

## DO EVALUATOR AND PROGRAM PRACTITIONER PERSPECTIVES CONVERGE IN COLLABORATIVE EVALUATION?

J. Bradley Cousins  
University of Ottawa  
Ottawa, Ontario

**Abstract:** Interest in collaborative and participatory forms of evaluation — evaluation that involves evaluators working directly with non-evaluator program practitioners or stakeholders — has increased substantially in recent years. Yet research-based knowledge about such approaches remains limited. Moreover, empirical studies have focused almost exclusively on the perspectives of evaluators or, to a lesser extent, non-evaluator stakeholders associated with the program. This study examines in a direct comparative way the convergence of evaluator and non-evaluator perspectives about collaborative evaluation. Sixty-seven pairs of evaluators and program practitioners, members of which participated on a common collaborative evaluation project, completed a questionnaire about the evaluation and their opinions concerning collaborative evaluation. Relative to their evaluator counterparts, program practitioners indicated they were more involved in technical evaluation activities, were more conservative in their views about evaluation consequences, and tended to feel more positively about the collaborative experience. They agreed, however, about evaluator involvement and the range of stakeholder groups participating in the program. In general, program practitioner and evaluator views and opinions about collaborative evaluation converged, although some differences regarding who should participate and the power and potential of collaborative evaluation were noted. Typically, program practitioners were more conservative in their opinions. The results are discussed in terms of their support for the integration of evaluation into program planning and development.

**Résumé:** L'intérêt pour les formes d'évaluation collaborative et participative (évaluations dans lesquelles les évaluateurs travaillent directement avec des praticiens ou des intervenants non-évaluateurs de programmes) a grandi considérablement ces dernières années. Pourtant, les connaissances basées sur la recherche concernant ces approches demeurent limitées. De plus, les études empiriques se concentrent presque exclusivement

ment sur les perspectives des évaluateurs ou, dans une moindre mesure, sur celles des intervenants non-évaluateurs associés au programme. Cet article compare directement la convergence des perspectives des évaluateurs et des non-évaluateurs sur l'évaluation collaborative. Soixante-sept paires d'évaluateurs et de praticiens qui ont participé à un projet d'évaluation collaborative commun ont rempli un questionnaire sur l'évaluation et ont donné leurs opinions sur ce type d'évaluation. Les résultats montrent que, par rapport aux évaluateurs, les praticiens participent plus aux activités d'évaluation technique, ont des points de vue plus traditionnels sur les conséquences de l'évaluation, et tendent à avoir une réaction plus positive à l'expérience affective d'évaluation collaborative. Les deux catégories s'entendent cependant sur la participation de l'évaluateur et sur la gamme de groupes d'intervenants participant au programme. Dans l'ensemble, les praticiens et les évaluateurs ont des opinions et des points de vue convergents sur l'évaluation collaborative malgré des différences constatées au niveau de qui devrait participer à l'évaluation et du pouvoir et du potentiel de l'évaluation collaborative. En général, les praticiens ont des opinions plus traditionnelles. L'article traite de leur préférence pour l'intégration de l'évaluation dans la planification et l'élaboration de programmes.

Approaches to evaluation and applied social research are increasingly relying on members of the research community (e.g., trained evaluators) working in collaboration with members of the community of practice (e.g., program managers or implementors). Several theoretical rationales for collaborative forms of evaluation have been offered, with justifications ranging from practical problem solving (Cousins & Earl, 1995; Mathison, 1994; Patton, 1997; Preskill, 1994), to considerations associated with the nature of knowledge and implications for its production (Donmoyer, 1991; Greene, 1992; Huberman & Cox, 1990; Mathison, in press), to issues of social justice, equity, and self-determination (Fetterman, 2000; McTaggart, 1991a; Sirotnik, 1990). Recently researchers have begun to study empirically the implementation of collaborative forms of evaluation as way to enhance our understanding of its dynamic complexity and intended and unintended consequences (Cousins & Earl, 1995; Donmoyer, 1991; McTaggart 1991b). Although such studies add greatly to our understanding, they are generally conducted from or focused on a particular perspective or vantage point, that of either the evaluator or the non-evaluator collaborator. Given the relative propensity of evaluators to engage in systematic inquiry, many studies are case reports or reflective essays provided by evalu-

ators (e.g., Greene, 2000; Lee, 1999; Mathie & Greene, 1997; Rowe & Jacobs, 1998). Other studies (e.g., Michalski & Cousins, 2000) have shown remarkable differences in views about evaluation among stakeholder groups. Few studies, however, have examined in a direct comparative way the convergence of evaluator and non-evaluator perspectives about the goals, processes, and consequences of collaborative forms of evaluation. The present study reports data from a modest yet diverse sample of evaluator and program practitioner pairs, members of which collaborated on an evaluation. Program practitioners were not trained in, nor did they have responsibility for, implementing evaluation logic and methods. The objective of the study is to assess the extent to which evaluator and practitioner perspectives converge and to add to our understanding about where and, perhaps more importantly, why they may differ.

## THEORETICAL FRAMEWORK

From an individual learning perspective, both the evaluator and the program practitioner stand to benefit from improved linkages between their respective communities (Huberman & Cox, 1990). Evaluators are likely to develop a deep, rich understanding of practical contexts and the needs of practitioners through sustained interaction with them. Naturally, such enhanced insight will probably influence design, delivery, and dissemination decisions within the evaluation (Greene, 1992). Program practitioners, too, are likely to benefit, as tighter linkages with evaluators are apt to improve their abilities to incorporate and integrate interpretations of evaluation data into their existing personal knowledge structures and frames of reference. These arguments are consistent with increasingly pervasive constructivist theories of learning and may be taken as a *philosophical* rationale for collaborative forms of systematic inquiry (Levin, 1993).

In our prior work (Cousins & Earl, 1992, 1995) we extended this individual learning perspective to a social-psychological framework based on the premise that knowledge is socially constructed and that direct participation in evaluation, inasmuch as collaborative evaluation activities (e.g., data collection, analysis, interpretation) are social acts, will serve as a forum for the integration of evaluation data into collective or shared knowledge representations. In our approach to collaborative evaluation (called practical participatory evaluation) primary users of evaluation data participate directly in the evaluation process from start to finish, including many of the

technical activities such as instrument development, data collection, processing, and interpretation and reporting. We suggest that engagement in such activities engenders deep levels of understanding, by evaluators and program practitioners alike, of the phenomenon under study situated within the local context (Cousins & Earl, 1995). To the extent that this approach fosters organizational norms of systematic inquiry as adaptive (assimilating externally produced knowledge) and generative (producing knowledge locally) learning systems, organizational learning capacity will be enhanced. Other scholars have developed a similar line of argument in reflecting on emerging roles for evaluation (Jenlink, 1994; Mathison, 1994; Owen & Lambert, 1995; Patton, 1999; Preskill, 1994). Collaborative evaluation of this sort is consistent with a utilization-oriented, problem-solving approach and with Levin's (1993) *pragmatic* justification for joint research.

Other forms of evaluation and action research are explicitly ideological in focus and seek to empower stakeholders. Empowerment evaluation (Fetterman, 2000) and participatory action research (McTaggart, 1991a) are exemplars of applied social research that are guided by goals of emancipation, liberation, illumination, and the amelioration of social injustice. In such approaches researchers act as facilitators who place control of the evaluation process with stakeholder groups and function to enable them to assume direct roles in the knowledge production function. These approaches find their justification in versions of critical theory (McTaggart, 1991a; Sirotnik, 1990) and empowerment theory arising from a community psychology framework (Fetterman, 2000). They are consistent with a third rationale for collaborative research identified by Levin (1993), the *political* justification.

Diversity of purpose and goals for collaborative evaluation is paralleled by diversity in its form. In our previous work (Cousins & Earl, 1992) we argued that three dimensions of form differentiate our pragmatic approach to collaborative evaluation from that typical of more conventional forms of stakeholder-based evaluation. These dimensions are control of evaluation project decision-making, stakeholder selection for participation, and depth of stakeholder participation in the evaluation project. In our more recent work (Cousins, Donohue, & Bloom, 1996; Cousins & Whitmore, 1998) we showed that types of collaborative evaluation and applied research could be differentiated according to these dimensions. We also showed empirically that in North America evaluations that tend to be controlled by evalua-

tors, involve a diverse group of stakeholders, and require deeper levels of participation by stakeholders tend to have a greater instrumental (decision support) and conceptual (learning) impact. Such forms of use are consistent with conventional frameworks of evaluation utilization that focus on the use of evaluation findings (Cousins & Leithwood, 1986; King & Pechman, 1984; Patton, 1986). Recent evidence, however, has begun to show that the process of evaluation may have important, possibly overlooked consequences for program stakeholders (Cousins et al., 1996; Cousins & Earl, 1995; Patton, 1997).

Granted that empirical research is developing, and much of this research is guided by or interpreted within principles of the theories mentioned above, much remains to be known. In particular, the interrelationships among characteristics of the evaluation project, the context within which the evaluation is carried out and, ultimately, the short- and long-term consequences of collaborative forms of evaluation findings are not well understood. What are the program-specific and extra-program consequences of collaborative approaches to evaluation? To what extent does the locus of control over the evaluation project (e.g., researcher dominated, practitioner dominated) explain such consequences? What are the implications of various dimensions of stakeholder participation (e.g., power to act on evaluation findings, representativeness of stakeholder groups, number of stakeholder groups, depth of participation) for evaluation activities and impact? What is the relative impact of collaborative evaluation on the use of evaluation findings as compared with consequences of evaluation processes?

Prior research designed to investigate these issues has relied heavily on interpretations or stories told from a particular perspective, usually that of the trained evaluator, and has relied on qualitative methodologies to provide richly detailed accounts of salient variables and their inter-relational properties (e.g., Barrington, 1999; Greene, 2000; Labrecque, 1999; Lee, 1999; Mathie & Greene, 1997). An interesting example is a study of a four-year collaboration of external evaluators and program coordinators to evaluate, understand, and improve a multi-site family literacy program in Colorado, reported by Torres et al. (2000). Although this report represented a single narrative, it was compiled by the evaluator and program practitioner members of the evaluation team. Tensions and issues associated with trust and credibility among members of the team were documented as factors impinging on evaluation implementation. In a different

approach, Robinson (1998) gathered participant observation data in his role as evaluator in a longitudinal study of an internal participatory evaluation of a Canadian national training program. Simultaneously, a research associate independently collected interview data from non-evaluator team members. Once the final evaluation report was released, Robinson was given access to the interview data. He was astonished to find that non-evaluator program practitioners had quite different perceptions from his own, particularly concerning control of evaluation decision-making and depth of participation. Interview respondents typically believed their involvement to be somewhat more limited than did Robinson, who was of the view that things were on a more or less equal footing.

These studies underscore the assertion that members of the community of practice perceive things differently than do their evaluator collaborators in participatory and collaborative evaluations. However, most such studies address the issue of convergence of views in only an incidental way. Placing a high premium on diversity of perspectives and the legitimacy of perceptions shaped by local context are hallmarks of constructivist approaches to empirical research. Ironically, case reports and reflective essays of collaborative evaluation experiences that are grounded in both evaluator and non-evaluator stakeholder perspectives are few in number. Indeed, it would appear that empirical research of any type on the issue of convergence remains limited in both scope and amount.

The present exploratory study seeks to add to the knowledge base by employing survey methods to examine the extent to which researcher and practitioner perspectives about collaborative evaluation converge. The specific objectives of the study are:

1. To describe differences and similarities between evaluators' and program practitioners' reported experiences of participating in a common collaborative evaluation project
2. To describe differences and similarities between evaluators' and program practitioners' reported opinions about the goals of collaborative evaluation and optimal roles for evaluator and non-evaluator participants

## METHOD

Data for the present study were taken from a larger North American survey study of collaborative evaluation practice (see Cousins,

Donohue, & Bloom, 1995, 1996). Sixty-seven pairings of evaluators and non-evaluator program practitioners who had participated together on a specific evaluation project made up the sample. The observed sample was achieved through a survey of evaluators taken from mailing lists of four professional associations: the American Evaluation Association, the Canadian Evaluation Society, the American Educational Research Association (Division H, School Evaluation and Program Development), and the Association of Educational Research Officers of Ontario. Evaluators participating in the survey were provided with the option of nominating a collaborative evaluation team member whose primary responsibility was *not* applied social research. The larger survey yielded usable data from 560 evaluators, of whom 115 (20.5%) nominated a practitioner counterpart. Of the 115 program practitioners subsequently surveyed, usable returns were received from 67 (58%).

It is very likely that the observed sample reported on evaluation projects that were considered by evaluators and practitioners as being positive experiences in some sense (e.g., process, perceived impact). The selection of the project (one that occurred within the past three years) was the evaluator's decision. Why select a project that was a disaster? The identification of the program practitioner survey candidate was also up to the evaluator. Why nominate someone for whom one knows the experience was negative or otherwise unpleasant? King (1995) identified three groups of collaborators who might be part of a given collaborative evaluation: those who are willing and able, those who are willing but unable (e.g., constrained by role responsibilities, workload demands), and those who are unwilling. It seems likely that respondents to the present survey would be members of the first group.

Evaluators and program practitioners received a common questionnaire that was tailored to their response group (i.e., evaluator, program practitioner) with minor adjustments in wording only. The questionnaire was divided into three sections and asked, using both closed-form and open-ended questions, for information on three specific areas:

1. *A specific evaluation project identified by the evaluator:* In this section of the questionnaire, closed-form questions were asked about participation practices of both evaluator and non-evaluator program practitioners (design and planning; instrument development; data collection, analysis, and in-

terpretation; reporting and follow-up); who participated and to what extent (program developers, managers, sponsors, implementers, intended beneficiaries, members of special-interest groups); and purposes, methods, audiences, and dissemination patterns. Questions also addressed perceptions about evaluation impact both in terms of use of findings and process use. Finally, respondents were queried about their affective experience with the collaborative process (reward, stress, frustration, enjoyment, implementation fluidity, etc.).

2. *General views and opinions about collaborative evaluation:* This section of the instrument gathered self-reported opinions and views about general evaluation issues (e.g., primary purposes, external versus internal evaluation, evaluator roles and functions) and views about non-evaluator program practitioner involvement in *formative* evaluation (e.g., diversity of stakeholder involvement, roles for non-evaluator program practitioners, consequences of participation, evaluator roles in working with non-evaluator participants). We decided to focus on formative evaluation on the basis of our belief that collaborative, participatory, and empowerment approaches are best suited to questions of program and organizational improvement (Cousins & Earl, 1995) and our desire to reduce ambiguity for respondents.
3. *Respondent background information:* Here we asked for minimal demographic and contextual information, such as organizational location, gender, education, organizational use of evaluation, and the like.

To test for differences in perceptions between evaluator and non-evaluator collaborators, we used multivariate analysis of variance with repeated measures and related *t*-tests.

## RESULTS

### Respondent Profiles

About one third of the evaluators were employed in university settings, with almost an equal number representing public-sector organizations, usually governmental. About 21% worked for private consulting firms, with only 6% working freelance. Almost half (47.7%) of the program practitioner respondents worked within government

bureaucracy, with another 13.8% employed in non-governmental, public-sector organizations. Almost 15% worked in private-sector organizations that depended heavily on public funds. Relatively few worked freelance or in other private-sector operations.<sup>1</sup>

About half of the evaluators were female, but almost two thirds of the program practitioners responding were women. This difference was statistically significant,  $t_{(67)} = 2.09, p < .05$ . The majority of the pairings were either all male (37.8%) or male evaluator/female practitioner. There were only six all-female pairings and an equal number of female evaluator/male practitioner collaborators. Evaluators tended to be more highly educated than program practitioners, with 62% holding a doctorate and 28% a Master's degree. The corresponding percentages for practitioners are 15.4% and 50.8%.

Evaluators typically had about 13 years' experience in that role; practitioners were somewhat more experienced in their field of expertise (16 years),  $t_{(63)} = -2.34, p < .05$ . Almost 60% of the program evaluators claimed that evaluation was their primary responsibility, and slightly over 80% of the practitioners indicated that evaluation is common in their organizations. Evaluators revealed that about 80% of the evaluations on which they work have involved program practitioners in carrying out the study. Similarly, practitioners indicated that about 73% of evaluations carried out in their organization involved members of the community of practice.

### Convergence in Perceptions of Experiences of Participating in a Collaborative Evaluation

Evaluators participating in the study were asked to identify a specific collaborative evaluation project on which they worked within the past three years and to nominate for the survey a practice-based colleague who had participated on the same project. In the opening section of the questionnaire, survey participants responded to questions about the nature of evaluation and their perceptions about the impact it had. Several questionnaire items were grouped according to a specific focus (e.g., evaluation tasks, perceived impact). For such question sets, multivariate analysis of variance with repeated measures was carried out in order to test for differences between program practitioner and evaluator samples. Table 1 reports the observed differences. The multivariate test is reported for all question areas; for those tests that showed significant differences, univariate F tests were reported to reveal where the mean differ-

ences lay at the item level. In addition to the multivariate analyses, related *t*-tests were conducted for individual items and scales that were not grouped with others.

Respondents were asked to rate the extent of participation by program practitioners in tasks associated with the evaluation process, ranging from scoping out the study to disseminating its results. Generally, practitioners appeared to participate in roles that are consistent with the conventional stakeholder-based model. They helped to frame and shape the evaluation and to interpret and disseminate results, but played a minor role in the technical aspects of the study. Table 1 shows that program practitioners tended to estimate their involvement in the study as greater than did evaluators. This was particularly the case for the technical data processing and analysis tasks.

Evaluators and program practitioners agreed as to the number of practitioner participants who participated in the study, with the median being 6. They also concurred that practitioner participants belonged to more than one stakeholder group and, as shown in Table 1, a multivariate test comparing perceived involvement from six different categories of stakeholders — ranging from program developers to program beneficiaries and special-interest groups — yielded no statistically significant differences.

The groups agreed about who controlled evaluation decision-making, with almost three quarters of the sample suggesting control was balanced; the remaining cases viewed evaluator control as greater than program practitioner control slightly more often. A series of 10 items solicited respondents' views about the evaluator's participation in the evaluation. Table 1 shows that no differences between program practitioner and evaluator views were evident. Generally, evaluators were thought to be most highly involved in providing technical support and expertise and less involved in chairing project meetings, disseminating results, helping practitioners develop evaluation skills, and educating them about the power and value of evaluation as a planned change strategy.

A variety of questionnaire items addressed the perceived impact of the collaborative evaluation project. Although the respondent views converged about the perceived extent of impact, as being modest with more impact to be expected, their views diverged about intended use by intended users. Evaluators tended to provide higher estimates

**Table 1**  
**Repeated Measures Multivariate Analysis of Variance Results for Perceptions about Collaborative Evaluation Experiences**

ITEM CATEGORY / Differentiating Items	Evaluator		Program Practitioner		F	df	Prob.
	Mean	SD	Mean	SD			
<b>PROGRAM PRACTITIONER PARTICIPATION</b> (10 items, 5 pt. frequency scale)							
Designing the study	3.50	1.13	3.87	1.07	3.43	1, 52	= .07
Processing for data analysis	2.15	1.23	2.81	1.43	10.74	1, 52	< .01
Data analysis	2.43	1.21	2.98	1.41	6.70	1, 52	< .01
<b>STAKEHOLDER GROUPS INVOLVED</b> (6 items, 5 pt. frequency scale)					NS		
<b>EVALUATOR PARTICIPATION</b> (10 items, 5 pt. frequency scale)					NS		
<b>EVALUATION CONSEQUENCES</b> (7 items, 5 pt. frequency scale)							
Basis for significant decisions	3.79	0.84	3.56	0.82	3.45	1, 37	= .07
Development of research skills	3.00	0.90	2.55	1.03	5.47	1, 37	< .05
Basic assumptions questioned	3.50	0.95	3.90	0.89	4.46	1, 37	< .05
<b>AFFECTIVE EXPERIENCE</b> (9 items, 5 pt. semantic differential)							
Stressful / not stressful	2.61	1.18	3.24	1.24	10.21	1, 61	< .01
Unmanageable / manageable	3.89	0.96	4.29	0.69	10.09	1, 61	< .01
Frustrating / encouraging	3.66	1.17	4.06	0.98	4.10	1, 61	< .05
Rough / smooth	3.32	0.95	3.77	0.84	8.83	1, 61	< .01

of intended uses than their program practitioner counterparts,  $t_{(60)} = 2.22, p < .05$ . On average, they believed that intended users were “taking steps toward action based on the data,” whereas practitioners were of the opinion that the “information is still being considered.” Ratings of seven different evaluation consequences (use of findings and impacts of the evaluation process) were analyzed and revealed an interesting pattern (see Table 1). Evaluators tended to estimate the instrumental consequences of the evaluation as higher than did program practitioners. They also estimated as higher the role that the evaluation process played in developing program practitioners’ evaluation skills. However, interestingly, program practitioners were more inclined than evaluators to suggest that working on the evaluation had caused them to question basic assumptions and beliefs about practice.

In rating nine semantic differential items concerning their affective experience arising from the collaborative project, practitioners appeared to find involvement much more satisfying and enjoyable than did evaluators. Table 1 indicates that program practitioners found the process to be less stressful, less frustrating, and more manageable and smooth than did their evaluator counterparts. On average, both groups provided positive self-reported ratings of their affective experience with the collaborative evaluation.

### Convergence in Views and Opinions about Evaluation

The next section of the questionnaire dealt with a variety of issues associated with evaluation in general, and particularly with formative, improvement-oriented evaluation. A four-point Likert-type scale was used. Opinions about evaluation in general — including evaluator roles, purposes, and conduct — converged, as shown in Table 2. Generally, these opinions were favourable. Taken together, respondents from both groups tended to disagree that the evaluator’s role is to bring about social justice.

A list of 17 items was associated with views about program practitioner participation in formative, improvement-oriented evaluation. In our prior work (Cousins et al., 1995), factor analyses of these items helped to categorize them into five different categories, and subsequently constructed linear combinations were found to be internally consistent. The categories are views about program practitioner participation in evaluation, formative evaluation’s role in fostering professional development, impact, issues associated with

**Table 2**  
**Repeated Measures Multivariate Analysis of Variance Results for Views and Opinions about Collaborative Evaluation**

ITEM CATEGORY / Differentiating Items	Evaluator		Program Practitioner		F	df	Prob.
	Mean	SD	Mean	SD			
<b>OPINIONS ABOUT EVALUATION IN GENERAL</b> (10 items, 4 pt. agreement scale)	—	—	—	—	—	—	—
<b>VIEWS ON PARTICIPATION</b> (7 items, 4 pt. agreement scale)	—	—	—	—	2.14	7, 41	= .06
Involve primary users	3.27	0.57	3.04	0.62	3.43	1, 47	= .07
Involve special interest groups	2.70	0.61	2.40	0.71	4.97	1, 47	< .05
<b>VIEWS ON PROFESSIONAL DEVELOPMENT</b> (5 items, 4 pt. agreement scale)	—	—	—	—	2.74	5, 54	< .05
Question fundamental beliefs	3.66	0.48	3.50	0.54	3.47	1, 58	= .07
Improve practice	3.78	0.41	3.51	0.57	8.49	1, 58	< .01
Educate re: power of evaluation	3.54	0.54	3.31	0.53	6.28	1, 58	< .01
<b>VIEWS ON IMPACT</b> (3 items, 4 pt. agreement scale)	—	—	—	—	NS	—	—
<b>VIEWS ON TECHNICAL QUALITY</b> (3 items, 4 pt. agreement scale)	—	—	—	—	NS	—	—
<b>VIEWS ON EMPOWERMENT</b> (5 items, 4 pt. agreement scale)	—	—	—	—	2.37	5, 45	< .05
Educate re: power of evaluation	3.54	0.54	3.34	0.52	3.50	1, 49	= .07
Involve special interest groups	2.78	0.62	2.44	0.76	6.26	1, 49	< .05

the technical quality of the evaluation, and formative evaluation's role in fostering empowerment among participants. Views on impact and technical quality converged over evaluator and practitioner groups, but differences were noted for each of the other categories (see Table 2).

Program practitioners were somewhat more skeptical than evaluators about the wisdom of involving primary users (individuals with a vital interest in the program and who are able to act on evaluation data) and stakeholders associated with special-interest groups. They were also less inclined to perceive that evaluation has the potential to enable practitioners to question fundamental beliefs about their practice or to improve their practice. Evaluators were more likely to perceive the evaluator's role as important in educating program practitioners about the power and value of evaluation, and their perceptions about the potential of evaluation to empower stakeholders were more optimistic.

In sum, the survey data reported above yield an interesting pattern, with four distinct themes emerging. First, relative to their program practitioner counterparts, evaluators tended to downplay the actual role that practitioners played in implementing technical evaluation tasks. Second, they also tended to overestimate the impact of the evaluation. Interestingly, they underestimated the extent to which participation in the process caused practitioners to reflect on their own practice. Third, the collaborative evaluation affective experience appeared to be somewhat more positive for practitioners than for evaluators, although both groups self-reported favourable experiences. Comparatively speaking, evaluators were more likely to perceive the process as stressful, difficult to manage, and frustrating. Finally, attitudes about evaluation purposes, processes, and roles were more favourable for evaluators than for the program practitioner respondent group.

## DISCUSSION

The present exploratory study is the first to address in a direct way the convergence of perceptions about collaborative evaluation of members of the evaluation community and members of the community of program practice. A unique feature of the design is that evaluator and program practitioner pairs who were members of the same collaborative evaluation team independently reflected on the same evaluation experiences. The findings contribute to knowledge about

where and to what extent evaluators and non-evaluator program practitioners perceive collaborative evaluation processes and consequences similarly. Do these perceptions converge? Findings from the present survey data suggest that in many instances they do. Program practitioners tended to agree with evaluators in their perceptions about stakeholder group involvement and evaluator participation in a specific collaborative evaluation project, and the groups held similar impressions regarding the potential impact and technical quality of collaborative formative evaluations. However, many interesting differences in evaluator and practitioner perspectives emerged and are worthy of further consideration.

Evaluators tended to emphasize the impact of the evaluation defined in conventional instrumental and conceptual terms. They were favourable about the support evaluation provides for decision-making and its role in fostering concrete action. They tended to underemphasize, relative to their program practitioner counterparts, the role that the collaborative process plays and its impact, quite apart from the findings that emerge. In comparison to evaluators, program practitioners spoke more favourably of collaborative evaluation as a rewarding and enjoyable set of activities, one that they perceived to be an useful professional development experience. They also perceived more than did evaluators that participation caused non-evaluator stakeholders to rethink their practice and to question basic assumptions about it. These data extend a growing body of literature in the evaluation utilization domain that highlights the benefits of the evaluation process as a distinct source of impact on programs and, in particular, stakeholders associated with them (Cousins & Earl, 1995; Jenlink, 1994; Owen & Lambert, 1995; Patton, 1997; Preskill, 1994). The present study supports the investigation of process effects of evaluation as a pressing research agenda. It is interesting to note, for example, that evaluators in our sample tended to perceive process-oriented consequences of collaboration as being less prominent than did their program practitioner counterparts. Why might this be the case? Further, the present results support the integration of evaluation into program planning and development, as opposed to being viewed as a separate, externalized activity (Patton, 1997, 1999).

Emerging evidence suggests that the logic of evaluation and systematic inquiry can be integrated into program, organizational, and community culture, but that such eventualities are most likely to occur within sustained evaluation activities that involve evaluators

working in collaboration with non-evaluator stakeholder participants (Barrington, 1999; Lee, 1999; Robinson, 1998; Rowe & Jacobs, 1998). According to Lee, for example, in her analysis of a sustained school improvement initiative involving significant collaborative evaluation activities, “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” (1999, p. 174). Torres et al. observed program-level consequences of collaborative evaluation in an extended multi-site family literacy program. The evaluation gave “participants greater understanding, confidence, and conviction about the objects of discussion, which in turn increases the likelihood that informed change will occur. It empowers participants to enact change on the basis of program issues and evaluation findings that have been thoroughly considered” (2000, pp. 36–37).

Such cultural change is worthy of continued investigation. We need to know much more about the conditions under which it is likely to occur, the barriers or factors that intrude on such change, and strategies that are effective for overcoming such barriers. This being said, increased sensitivity on the part of evaluators to the perceptions held by program practitioners and other non-evaluator stakeholder groups will help them to accomplish the goal of integrating evaluation into existing program, organization, or community cultures.

Although the present study adds to understanding about the convergence of perceptions held by members of the evaluation community and those of the community of practice, several limitations are worth noting. As observed above, the achieved sample is likely to be biased towards collaborative evaluation projects that are in some sense successful or productive. It is equally important to understand the tensions and less favourable differences that emerge between evaluators and non-evaluator practitioners in evaluation contexts. Such bias will be difficult to overcome in a large-scale survey approach, but one can imagine longitudinal case studies that carefully examine the convergence of stakeholder and evaluator perceptions.

Another limitation of the present study is that the views of only one non-evaluator stakeholder participant were obtained. Yet in most collaborative evaluations members of multiple and different stakeholder groups participate in the evaluation. The stakeholders in the present case, given their program practitioner status, might be considered primary stakeholders, or those with a vested interest

in the evaluation and its findings. But many collaborative evaluations deliberately seek to engage members of the intended program beneficiary stakeholder group, people who typically differ considerably in values and perceptions from their program implementer/developer counterparts (Brandon, 1998; Gaventa, Creed, & Morrissey, 1998; Greene, 2000). Other studies have shown quite significant variation in perceptions about evaluation purposes, processes, and consequences among different stakeholder groups (Michalski & Cousins, 2000). Ongoing empirical research in this domain would do well to take into account the diversity of values and perspectives held among non-evaluator stakeholder groups.

Despite these limitations, the present study does represent a serious empirical investigation of variation in perceptions, and therefore contributes to bridging the gap between theory and practice in collaborative, participatory, and empowerment evaluation. Much more empirical work is needed in order to reduce this gap. Understanding needs to be developed about the conditions and factors that enable effective collaborative evaluation to come about, the barriers that will intrude on realizing this goal, and strategies for ameliorating these barriers. It seems likely that heightened sensitivity of evaluators to the potential process effects of evaluation will be beneficial. This implies that evaluators might attend more directly to the task of developing among non-evaluator stakeholders skills in systematic inquiry and evaluation logic than may presently be the case.

## ACKNOWLEDGEMENTS

Previous versions of this article were presented at the annual meetings of the American Educational Research Association and the American Evaluation Association. The research was funded by the Social Sciences and Humanities Research Council of Canada (grant no. 410-92-0983). The views expressed in this work are the author's and do not necessarily reflect those of the Council. The author is indebted to John Donohue and Gordon Bloom for their contribution to the data collection and analysis phases of the project.

## NOTE

- 1 An anonymous member of the peer review panel for this article expressed an interest in whether differences in perceptions might be attributable to the organizational location of the evaluator or pro-

gram practitioner. Post hoc analyses using multivariate analysis of variance were run to test for such differences between such groups. Within subgroups no differences in perceptions, views, or opinions were found for evaluators (university  $n = 23$ ; public sector  $n = 20$ ; private sector  $n = 24$ ) nor for program practitioners (public sector  $n = 33$ ; private sector  $n = 32$ ).

## REFERENCES

- Barrington, G.V. (1999). Empowerment goes large scale: The Canada Prenatal Nutrition Experience. *Canadian Journal of Program Evaluation, Special Issue*, 179–192.
- Brandon, P. (1998). Stakeholder participation for the purpose of helping ensure evaluation validity: Bridging the gap between collaborative and non-collaborative evaluations. *American Journal of Evaluation*, 19(3), 325–337.
- Cousins, J.B., Donohue, J.J., & Bloom, G.A. (1995, November). *Collaborative evaluation: Survey of practice in North America*. Monograph presented at the joint meeting of the Canadian Evaluation Society and the American Evaluation Association, Vancouver, B.C. (ERIC Document Reproduction Service No. ED 182 465)
- Cousins, J.B., Donohue, J.J., & Bloom, G.A. (1996). Collaborative evaluation in North America: Evaluators' self-reported opinions, practices and consequences. *Evaluation Practice*, 17(3), 207–226.
- Cousins, J.B., & Earl, L.M. (1992). The case for participatory evaluation. *Educational Evaluation and Policy Analysis*, 14(4), 397–418.
- Cousins, J.B., & Earl, L.M. (Eds.). (1995). *Participatory evaluation in education: Studies in evaluation use and organizational learning*. London: Falmer.
- Cousins, J.B., & Leithwood, K.A. (1986). Current empirical research on evaluation utilization. *Review of Educational Research*, 56(3), 331–365.
- Cousins, J.B., & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation* (pp. 5–23). San Francisco: Jossey-Bass.

- Donmoyer, R. (1991). Postpositive evaluation: Give me a for instance. *Educational Administration Quarterly*, 27(3), 265–296.
- Fetterman, D.M. (2000). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Gaventa, J., Creed, V., & Morrissey, J. (1998). Scaling up: Participatory monitoring and evaluation of a federal empowerment program. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation: New directions in evaluation* (pp. 81–94). San Francisco: Jossey-Bass.
- Greene, J. (2000). Challenges in practicing deliberative democratic evaluation. In K.E. Ryan & L. DeStefano (Eds.), *Evaluation as a democratic process: Promoting inclusion, dialogue and deliberation* (pp. 13–26). San Francisco: Jossey-Bass.
- Greene, J.C. (1992). The practitioner's perspective. *Curriculum Inquiry*, 22(1), 39–45.
- Huberman, M., & Cox, P. (1990). Evaluation utilization: Building links between action and reflection. *Studies in Educational Evaluation*, 16, 157–179.
- Jenlink, P.M. (1994). Using evaluation to understand the learning architecture of an organization. *Evaluation and Program Planning*, 17(3), 315–325.
- King, J.A. (1995). Involving practitioners in evaluation studies: How viable is collaborative evaluation in schools? In J.B. Cousins & L.M. Earl (Eds.), *Participatory evaluation in education: Studies in evaluation use and organizational learning* (pp. 86–102). London: Falmer.
- King, J.A., & Pechman, E.M. (1984). Pinning the wave to the shore: Conceptualizing evaluation use in school systems. *Educational Evaluation and Policy Analysis*, 6 (3), 241–253.
- Labrecque, M. (1999). Development and validation of a needs assessment model using stakeholders involved in a university program. *Canadian Journal of Program Evaluation*, 14(1), 85–102.
- 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.

- Levin, B. (1993). Collaborative research in and with organizations. *Qualitative Studies in Education*, 6(4), 331–340.
- Mathie, A., & Greene, J.C. (1997). Stakeholder participation in evaluation: How important is diversity? *Evaluation and Program Planning*, 20(3), 279–285.
- Mathison, S. (1994). Rethinking the evaluator role: Partnerships between organizations and evaluators. *Evaluation and Program Planning*, 17(3), 299–304.
- Mathison, S. (in press). Evaluation as a democratizing force in schools. *International Journal of Social Education*.
- McTaggart, R. (1991a). Principles for participatory action research. *Adult Education Quarterly*, 41(3), 168–187.
- McTaggart, R. (1991b). When democratic evaluation doesn't seem democratic. *Evaluation Practice*, 12(1), 9–21.
- Michalski, G.M., & Cousins, J.B. (2000). Differences in stakeholder perceptions about training evaluation: A concept mapping/pattern matching investigation. *Evaluation and Program Planning*, 23, 211–230.
- Owen, J.M., & Lambert, F.C. (1995). Roles for evaluators in learning organizations. *Evaluation*, 1(2), 237–250.
- Patton, M.Q. (1986). *Utilization-focused evaluation* (2nd ed.). Newbury Park, CA: Sage.
- Patton, M.Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Newbury Park, CA: Sage.
- Patton, M.Q. (1999). Organizational development and evaluation. *Canadian Journal of Program Evaluation, Special Issue*, 93–114.
- Preskill, H. (1994). Evaluation's role in enhancing organizational learning. *Evaluation and Program Planning*, 17(3), 291–297.
- Robinson, T. (1998). *An investigation of the organizational effects of internal participatory evaluation*. Unpublished doctoral dissertation, University of Ottawa, Ottawa.

- Rowe, W.E., & Jacobs, N.F. (1998). Principles and practices of organizationally integrated evaluation. *Canadian Journal of Program Evaluation*, 13(1), 115–138.
- Sirotnik, K. (Ed.). (1990). *Evaluation and social justice: Issues in public education*. San Francisco: Jossey-Bass.
- Torres, R.T., Stone, S.P., Butkus, D.L., Hook, B.B., Casey, J., & Arens, S.A. (2000). Dialogue and reflection in a collaborative evaluation: Stakeholder and evaluator voices. In K.E. Ryan & L. DeStefano (Eds.), *Evaluation as a democratic process: Promoting inclusion, dialogue and deliberation* (pp. 27–38). San Francisco: Jossey-Bass.