initial survey study, it was the focus of a key section of the survey, allowing for a secondary data analysis to address specific ECB-related research questions.

RESULTS

This article offers descriptive data from US members of the AEA concerning their attitudes and perceptions related directly and indirectly to ECB. In this section, responses to questions explicitly related to ECB will be described first, followed by questions concerning concepts associated with ECB, specifically evaluation use and organizational learning and change.

Evaluation Capacity Building as the Most and Least Important Evaluation Approach

On a scale of 1 to 5 (1 = *not at all important* and 5 = *extremely important*), respondents were asked to rate the importance of eight evaluation approaches: objectives-based, participatory, theory-driven, user-focused, experimental/quasi-experimental, social justice/democratic, empowerment evaluation, and evaluation for capacity building (listed in that order on the survey). Study participants were directed to reflect on the current field of evaluation when rating each approach. The mean response for each item is presented in Table 1.

For this set of items, respondents were given the option to select "I am unfamiliar with this approach" instead of rating its importance to

the field. Based on the percentage of respondents who indicated they were unfamiliar with each approach, respondents were most unfamiliar with the "social justice" approach (16%), while they were least unfamiliar with the "experimental/quasi-experimental" approach (2%). A relatively small percentage of respondents (6%) indicated that they were unfamiliar with "evaluation for capacity building" as an evaluation approach.[1]

After rating each approach, respondents were asked to select the most and least important approaches to the evaluation field from the list of eight. We are aware that these eight approaches are not mutually exclusive. For example, a theory-driven evaluation can utilize an experimental or quasi-experimental design and an empowerment evaluation is a form of participatory evaluation. Our

Table 1
The Most and Least Important Evaluation Approaches & Mean Importance Rating for Each

	Most important N = 1,102	Least important N = 1,067	Mean* N = 937–1,099	SD	Rank	95% C.I. Lower	95% C.I. Upper
Objectives-based	24% (n = 264)	7% (n = 76)	4.23	.879	2	4.18	4.28
User-focused	21% (n = 235)	2% (n = 24)	4.25	.838	1	4.20	4.30
Participatory	14% (n = 155)	4% (n = 38)	4.03	.964	3	3.97	4.09
Experimental/ quasi-experimental	12% (n = 127)	21% (n = 225)	3.65	1.06	6	3.59	3.71
Theory-driven	11% (n = 122)	18% (n = 188)	3.68	1.06	5	3.62	3.74
Evaluation for capacity building	10% (n = 113)	7% (n = 73)	4.01	.964	4	3.95	4.07
Empowerment evaluation	5% (n = 52)	16% (n = 175)	3.35	1.13	7	3.28	3.42
Social justice/ democratic	3% (n = 34)	25% (n = 268)	3.17	1.11	8	3.10	3.24

* 1 (not at all important) to 5 (extremely important)

intent was to gather feedback on a wide variety of ways evaluators approach conducting their evaluation work, but in doing so we developed a list of "approaches" that belong to separate and overlapping categories. Although it is likely that practicing evaluators draw upon the conceptual tenets of more than a single evaluation approach or model, we believe it is valuable to require respondents to select an approach that they feel is fundamentally least or most important to facilitate comparisons between approaches. Thus, we argue that items within a list do not have to be conceptually the same to select a least or most important item from the list. This issue presents more of a dilemma with the rating of each approach, which is why we use the forced-choice responses rather than the Likert-type responses in further data analysis. As seen in Table 1, 10% (n = 113 of 1,102) of respondents selected ECB as the most important approach of the eight, while 7% (n = 73 of 1,067) of respondents selected ECB as the least important approach.

To better understand how respondents who considered ECB the most important approach, respondents who considered ECB the least important approach, and respondents who did not indicate ECB as either the most or least important approach differed in their responses, a series of Fisher's Exact Tests were conducted comparing these three groups on a series of background variables. We used a Fisher's Exact Test rather than a Chi-Square Test to account for the small cell sizes in the contingency tables when comparing variables with multiple levels. We compared the three-level ECB variable to a series of background variables that included professional identity, knowledge and experience level, role as an evaluator, organization they predominantly work within, the number of evaluations they conduct per year, the size of evaluations they typically work on, years conducting evaluations, and membership in the Evaluation Use TIG. There were no statistically significant differences between individuals who selected ECB as the least important approach, those who selected ECB as the most important approach, and those who did not select ECB as least or most important on any of the background variables.

Organizational Learning

To further compare those who consider ECB most important and those who consider it least important, we examined perceptions of the actual and ideal purposes of evaluation for these two groups. Specifically, from a list of six, respondents ranked what they believed

to be the top three purposes of evaluation based on the *actual* and *ideal* importance of each approach to the field of evaluation. The list of six included investigate the merit or worth of the program, provide information for decision-making, improve programs, generate new knowledge about programs, facilitate organizational learning, and promote social justice (in that order on the survey). Frequencies on these items were compared between the two groups (ECB as the most important approach and ECB as the least important approach). A greater proportion of respondents who selected ECB as the most important evaluation approach ranked "to facilitate organizational learning" as an actual and ideal purpose of evaluation when compared to respondents who selected ECB as the least important evaluation approach. However, for both groups, more respondents rated the purpose of evaluation "to facilitate organizational learning" as ideally important rather than actually important.

Perceived Role of the Evaluator in Promoting Evaluation Capacity Building

Several factors can influence the role of an evaluator, including the purpose of the evaluation, stakeholders' information needs, the evaluator's epistemological preferences, and the evaluation approach used. To better understand AEA members' perceptions of the evaluator's role, respondents were asked to rate the extent to which they agree or disagree (on a scale from 1 to 5 where 1 = *strongly disagree* and 5 = *strongly agree*) that an evaluator should be responsible for 10 different activities, including building evaluation capacity. Table 2 presents the frequencies and mean ratings for each activity.

Overall, respondents strongly endorsed each of the 10 activities. The mean score for "building evaluation capacity" was the fifth highest mean ($M = 4.15$). The activity "becoming program advocates" received a notably lower mean score rating ($M = 2.53$).

Initially, we conducted a series of one-way ANOVAs to examine differences on key variables between three groups: (a) those who selected ECB as the most important approach, (b) those who selected ECB as the least important approach, and (c) those who did not select ECB for either the most or least important approach. We found that mean differences between groups focused on the difference between the first two groups and not the third; therefore, we conducted targeted *t*-tests to look at group differences between the two groups on three variables of interest (that showed statistical significance from the

Table 2
The Evaluator's Role (N = 1,047–1,113)

		Response Distribution*					Mean	SD	Rank	95% C.I.	
	N	(1)	(2)	(3)	(4)	(5)				Lower	Upper
Involving stakeholders in the evaluation process	1,113	<1%	<1%	2%	21%	76%	4.73	.542	1	4.70	4.76
Making the interests of the economically and politically disadvantaged a priority	1,064	2%	9%	33%	36%	20%	3.62	.974	9	3.56	3.68
Facilitating organizational learning	1,103	<1%	2%	9%	46%	43%	4.30	.713	3	4.26	4.34
Becoming program advocates	1,096	16%	36%	31%	13%	4%	2.53	1.04	10	2.47	2.59
Maximizing intended use by intended users	1,091	<1%	3%	10%	35%	51%	4.33	.821	2	4.28	4.38
Being accountable to intended users of the evaluation for intended uses of the evaluation	1,047	2%	8%	14%	37%	40%	4.05	1.01	6	3.99	4.11
Formulating recommendations from evaluation studies	1,103	<1%	3%	11%	38%	48%	4.30	.806	4	4.25	4.35
Becoming involved in evaluation follow-up activities	1,103	<1%	8%	27%	41%	23%	3.77	.921	7	3.72	3.82
Attending to the social consequences of the evaluation	1,085	2%	10%	25%	43%	21%	3.72	.955	8	3.66	3.78
Building evaluation capacity	1,089	<1%	2%	16%	45%	36%	4.15	.797	5	4.10	4.20

Mean = 3.95

1 (strongly disagree) to 5 (strongly agree)

analysis of variance). Respondents who selected ECB as the most important approach agreed more (M = 4.57) that the "evaluator is responsible for facilitating organizational learning" than those who selected ECB as the least important approach (M = 4.10) at a statistically significant level; $t(111)$ = 4.41, p < .001. Also, not surprisingly, respondents who selected ECB as the most important approach agreed more (M = 4.64) that the "evaluator is responsible for building evaluation capacity" than those who selected ECB as the least important approach (M = 3.52) at a statistically significant level; $t(107)$ = 9.92, p < .001.

Evaluation Capacity Building as an Activity That Influences Use

The question of what promotes evaluation use is not new; in fact, use is an area within the field of evaluation that has an extensive research base. Our survey asked respondents questions about whether evaluation capacity building activities, such as "interweaving the evaluation into organizational processes and procedures" and "linking the evaluation to an established pattern of organizational life," increase the likelihood of evaluation use. Specifically, respondents were asked to rate on a scale of 1 to 5 (1 = *not at all influential* and 5 = *extremely influential*) the extent to which 15 different factors increase evaluation use.

As seen in Table 3, on average, all of the factors listed were considered at least *somewhat influential*. The two factors related to organizational processes and life were not endorsed as strongly as other factors: "interweaving the evaluation into organizational processes and procedures" had the 7th highest average out of 15 items and "linking the evaluation to an established pattern of organization life" had the 10th highest average out of 15 items. An independent-samples t-test showed that respondents who selected ECB as the least important approach agreed more (M = 4.37) that "identifying and prioritizing intended users of the evaluation" was the responsibility of the evaluator than those who selected ECB as the most important approach (M = 3.96) at a statistically significant level; $t(150)$ = 3.26, p < .001.

Evaluation and Organizational Outcomes

Respondents were asked to rate the extent to which they believe evaluation fosters 10 separate organizational outcomes, presented in Table 4. Overall, the items received fairly high ratings. The most notable departure from this was the item "evaluation changes social

Table 3
Factors That Influence Use (N = 897–1,004)

		Response Distribution*					Mean	SD	Rank	95% C.I.	
	N	(1)	(2)	(3)	(4)	(5)				Lower	Upper
Planning for use at the beginning of the evaluation	992	<1%	2%	7%	33%	58%	4.46	.752	1	4.41	4.51
Identifying and prioritizing intended *users* of the evaluation	995	<1%	3%	11%	38%	48%	4.30	.813	5	4.25	4.35
Identifying and prioritizing intended *uses* of the evaluation	991	<1%	1%	9%	41%	47%	4.34	.739	2	4.29	4.39
Soliciting and using suggestions from stakeholders on the evaluation design	1,001	<1%	5%	18%	40%	36%	4.05	.899	8	3.99	4.11
Involving stakeholders in the evaluation process	1,004	<1%	3%	11%	38%	48%	4.30	.813	4	4.25	4.35
Establishing a balance of power among stakeholders	897	6%	15%	35%	31%	13%	3.31	1.06	15	3.24	3.38
Designing the evaluation within resource limitations	978	2%	6%	17%	35%	40%	4.04	.996	9	3.98	4.10
Demonstrating that the benefits of the evaluation will outweigh the costs	954	4%	11%	29%	35%	22%	3.60	1.05	14	3.53	3.67
Adhering to high standards of methodological rigor	1,001	1%	9%	25%	34%	31%	3.85	.998	11	3.79	3.91
Interweaving the evaluation into organizational processes and procedures	985	<1%	2%	15%	42%	40%	4.19	.809	7	4.14	4.24
Linking the evaluation to an established pattern of organizational life	933	1%	6%	22%	42%	30%	3.94	.914	10	3.88	4.00
Developing a communicating and reporting plan	998	<1%	3%	11%	44%	42%	4.25	.785	6	4.20	4.30
Communicating findings to stakeholders as the evaluation progresses	997	<1%	3%	10%	37%	51%	4.34	.804	3	4.29	4.39
Maintaining significant involvement in evaluation follow-up activities	974	1%	9%	25%	41%	24%	3.77	.960	12	3.71	3.83
Taking steps to prevent the misuse of evaluation findings	947	3%	9%	27%	35%	26%	3.72	1.05	13	3.65	3.79

Mean = 4.03

* 1 (not at all influential) to 5 (extremely influential)

power structures," which received an average rating of 2.56 as well as the largest percentage of respondents indicating that they had "no basis to judge/not sure" (10%).

We divided the 10 items into two categories (see category assignment in Table 4): items related to learning (α = .817) and items related to change (α = .813). A factor analysis of the 10 items supported our categorization, although only 6 of the 10 items were included in the two-factor solution. While overall means ratings were high, learning-focused organizational outcomes were rated more favourably (M = 3.84) than change-focused organizational outcomes (M = 3.34) at a statistically significant level; $t(1,018)=2.59, p < .01$.

DISCUSSION

Our study offers a descriptive view of how a sample of US evaluators view evaluation capacity building and the connections between capacity building, evaluation use, and organizational learning and change. An assumption of our study is that not only are organization-related factors important to consider (and to research), but so are the evaluator-related factors that we examine in this study, such as the value of ECB as an evaluation approach, whether they facilitate/plan for evaluation use, and whether they consider organizational factors in their evaluation practice with members of an organization. Our findings lend some provisional support for the emerging ECB models as they have been depicted. Without question, further research is needed on ECB, and we hope that the results of this survey have identified some areas for future empirical work.

Perceptions of ECB as an Evaluation Approach

As evaluation capacity building scholarship develops, new models and frameworks describing ECB are emerging that help to advance our understanding of the evaluator's role in ECB-related activities, the relationship between ECB and evaluation use, and organizational learning and change outcomes. We anticipate that these models and frameworks will help define more precisely what ECB is and what it is not, as well as offer guides for how evaluators can develop and promote ECB in practice.

The numbers are promising concerning evaluators' familiarity with ECB: fewer than 6% of respondents reported being unfamiliar with ECB as an evaluation approach. The remaining 95% of the sample

Table 4
Evaluation Fostering Organizational Learning [L] and Change [C] Outcomes: Learning Items (*N* = 912–1,005)

Evaluation....	*N*	Response Distribution*					Mean	*SD*	Rank	95% C.I.	
		(1)	(2)	(3)	(4)	(5)				Lower	Upper
contributes to an organization's ability to learn from its experiences [L]	1,005	<1%	1%	14%	45%	40%	4.22	.753	1	4.17	4.27
plays a pivotal role in transforming organizations [C]	985	1%	14%	41%	32%	13%	3.42	.910	8	3.36	3.48
enhances individual learning [L]	995	<1%	5%	28%	48%	18%	3.78	.810	6	3.73	3.83
enhances group learning [L]	987	<1%	5%	26%	52%	16%	3.80	.781	5	3.75	3.85
results in fundamental changes in how organizations achieve their goals [C]	983	2%	16%	44%	30%	8%	3.27	.890	9	3.21	3.33
develops users' systematic inquiry skills [L]	970	2%	13%	34%	41%	11%	3.47	.907	7	3.41	3.53
can be a powerful planned change strategy [C]	976	<1%	4%	17%	46%	32%	4.04	.861	2	3.99	4.09
helps intended users question basic assumptions about their practice [L]	997	<1%	4%	22%	51%	23%	3.90	.822	3	3.85	3.95
changes social power structures [C]	912	10%	40%	37%	11%	3%	2.56	.912	10	2.50	2.62
increases users' evaluative thinking [L]	1,000	<1%	5%	23%	51%	20%	3.84	.818	4	3.79	3.89
							Mean=3.84				

*1 (no extent) to 5 (great extent)

felt comfortable rating the importance of this approach, and of those who did, approximately 75% rated ECB as an important/extremely important evaluation approach. Additionally, 82% of the sample considered ECB a responsibility of the evaluator. These data support the assertion made by some (e.g., Preskill & Boyle, 2008) that over the last decade, ECB has emerged as a relevant concept in evaluation, one that is worthy of additional thought, research, and practice.

Our data suggest an emergence of ECB as an important approach to evaluation; however, evidence supporting wide acceptance of ECB as an important focus of the field at this point in time is more open to debate. Respondents considered the importance of ECB to the field as being more of an ideal rather than a reflection of what the field currently values in evaluation approaches. As a relatively new topic, this finding recognizes ECB as an area not as developed as other evaluation practice areas, but identifies ECB as an area with value to the field.

Connections Between ECB and Evaluation Use

Alkin and Christie (2004) argue that ECB theories are derived from and build upon theories of use, specifically, Patton's (2008) earlier notions about process use. Recent refinements of ECB models have focused attention not only on building the capacity of individuals and organizations to evaluate but also their capacity to use evaluation. With ECB as an outgrowth of use theories (e.g., Alkin & Christie, 2004; Christie & Alkin, 2008), connections between evaluation use and ECB seem natural and are now well depicted in models of ECB, such as Cousins et al.'s (2007a) ECB conceptual framework and Preskill and Boyle's (2008) multidisciplinary model of ECB.

In our study, the data do not clearly endorse the connection between evaluation use and ECB. For instance, although there was general agreement amongst evaluators that the factors listed as potentially influencing use were considered to be, on average, at least *somewhat influential* in increasing evaluation use, the two factors most related to organizational processes and life were not endorsed as strongly as the other factors. In addition, respondents who selected ECB as the least important approach were more likely to agree that "identifying and prioritizing intended users of the evaluation" was the responsibility of the evaluator than were those who selected ECB as the most important approach at a statistically significant level. To us, these data suggest a need for further research exploring the relationship between evaluation use and ECB.

Organizational Learning, Change, and ECB

Along with evaluation use, organizational learning and change are integral to ECB. Accordingly, they have been identified as key components of the few existing models for ECB. Included in our survey were 10 items relating to organizational level outcomes that could be fostered through evaluation. These items were grouped into two categories: organizational learning and organizational change. Respondents in our study rated the organizational learning items more favourably than the organizational change items at a statistically significant level. This difference may be explained by a perception that the more typical evaluation processes in which evaluators engage (i.e., engaging stakeholders) are limited in the ways in which they can impact ECB, particularly because ECB is thought to require intentional action (Stockdill et al., 2002).

Respondents' stronger endorsement for the organizational learning items may point to evaluators' beliefs about the limited impact of a more traditional evaluation study—that is, a study where neither ECB nor organizational change is an intended outcome of the evaluation. If we view organizational learning as being closely linked to process use, then evaluation activities that are used during a traditional evaluation are more likely to result in organizational learning (than change). Thus, facilitation of learning is more closely tied to the purview of the more common evaluation design, while transforming organizations more generally is not. Moreover, because ECB is context dependent (Stockdill et al., 2002), it may be more difficult to achieve organizational change through evaluation activities alone, and thus factors beyond the evaluation may be more likely to impact whether organizational change will occur. Facilitating the "transfer of learning" involved in sustained ECB and organizational change is complex and time-consuming, requires greater skill from the evaluator (or at least different skills), and is influenced by several organizational factors, including whether the organization is "ready" for the change, as Cousins, Elliott, et al. (2008) states. This underscores the importance of ECB as being intentional, as argued by Stockdill et al. (2002). The distinction between organizational learning and organizational change, as well as the notion of intentionality, are important and thus should be noted as areas for future research. Perhaps these findings also point to a distinction evaluators recognize between how much or how well evaluation can promote learning within an organization and the actual change of an organization's culture, processes, or procedures. This is consistent with Cousins, Elliott, et al.'s (2008)

distinction between the capacity to *use* evaluation (for organizational change) and the capacity to *do* evaluation—again, another important area for future research.

Study Limitations

An important limitation of the current study is that the majority of items on the survey ask about attitudes or perceptions, not behaviours. Thus, high endorsement of an item does not necessarily link to corresponding behaviours. Additionally, these data reflect the attitudes, opinions, and perceptions of 30% of US members of AEA. This is an admittedly limited sample from which to draw conclusions about evaluators' beliefs more generally. The results presented in this article should be viewed in the context of these limitations.

Future Research

In addition to the areas already noted, we see some additional lines of future research. Future research should unpack the "process" of using evaluation to promote organizational learning and to create organizational change. We know this process involves organizational factors (such as resources, support systems, organizational culture, etc.), but less is known about how the characteristics of an evaluator (attitude, favoured evaluation approach, style of interaction with stakeholders, etc.) relate to these organizational factors already identified in existing conceptual frameworks. Evaluator-related characteristics should be explored in much greater depth so that we can better understand how, why, and under what conditions ECB is possible and achievable. A natural next step is to examine actual evaluation practice (e.g., case study analysis, observation, client-report of evaluation activities) to see if the relationships we (and others) have found between ECB and use and organizational learning and change are articulated in practice.

NOTE

1.   The phrase "Evaluation for Capacity Building" was intended to refer to indirect capacity building characterized by developing evaluation knowledge and skill through doing or being in proximity to an evaluation rather than direct capacity building characterized by developing evaluation knowledge and skills through courses and workshops. Within the context of the evaluation use survey, it was also thought

that respondents would interpret "Evaluation for Capacity Building" as increasing the capacity to do and use evaluation. We cannot be sure respondents interpreted the phrase in this manner, and this ambiguity should be considered a potential limitation.

REFERENCES

Alkin, M. C., & Christie, C. A. (2004). An evaluation theory tree. In M.C. Alkin (Ed.), *Evaluation roots: Tracing theorists' views and influences* (pp. 12–65). Thousand Oaks, CA: Sage.

American Evaluation Association. (2008, April). *American Evaluation Association internal scan report to the membership*, by Goodman Research Group. Retrieved from <http://www.eval.org>www.eval.org.

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

Christie, C. A., & Alkin, M. C. (2008). Evaluation theory tree re-examined. *Studies in Educational Evaluation, 34*(3), 131–135.

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, 93*, 47–62.

Cousins, J. B., Elliott, C., Amo, C., Bourgeois, I., Chouinard, J., Goh, S. C., & Lahey, R. (2008). Organizational capacity to do and use evaluation: Results of a pan-Canadian survey of evaluators. *Canadian Journal of Program Evaluation, 23*(3), 1–000.

Cousins, J. B., Goh, S. C., Clark, S., & Lee, L. E. (2004). Integrating evaluative inquiry into the organizational culture: A review and synthesis of the knowledge base. *Canadian Journal of Program Evaluation, 19*(2), 99–140.

Cousins, J. B., Goh, S. C., & Elliott, C. (2007a, November). *Integrating consequences of evaluation capacity building inquiry*. Paper presented at the annual meeting of the American Evaluation Association, Baltimore, MD.

Cousins, J. B., Goh, S. C., & Elliott, C. (2007b, June). *Organizational capacity to do and use evaluation: Preliminary results of a Pan-Canadian survey of evaluators*. Paper presented at the annual meeting of the Canadian Evaluation Society, Winnipeg, MB.

Cousins, J. B., Goh, S. C., Elliott, C., & Aubry, T. (2008, May). *Government and voluntary sector differences in organizational capacity to do and use evaluation*. Paper presented at the annual meeting of the Canadian Evaluation Society, Quebec.

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.

Fleischer, D., & Christie, C.A. (2009). Evaluation use: A survey of U.S. American Evaluation Association members. *American Journal of Evaluation, 30*(2), 158–175.

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

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, 93*, 81–107.

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

Owen, J. M., & Rogers, P. (1999). *Program evaluation: Forms and approaches*. Thousand Oaks, CA: Sage.

Patton, M. Q. (2008). *Utilization-focused evaluation* (4th ed.). Newbury Park, CA: Sage.

Preskill, H., & Boyle, S. (2008). A multidisciplinary model of evaluation capacity building. *American Journal of Evaluation, 29*(4), 443–459.

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. T. (1999). *Evaluative inquiry for learning in organizations*. 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.

Russ-Eft, D., & Preskill, H. (2001). *Evaluation in organizations: A systematic approach to enhancing learning, performance, and change*. New York, NY: Perseus.

Senge, P. M. (1990). *The fifth discipline: The art & practice of the learning organization*. New York, NY: Doubleday.

Shulha, L. M. (2000). Evaluative inquiry in university-school professional learning partnerships. In V. J. Caracelli & H. Preskill (Eds.), *The expanding scope of evaluation use*. *New Directions for Program Evaluation, 88*, 39–53.

Stockdill, S. H., Baizerman, M., & Compton, D. (2002). Toward a definition of the ECB process: A conversation with the ECB literature. In D. W. Compton, M. Baizerman, & S. H. Stockdill (Eds.), *The art, craft and science of evaluation capacity building. New Directions for Evaluation, 93*, 7–26.

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.

Taut, S. (2007). Studying self-evaluation capacity building in a large international development organization. *American Journal of Evaluation, 28*(1), 45–99.

Dreolin Nesbitt Fleischer is pursuing her doctorate in Evaluation and Applied Research Methods at Claremont Graduate University. She currently works as an independent consultant and resides in Cambridge, Massachusetts, USA. Her research interests are evaluation use, stakeholder involvement, and evaluation practice.

Katherine B. LaVelle is an applied developmental researcher and practicing program evaluator at Five Acres—The Boys' and Girls' Aid Society of Los Angeles in Altadena, California, USA. In 2008 she was a doctoral candidate at Claremont Graduate University. Her work focuses on evaluation of clinical interventions and educational programs, as well as research on the promotion of resilience in children and families.

Christina A. Christie is Associate Professor, Department of Education, Graduate School of Education and Information Studies, University of California, Los Angeles. In 2008 she was on the faculty at Claremont Graduate University.