

American Journal of Evaluation

<http://aje.sagepub.com>

Multiple Perspectives on Training Evaluation: Probing Stakeholder Perceptions in a Global Network Development Firm

Greg V. Michalski and J. Bradley Cousins
American Journal of Evaluation 2001; 22; 37
DOI: 10.1177/109821400102200105

The online version of this article can be found at:
<http://aje.sagepub.com/cgi/content/abstract/22/1/37>

Published by:

 SAGE Publications

<http://www.sagepublications.com>

On behalf of:

American Evaluation Association

Additional services and information for *American Journal of Evaluation* can be found at:

Email Alerts: <http://aje.sagepub.com/cgi/alerts>

Subscriptions: <http://aje.sagepub.com/subscriptions>

Reprints: <http://www.sagepub.com/journalsReprints.nav>

Permissions: <http://www.sagepub.com/journalsPermissions.nav>

Citations (this article cites 11 articles hosted on the SAGE Journals Online and HighWire Press platforms):
<http://aje.sagepub.com/cgi/content/refs/22/1/37>

Multiple Perspectives on Training Evaluation: Probing Stakeholder Perceptions in a Global Network Development Firm

GREG V. MICHALSKI AND J. BRADLEY COUSINS

ABSTRACT

Although stakeholder-based evaluation has been fairly well developed in the general program evaluation literature, it remains barely recognized in training evaluation practice. This article aims to contribute to our understanding of multiple stakeholder perceptions about training evaluation in an organizational context. Extending prior empirical work that described differences in stakeholder perceptions about valued training outcomes, the current case study examines multiple stakeholder perspectives of the purposes, processes, and consequences of evaluation in a global telecommunications network development firm. Semistructured interviews with 15 individuals in three stakeholder groups were audio recorded, transcribed, and coded for analysis. Diverging from common training evaluation practice, which generally fails to acknowledge and frequently limits stakeholder involvement, the current results provide evidence for a multiplicity of stakeholder views. These views were found to be related to the job role of a stakeholder in the organization. For example, training sponsors (line managers with budgetary discretion) described largely formative evaluation purposes, as well as instrumental and symbolic forms of utilization, to improve training for their employees. Training participants (employees with direct or indirect reporting relationships with the managers interviewed) described mixed (formative-summative) purposes, as well as instrumental and conceptual forms of utilization, to improve specific courses and to make informed course selections. Training providers (internal training specialists who develop and deploy training for the previous two groups as internal organizational clients) described evaluation in terms of mixed purposes, and instrumental and symbolic forms of utilization, mostly to highlight training merit and worth and to sustain and expand training budgets. These differing perspectives are discussed in terms of a three-dimensional schema of collaborative inquiry and evaluation. Implications for training and program evaluation research and practice are also discussed.

Greg V. Michalski • ACT, Inc., Training Evaluation Services (TES), Professional Development Division, 2255 North Dubuque Road, P.O. Box 168, Iowa City, IA 52243-0168; Tel.: (319) 341-2425; E-mail: michalsg@act.org.

American Journal of Evaluation, Vol. 22, No. 1, 2001, pp. 37–53. All rights of reproduction in any form reserved.
ISSN: 1098-2140 Copyright © 2001 by American Evaluation Association.

INTRODUCTION

The training evaluation and program evaluation literatures have developed largely in parallel, with few points of intersection. The general program evaluation literature includes a steadily developing discussion of stakeholder perspectives in terms of the conduct, use, and impact of evaluation (Alkin, Hofstetter, & Ai, 1998; Bryk, 1983; Greene, 1988; Mark & Shotland, 1985). However, despite this work and its emergent application in training evaluation research (Brown, 1994; Darrah, 1995; Michalski & Cousins, 2000), stakeholder-based evaluation, for the most part, remains outside the mainstream of training evaluation practice. Given the large and growing body of training literature and the amount of resources devoted to corporate training and its evaluation (Bassi & Van Buren, 1998), we find the underdevelopment of work on stakeholder diversity in the field to be striking. Particularly in an age when the effectiveness and competitiveness of business organizations are increasingly described in terms of their collective intellectual capital (Stewart, 1999), we see opportunities to improve training evaluation practice through a better understanding of multiple stakeholder perspectives.

The present study has two primary purposes. The first is to contribute to the underdeveloped empirical literature on stakeholder diversity in training evaluation. By definition, stakeholder-based evaluation is politically complex, as it seeks to use the perspectives of multiple or varied constituencies. Organizations are also complex and political. The paucity of empirical literature on stakeholder issues in training evaluation is even more problematic because general discussions of politics in training evaluation are even rarer (virtually nonexistent) in the literature. The lack of discussion in both of these areas has limited the scope and robustness of training evaluation practice, as many evaluators continue to evaluate training from a unidimensional perspective in terms of, for example, return on investment (Lombardo, 1989; Phillips, 1996; Tesoro, 1998). We suggest instead that, as in many other areas, different stakeholder groups may hold different views about training that evaluators should take into account.

A second purpose is to extend and more deeply probe the findings of previous related research. In that research, we employed quantitative methods of concept mapping and pattern matching techniques (Trochim, 1989) to investigate stakeholder differences in valued training outcomes relative to evaluation (Michalski & Cousins, 2000). We examined the views of three stakeholder groups: training sponsors, training participants, and training providers. All stakeholders identified a similar set of valued training outcomes for the organization. However, greater differences were found regarding the relative importance that each group ascribed to these outcomes as criteria for training evaluation. The patterns of variability that were revealed correspond well with the specialized organization role of each stakeholder group, particularly in terms of multiple-constituency views of organizational effectiveness, power theory, and discussions of the politics of evaluation. In this paper, we attempt to extend the previous research and offer a richer view of multiple stakeholders' perspectives on training evaluation.

CONCEPTUAL FRAMEWORK

In this section, we briefly review the main elements of the conceptual framework employed in the present research. This framework was similarly used to guide our prior work, and a

more detailed view, as well as a review of literature associated with the framework, is presented in Michalski and Cousins (2000). We begin by considering the complex organizational context, which includes multiple-constituency views of organizational effectiveness specifically from the perspective of knowledge workers. Three stakeholder groups—each composed of highly skilled, knowledge workers possessing substantial intellectual capital—were defined in terms of their roles in the organization with respect to training. In our case study organization, one stakeholder group was the training sponsors, which included line managers with direct responsibility for product development and budgetary discretion for training. The second stakeholder group, training participants (trainees), included various nonmanagement, product development personnel, such as engineers, hardware and software designers, and other technical workers with direct or indirect reporting relationships to the sponsor group. Both the sponsor and training participant groups are members of the same product development organization. The final stakeholder group of interest were training providers, defined as the training professionals who designed, developed, deployed, specified, and/or facilitated training for the previous two client groups.

The results of training were described by stakeholder perceptions of valued outcomes that would contribute to the success of the organization. As noted in our prior work, we found general agreement among stakeholder groups regarding these outcomes. However, there were substantial group differences regarding the relative importance ascribed to these in terms of training evaluation.

In addition to describing what training outcomes they value, interviewees were asked for their views about the purposes, processes, and consequences of training evaluation. Our classification of evaluation purposes for this research include formative (related to program improvement), summative (related to the judgment of program merit and worth), and mixed (formative-summative) forms. Evaluation purposes might also include a broader range of possibilities, including knowledge generation and organizational learning (Patton, 1996, 1997; Preskill & Torres, 2000). Evaluation processes that we asked about include the planning, collection and analysis of evaluation data, and the presentation of evaluation findings. The consequences of evaluation include primarily (1) instrumental uses related to specific decisions affecting training, (2) conceptual uses for which the evaluation serves an educative function, and (3) symbolic uses involving the use of evaluation to signify or connote a set of results (Caracelli & Preskill, 2000; Shulha & Cousins, 1997).

CASE ORGANIZATION

This was a single case study. A case study is not a methodological choice, but a choice of object to be studied. Its central purpose is to optimize understanding of the case rather than generalize beyond. A case study involves both the process of learning about the case and the product of our learning (Stake, 1994, 1995).

The case organization we examined is a major research and development (R&D) division of a global network design and engineering company based in Ontario, Canada. The company as a whole has a total workforce in excess of 50,000 employees, and operates in over 150 countries and territories worldwide. With a substantial investment in the collective intellectual capital of its employees, the company strives to maintain and contribute to their knowledge and skills through a range of training, education, and development opportunities, including computer and web-based (Michalski, 2000). Figure 1 offers a diagram of the main

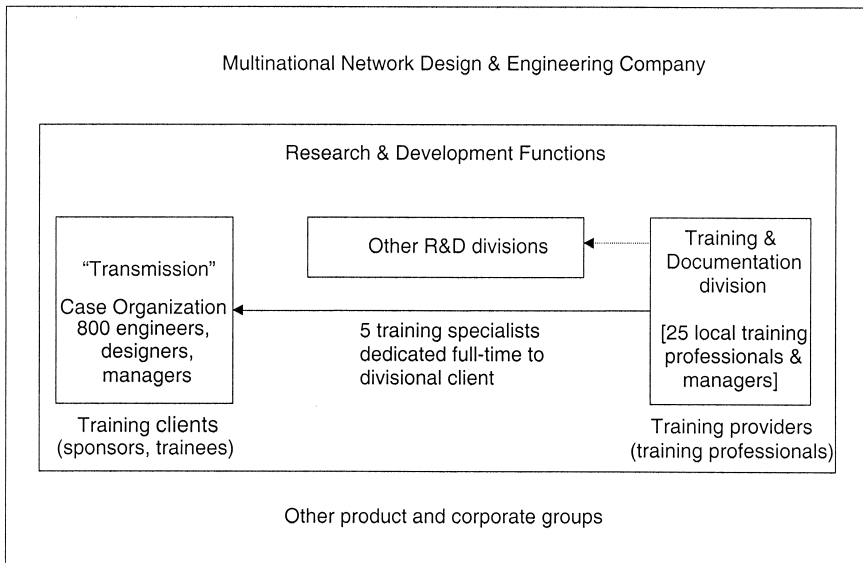


Figure 1. Situation of case organization within the company.

functional components within the case organization at the time of the study. The transmission group is situated among several divisions that provide various R&D functions for the company. The main activities of this division—known simply as “transmission”—revolve around the design and development of technologies and products for market to the telecommunications industry, specifically, high-speed, synchronous optical network (SONET) hardware and software (see Winch, 1998). Along with other parallel development groups, transmission receives training support from a separate internal training and documentation staff. During the study, the training staff consisted of approximately 25 locally situated training professionals and training managers, while the transmission organization was staffed by approximately 800 local, full-time personnel.

METHODS

Qualitative interviewing provided the principal means for data collection and analysis. In addition to its well-established use for field research in the social sciences, this method also has an emerging history of application for learning-related research in technology-based organizations (see e.g., Brown & Duguid, 1991; Sacks, 1994).

Sample

Of the 15 study participants who constructed concept maps in our prior research (Michalski & Cousins, 2000), eight were randomly selected from each of the three groups and invited to participate in interviews. Five individuals from each of the three stakeholder groups agreed to be interviewed. Individual, private interview sessions were scheduled. Before each interview, permission was specifically requested and obtained to record the

session on audiotape. Realizing that actual session times would vary (depending on the length, depth, and complexity of any given respondent's reply), each interview was scheduled for one hour in duration. All interviews were conducted on the premises of the case organization by the senior author who, during the period of the study, was employed as a training and performance project manager with the company.

Interview participants were familiar with the general background and purpose of the interviews from their participation in our prior research. All three groups had comparable demographic representation based on age, gender, years with the company, years in current job, and job classification level. Only one individual indicated his age to be less than 30 years; seven of the individuals indicated their age in the 30 to 39 year range; four in the 40 to 49 year range; and three in the 50 to 60 year range. Thirteen of the study participants were male, and all but one participant indicated being with the firm at least three years. Five participants indicated they had more than ten years with the company.

Instrument

A semistructured interview guide was developed in view of the conceptual framework described. Pilot testing of the guide was done with six individuals, two from each of the three stakeholder groups, who did not participate in the study. The pilot-testing process involved using the guide in trial interview sessions. Results of these sessions were combined with information obtained from participant debriefing about the quality and clarity of the interview questions. Notes were taken, reviewed, and used to improve the phrasing and content of the interview guide. The guide was refined and updated using the pilot-test results.

The final interview guide included questions soliciting respondent views about the purposes, processes, and consequences of training evaluation. Several questions also allowed respondents to describe and comment on perceived differences among all stakeholder groups. Additionally, although largely outside of the main focus of the present discussion, several questions were also included to corroborate each individual's views about valued training outcomes with those of their respective stakeholder group based on previous concept mapping results (Michalski & Cousins, 2000). As discussed in detail in our prior work, a concept map is both an individual and group construction developed using multidimensional scaling and cluster analysis techniques. An integrated software application was used to organize, weight, and graphically present conceptual categories as a concept map. Pattern matching comparisons also allowed us to measure the degree of similarity/dissimilarity in views (e.g., on the importance of a given training outcome as a criteria for evaluation) between various combinations of groups and individuals (see Michalski & Cousins, 2000), and a secondary purpose of the interviews was to assess our participants' agreement with the concept map developed for their stakeholder group.

Analysis

All interviews were audiotaped and transcribed verbatim to enhance descriptive validity (see, e.g., Maxwell, 1996). These data were analyzed following techniques described by Miles and Huberman (1994). These authors favor deriving a start list of codes from the conceptual framework and research questions guiding the study before fieldwork. A list of "start codes" was developed corresponding to the main elements of the conceptual framework described. The coding list underwent several revisions both before and during field-

work. After preliminary analysis, several “add on” and “pattern” codes were further developed. To expedite the analysis, we used a popular qualitative data analysis software program to perform the coding and indexing, while accommodating our conceptual framework (Richards & Richards, 1994).

RESULTS

The interview sessions ranged between 30 and 90 minutes, with most requiring approximately an hour. The shortest interview occurred with a relatively new designer who had less than five years of experience in his profession. This individual was not very opinionated about training because, as he noted, this was not his area of expertise. The longest interview occurred with the very opinionated director of the training provider group, who had nearly 30 years experience with the company. Although the following discussion focuses primarily on stakeholder views regarding training evaluation, we first briefly present individuals' views of training outcomes relative to their group's concept mapping results.

Valued Training Outcomes

Corroborating the results of our prior work (Michalski & Cousins, 2000), all individuals generally agreed with the training outcomes depicted on their group's concept map. Although there were differences across the groups in terms of the number and labeling of conceptual clusters, all participants identified valued training outcomes in terms of customer and employee satisfaction, as well as employee learning relevant to the job and/or profession.

All groups described a relatively general set of valued outcomes; however, each group emphasized different aspects in terms of training evaluation. These differences corresponded quite closely with a stakeholder's organizational role and relationship to the training function. Trainees emphasized their own professional and career development. Sponsors exhibited a somewhat more balanced view, describing the interdependencies across training outcomes, especially in terms of business, customer, and employee performance. Training providers also recognized external customers, but were most keenly concerned with training results in terms of the satisfaction of sponsors and participants as internal training clients. Stakeholder views of the purposes, processes, and consequences of training evaluation are described next.

Stakeholder Views of Training Evaluation

A summary and comparison of stakeholders' perspectives about training evaluation is presented in Table 1, which is derived from the results of the interviews, coding, and analysis performed. Details of the results summarized in the table follow.

Purposes. Study participants were asked for their views of the main purpose(s) of training evaluation in formative and summative terms. Based on this, interviewees discussed evaluation purposes largely in formative, summative, and mixed (formative and summative) terms. Members of the training sponsor group described evaluation using predominantly formative terms, emphasizing course and program improvement. As managers in a highly technical product development environment, program sponsors generally viewed training as a necessary investment (or perhaps sunken cost). Viewing themselves as decision-makers

Table 1.
A Summary and Comparison of Multiple Stakeholder Views About Training Evaluation

<i>Training Evaluation</i>				
<i>Group</i>	<i>Purposes</i>	<i>Processes</i>	<i>Consequences</i>	
Sponsors	Formative (PPP)	Sponsors as primary users (PPP)	Instrumental utilization (PPP)	
	Mixed formative and summative (PP)	Trainees as data sources (PPP)	<ul style="list-style-type: none"> ● improve program 	
		Survey data & analysis (PPP)		
	Report to all stakeholders, brief format (PPP)			
Summative (NN)	Providers as users for program improvement (PP)		Symbolic utilization (PP)	
	Qualitative data, e.g., interviews, focus groups (PP)		<ul style="list-style-type: none"> ● enhance program credibility both above and below in the line organization hierarchy 	
	Report results using web (NN)		<ul style="list-style-type: none"> ● encourage program participation among employees 	
	Long detailed report (NNN)			
Trainees	Mixed formative and summative (PP)	Trainee & sponsor involvement (PPP)	Instrumental utilization (PPP)	
		Survey data & analysis (PPP)	<ul style="list-style-type: none"> ● assist employee course selection ● support professional development 	
	Summative, largely course-level evaluation (P)	Report to all stakeholders, web format (PPP)		
		Provider involvement (PP)		Conceptual (PP)
Report results using e-mail (NN)		<ul style="list-style-type: none"> ● educate providers to improve program 		
Providers	Mixed formative and summative (PPP)	Provider group to initiate and lead evaluation (PPP)	Instrumental utilization (PPP)	
		Sponsors as information sources to frame evaluation (PPP)	<ul style="list-style-type: none"> ● sustain program ● evolve program ● secure program funding ● advertise training 	
	Summative (PP)	Survey data & analysis (PPP)		Symbolic utilization (PP)
		Selective reporting to all stakeholders (PPP)		<ul style="list-style-type: none"> ● highlight program success ● demonstrate program value ● demonstrate training expertise
	Trainees as information sources (PP)			
	Qualitative data (PP)			
Knowledge testing (P)				
Selective reporting using web (P)				
Long detailed report for a single audience (NN)				

Key: PPP = strong positive; PP = moderate positive; P = weak positive; N = weak negative; NN = moderate negative; NNN = strong negative.

who essentially “pay” for training from departmental budgets, sponsors saw a central purpose of evaluation as being to continuously improve training. The following comments from two different members of this group represent this perspective:

“I think that improvement is a much more positive angle. Judgment is kind of saying, ‘well, we experienced this set of training and now we are going to decide whether it is good or bad.’ It doesn’t imply any forward movement in terms of taking that value judgment and doing something with it.”

“The next step is to determine ways to improve what we have . . . I mean, yes we need training, so we need it! So who cares whether it’s good, bad, or indifferent, just make it better and let’s deliver it . . . so I think the evaluating is contained in determining ways to improve it.”

In contrast, members of the trainee group had a somewhat narrower focus compared to sponsors. As direct consumers who participate in a wide range of training interventions, trainees exhibited a more mixed formative-summative view of evaluation purposes. Perhaps as a result of this group’s deeply entrenched image of “training evaluation” in terms of typical end-of-course-satisfaction-forms, which are generally more summative than formative in tone, members of the trainee stakeholder group consistently emphasized a balanced mixture of evaluation purposes. The following excerpts serve to illustrate this view:

“I would definitely say that it is a mix of both [formative and summative]. . . . You can’t really make a suggestion for improvement without criticizing or judging, and you can’t really judge unless you think there is room for improvement.”

“I think it’s both. . . . so maybe you can improve what is there already, and maybe you can say that it is not necessary and introduce something else.”

The third group, training providers, described some formative purposes for training evaluation. However, they also expressed rather strong summative views that were mostly intended to demonstrate the value of training, in financial terms, to training sponsors. The summative aspect of this group’s perspective might be further explained by their expressed interest in favorable evaluation results—namely, those that could be used as evidence of merit and worth to support continued funding and expansion of training. The comments of two different training providers capture this perspective well:

“Both [formative and summative]. . . . [training] has to have a benefit. What is it worth to the organization? That’s a judgment. So we are spending that much money on investing in your people for a reason. You need to know what that reason is. At the same time, how we can make it better? So you use it also to make it better. . . . So judge and improve.”

“[The purpose of training evaluation is] to make the decisions on the investment up front, if you have a limited amount of cash to invest.”

As discussed in a later section, training providers also echoed similar views in describing evaluation consequences—particularly those leading to slow, sustained program change and evolution.

Processes. Each study participant was asked for their views about training evaluation processes in terms of planning and involvement, data collection and analysis, and the dissemination or reporting of results.

Planning and involvement. Multiple stakeholder involvement was generally favored across the groups. Sponsors expressed an interest in diverse involvement and, while members

of this group did mention the involvement of all three stakeholder groups, they emphasized involvement by themselves and training participants. To a large extent, sponsors viewed participants as the main training consumers. As managers with discretion about funding of training, they viewed themselves as the primary evaluation users, acting on behalf of training participants (i.e., their employees). Hence, sponsors consistently mentioned the involvement of both trainees and sponsors:

“The participants, [and] . . . managers [as] the people in the organization who are responsible for moving its capability forward. The people who deliver the training, in order to evolve a training event to make it better next time.”

“The individuals themselves. . . the individual’s manager. . . But the other ones could be catalysts for it—like the people who give the training, [they] have to respond to it in some way [to] improve it.”

While there was also mention of provider involvement, this was almost an afterthought as illustrated by the following sponsor comment:

“The people who take the training. . . then of course, the training people [providers] from a ‘How do I best deliver it?’ point of view.”

This particular sponsor also did not seem to have much faith in the idea of having members of the training provider group be principally responsible for the planning and execution of training evaluation:

“There could be another independent group of individuals who can provide input on what constitutes effective training. People who have enough knowledge of training techniques and training approaches to be able to advise the people who are trying to evolve the training as to what works and what doesn’t.”

Such a view suggests a role for an external training evaluation service or possibly an independent internal evaluation function (Love, 1991).

Trainees generally saw themselves as evaluation data sources, and largely viewed trainees and sponsors as central in the evaluation process. In addition, this group also expressed a range of other opinions. For example, in responding to the question about who should be involved in training evaluation planning, one trainee commented:

“The people being trained [trainees], the people who send people to be trained [managers], and also the people providing the training [providers], because all three are inter-related.”

Another trainee expressed a lack of understanding about the role of management in the evaluation process by stating, “I don’t have much of an idea of how the managers are evaluating the training process.” A different trainee flatly stated that training providers need not be involved in training evaluation. This individual indicated that highly experienced line managers who have, for example, at least ten years of experience would be in the best position to evaluate training.

In contrast to the previous two groups, training providers viewed themselves as central in the planning and execution of evaluation. While widely discussing a need to involve all stakeholders, providers saw evaluation as a means to promote training as their main service, and a potential way to expand their services (e.g., by providing evaluation services) to internal organizational clients. One experienced senior training manager suggested that the evaluation process should be the exclusive domain of the training provider group. This respondent closely linked evaluation to the existing training needs assessment services that

the group offered. In discussing evaluation planning and execution, another training provider stated, "I think that the training specialists have to drive [the evaluation process], absolutely."

On the whole, providers viewed sponsors' perspectives as primarily useful in helping to specify the purpose, scope, and focus of the evaluation. Trainee perspectives were also mentioned, but the emphasis was on sponsors' views. The following training provider comment illustrates this:

"I wouldn't want it to be left to me to try to tell them [sponsors] what is going to be the benefit of this, I would like them to be able to work with me and decide. If [sponsors] can tell me up front [what the goals of training are] then we'll measure [that]. . . and come up with a way that we're both going to believe what we find out. So, number one, the sponsor has to decide with you. The sponsors and obviously the [trainees] are going to be integral in making that decision, but I think the sponsor has to be the one to give you that end result. If [in the sponsor's view, positive results are achieved] then you have done your job."

Other training providers expressed similar views regarding the primary role of sponsors in performing the evaluation.

Data collection and analysis. Respondents were asked their views about data collection and analysis in the training evaluation process. Perhaps because of the pervasive use of traditional questionnaires to collect user satisfaction data related to training (and other services) in the organization, all groups identified the use of such surveys aimed primarily at the trainee population. Sponsors and participants were closely aligned in their views about evaluation data collection and analysis. The following sponsor comment is illustrative:

"I think survey forms are quite effective [although] there is a danger of being surveyed to death. But if it is a form that can be filled in quite quickly, and with skillfully designed questions and multiple choice answers [and] if the words are phrased correctly; and it can very easily key into potential answers that people will want to give, then it is quite an effective way to collect data quickly and efficiently."

There was also general recognition of alternate means of collecting training evaluation data. This included the use of more qualitative methods, such as interviews and focus groups. After reflecting on some of the problems he perceived in using strictly quantitative means (i.e., scaled questionnaire items), a training participant commented:

"Maybe a mixture of a survey and group discussions. . . . You would have a focus group of people that would be able to talk to each other, brainstorm, and send feedback directed to a specific area. I would capture that and then send it back as part of the evaluation."

Similarly, a training sponsor expressed his openness to alternative methods by commenting:

". . . if it involves a large group and/or it's very critical, then probably the focus group type of approach comes into play there, because you are going to get more insight by having a group discussion than you are just asking people independently of each other to fill in a form."

Although there was a general sense among both sponsors and participants that evaluation data collection and analysis was outside their respective areas of expertise, training providers were much more comfortable asserting their opinions on this topic. In addition to

generally mentioning survey techniques, this group favored data that produced what they perceived as “hard evidence” to demonstrate the value and worth of training, especially to funding sponsors. The following comments by a training specialist exemplifies this perspective:

“A questionnaire to a sample [of trainees] will work [but] I think that we’ve got to increase the amount of testing that we do, especially with the volume of people that seem to come into line groups . . . You’ve got to be able to competently say someone can do something . . . and the value of that [is that it] shows the sponsor that we’re addressing the training that they need.”

This perspective also corresponds well to the training provider view (discussed earlier) that providers should “drive” the evaluation process, based on the widely held belief that they are obviously the best positioned of the three groups to implement, facilitate, or otherwise specify appropriate measures (e.g., knowledge and performance tests).

Dissemination of results. Respondents were asked about appropriate methods and audiences for disseminating or reporting training evaluation results. Individuals from all groups generally favored reporting training evaluation results in an open manner to all stakeholders. All sponsors interviewed mentioned themselves as report recipients. In the following comment, a sponsor indicated his preference to openly share evaluation results, while also reinforcing the formative evaluation perspective held among the sponsor group:

“. . . report the results to somebody who can actually take the action to do improvement, and two, we’d also want to report the results to the sponsors of the whole training program. . . . The third audience would be the population at large. They need to know what has changed, what is new, and what is improved.”

Another clear theme among this group was an aversion for large (especially paper-based) training evaluation reports. Alternatives mentioned included live presentations of results and especially web-based summaries (e.g., accessible on the corporate intranet), as illustrated by the following sponsor comment:

“So typically [evaluation results should be reported using] web pages where information about courses is available. . . . I think the last thing that [sponsors] need is a 10 page, closely type-written report going into a lot of detail.”

Trainees were more indifferent about the details of evaluation reporting. This group generally viewed all three stakeholder groups as legitimate recipients of evaluation results, but considered sponsors to be the primary users of the information. In stating that results should be reported “to the people who provide and will decide the future of the training program,” one trainee implied a belief that trainees, individually or collectively, did not generally have much to do with deciding the future of training policy. This group further favored using the internal corporate intranet (web) as a means to disseminate evaluation results. Two individuals also mentioned their dislike for e-mail as a reporting method, while another suggested that, despite the presence of formalized reporting methods, training evaluation information flows among employees through informal interpersonal networks (e.g., word of mouth, personal e-mail messages, and so on). This is especially true among training participants regarding comments about the quality and usefulness of specific courses as a way to assist peers in selecting worthwhile courses.

Training providers purportedly favored wide dissemination of evaluation results. Ac-

ording to one training manager, "I firmly believe that every person who contributed to the study should be given the opportunity to see the end results." But it was also clear that this group viewed themselves as the principal owners of evaluation results. Hence, this group favored retaining a degree of selective control over dissemination and reporting, especially where they perceived this to affect training funding decisions. The following comments of two different training managers illustrated this view:

"[Training evaluation reporting] would depend on what you are reporting and what parts of the data you are reporting. . . it will probably have to be reported many different ways to different audiences and probably at different times."

"I am a [training] provider and my opinions are totally different when it comes to dealing with training from the person who is paying the money."

This group did not support the use of a single, formal evaluation report document to be distributed to all stakeholders. Rather, they suggested a range of "tailored" reporting mechanisms, including e-mail, the selective placement of information on the corporate Intranet, and live presentations of evaluation results to sponsors, particularly to positively influence their training budget decisions.

Consequences

Respondents were asked about their views of the ideal consequences of evaluation in terms of how results could be used to bring about beneficial program change. Responses referred to a variety of instrumental, conceptual, and symbolic evaluation consequences that ranged from rational decision-making to largely political.

Training sponsors described evaluation consequences predominantly in terms of instrumental and symbolic forms of utilization. As a logical extension of their views around formative evaluation purposes (discussed above), instrumental utilization patterns were described in terms of training success and improvement. The following two sponsor views illustrate this perspective:

"The end result [of evaluation would be a] successful training program. . . everybody is able to learn the things they feel are important, and as a result of that we do the right things."

"Show that there is change and adaptation, constant, continuous improvement taking place, monitoring yourselves, and trying to make things better. Just so you realize that the courses are not stagnant. . . it would lead to an improvement."

But in addition to views about consequences in terms of instrumental utilization, sponsors also expressed symbolic utilization views. A form of symbolic utilization is exemplified by comments that describe evaluation as boosting training's credibility with individuals both above and below the line in the organization's hierarchy. This is illustrated by the following excerpts from two sponsors:

"The consequence of it is that people would put some credibility into the training, that if I invest and send my employee to this training he is going to be better for it, you need to build confidence up."

"A consequence of the training program evaluation process would be that [executive management] would feel comfortable that something is in place that would raise a flag when necessary, but not kind of put it in their face all the time when not necessary."

Trainees were mostly concerned with using evaluation for course improvement. As the group with the least accountability for evaluation results, rather than discuss consequences in terms of symbolic utilization, this group described utilization in strongly instrumental and moderately conceptual terms. The following comments from two trainees typify this view:

“We will see an improvement in the quality of training being provided to [hardware and software] designers, and the quality of work coming out of the design groups. I’m focusing specifically on design because that’s where I work, but I think that extends very well into other areas as well.”

“Eventually, I think that the training will be more suited to the needs of workers and to the needs of the [company]. Because the existing courses will be evaluated and necessary actions will be taken, this will satisfy the workers and the people who are taking the courses.”

Training providers were by far the most opinionated about the consequences of evaluation. Members of this group discussed evaluation consequences principally in terms of instrumental and symbolic forms of utilization. This involved continuously improving training through slow, steady evolution, and highlighting positive training results as “advertising” for training in the organization. The following provider comments reflect these perspectives:

“If the evaluation is effective, we’ll find that a curriculum evolves. So [training delivery methods and media] changes, the frequency of delivery changes, the modularity of training changes. So there is always some sort of movement. Things somewhat change [and are] not just constant.”

“In a nutshell, the training will be hopefully revamped; when I say revamped I don’t mean a huge change, but even if it is a minor change and kept updated, if it is done on a yearly basis.”

“[Evaluation results should be used mainly] for recognition, for advertising. If we do a good job evaluating training, then we can make the training better and better. . . . So it will lead to changes in improvement to the course marketing, you can get more people coming in.”

Such views correspond well with this group’s preference for mixed (formative-summative) evaluation purposes, as well as their desire to manage and influence the evaluation process.

DISCUSSION

Each stakeholder group articulated their views of the purposes, processes, and consequences of training evaluation by differentially emphasizing certain valued training outcomes over others. The organizational role and expectations of the group (see, e.g., McLinden & Trochim, 1998) were observed to influence stakeholder views of evaluation. These findings support our previous results in which concept mapping and pattern matching techniques revealed group similarities regarding valued training outcomes, but differences regarding group views of training evaluation.

As an aid to understanding the current findings, the evaluation perspectives of each stakeholder group can be categorized using a three-dimensional schema of collaborative inquiry (Cousins & Whitmore, 1998). The three axes (diametrical extremes) of the schema

are defined in terms of (1) who has control of evaluation process (researcher/evaluator to program practitioner); (2) stakeholder selection for participation (all legitimate groups to primary users only); and (3) depth of participation (consultation to deep participation). Based on the present case study, both training client groups (i.e., sponsors and trainees) favored (1) researcher (evaluator) control, (2) involving all legitimate stakeholder groups, and (3) a consultative depth of participation. These coordinates describe what is typically described as stakeholder-based evaluation.

On the other hand, training provider views would be most accurately described as (1) favoring program practitioner (in this case training provider) control, (2) involving only primary users (providers and sponsors), and (3) a consultative depth of participation. Such a view tends to be consistent with at least some applications of empowerment evaluation (Fetterman, 2000). Both of these forms of evaluation involve considerations of politics and—although beyond the scope of the present study and as noted earlier—the politics of training and its evaluation is an area ripe for further investigation. Beyond this, the current findings have several implications for training and program evaluation. These are discussed next.

Implications for Research and Practice

Recent research suggests that training evaluation is underutilized for decision-making relative to organizational training (Twitchell, Holton III, & Trott, 2000). We believe that, for training evaluation to have fuller importance, assumptions about stakeholder homogeneity in training evaluation practice need to be questioned more thoroughly. Our results suggest the presence of a multiplicity of stakeholder views regarding the purposes, processes, and consequences of training evaluation. This multiplicity appears to be inadequately recognized, largely, we believe, because of the lack of integration of the training evaluation and program evaluation literatures. Well-established program evaluation theory and practice remains largely outside of the mainstream in training evaluation practice. As popular training evaluation practice continues to frame evaluation as a largely singular stakeholder endeavor (usually based on an assumed view of the evaluation sponsor), training may well be generating unrecognized organizational effects beyond this limited view. Training evaluators can address this by familiarizing themselves with the program evaluation literature, and by taking measures to define stakeholder groups in connection with a training evaluation project.

Contemporary program evaluation tells us that there are many intended and unintended consequences of evaluation and modes of utilization. Recent conceptions of evaluation use include those related to process use and organizational learning (Caracelli & Preskill, 2000; Daft & Huber, 1987; Shulha & Cousins, 1997). Such notions represent a significant development because they augment and extend prior conceptualizations of evaluation utilization described by the instrumental, conceptual, and symbolic modes. For example, process use (see also Patton, 1997; Preskill & Caracelli, 1997) refers to the impact of the evaluation process, not just its outcomes or results for short-term decision making. Through their involvement in the evaluation process, participants can learn about the organizational effects of training. Such learning can ultimately affect organizational practices and behaviors in ways that transcend limited views of use for short-term decision making. Program evaluation researchers might fruitfully examine why program evaluation ideas have not influenced training evaluation to a greater degree. Furthermore, skilled program evaluation practitioners

may want to take a fresh look at the sizable area of training evaluation as an applied field offering substantial opportunities.

Limitations

Because our interest was in performing training evaluation case research, as opposed to a generalized training evaluation, there are a number of other methodological and design considerations that our findings were not intended to address. These involve close considerations of the relative strengths and weaknesses of alternative designs for training evaluation and specific validity threats (see, e.g., Eckert, 2000). Moreover, the limitations frequently cited in connection with single case studies also apply here. These typically involve limitations to the generalizability of findings, especially outside of a particular case setting. Nevertheless, some generalizations may be derived from our findings. These are perhaps best described as *petite* generalizations (Stake, 1995) which might be best used to inform further research, for example, using a range of methodological approaches in a variety of organizational settings.

CONCLUSIONS

The problem of identifying and accounting for multiple stakeholder views in training evaluation both deserves and requires increased attention on a number of fronts. Experienced program evaluators may legitimately inquire into the reasons that concepts and frameworks so commonly used in program evaluation have not widely “cross-pollinated” the area of training evaluation. At the same time, rather than lament the lack of evaluation use, training evaluators need to look beyond the set of common training evaluation practices so frequently recycled in their field. As noted by other recent researchers (Twitchell, Holton III, & Trott, 2000, p. 105), training evaluation in business and industry has been stymied by such practice:

“After at least 40 years of bemoaning the lack of evaluation, promoting the value of evaluation, developing methods of evaluation, and pushing the evaluation cause, there appears to have been only modest changes in the amount, types, or quality of evaluation in business and industry.”

We suggest that a solution lies in bringing to bear the collective intellectual capital of the organization to the training evaluation problem through the involvement of multiple stakeholders, and the application of knowledge widely accepted in the broader evaluation field. In particular, training evaluators need to become familiar with modes of evaluation use beyond that of simple, short-term decision making.

Finally, in helping to explain our prior findings, the present results also suggest the need for further inquiry. As noted, most investigations that directly or indirectly result in data about stakeholder diversity in training evaluation have employed qualitative methods. While these have produced results that are extremely useful in identifying diverse stakeholder perspectives, further investigation would be clearly beneficial. Such inquiry might include quantitative techniques based on surveys of various stakeholder populations. Together with the results produced so far, additional study holds much promise to extend and complement our understanding in this emerging area.

REFERENCES

- Alkin, M. C., Hofstetter, C. H., & Ai, X. (1998). Stakeholder concepts in program evaluation. In A. Reynolds & H. Walberg (Eds.), *Advances in educational productivity*, 7 (pp. 87–113). Greenwich, CT: JAI Press.
- Bassi, L. J., & Van Buren, M. E. (1998). The 1998 ASTD state of the industry report. *Training & Development*, 52, 21–43.
- Brown, D. C. (1994). *How managers and training professionals attribute causality for results: Implications for training evaluation*. Unpublished Doctoral Dissertation, College of Education, University of Illinois, Urbana-Champaign.
- Brown, J. S., & Duguid, P. (1991). Organizational learning and communities-of-practice: Toward a unified view of working, learning, and innovation. *Organization Science*, 2, 40–57.
- Bryk, A. S. (Ed.). (1983). *Stakeholder-based evaluation*, New Directions for Program Evaluation, 17. San Francisco: Jossey-Bass.
- Caracelli, V. J., & Preskill, H. (Eds.). (2000). *The expanding scope of evaluation use*, New Directions for Evaluation, 88. San Francisco: Jossey-Bass.
- Cousins, J. B., & Whitmore, E. (1998). Framing participatory evaluation. In E. Whitmore (Ed.), *Understanding and practicing participatory evaluation*, New Directions for Evaluation, 80 (pp. 5–23). San Francisco: Jossey-Bass.
- Daft, R. L., & Huber, G. P. (1987). How organizations learn: A communication framework. *Research in the Sociology of Organizations*, 5, 1–36.
- Darrah, C. N. (1995). Workplace training, workplace learning: A case study. *Human Organization*, 54, 31–41.
- Eckert, A. (2000). Situational enhancement of design validity: The case of training evaluation at the World Bank Institute. *American Journal of Evaluation*, 21, 185–193.
- Fetterman, D. (2000). *Foundations of empowerment evaluation*. Thousand Oaks, CA: Sage.
- Greene, J. C. (1988). Stakeholder participation and utilization in program evaluation. *Evaluation Review*, 12, 91–116.
- Lombardo, C. A. (1989). Do the benefits of training justify the costs? *Training and Development Journal*, 43, 60–64.
- Love, A. J. (1991). *Internal evaluation: Building organizations from within*. Newbury Park, CA: Sage.
- Mark, M. M., & Shotland, R. L. (1985). Stakeholder-based evaluation and value judgments. *Evaluation Review*, 9, 605–626.
- Maxwell, J. A. (1996). *Qualitative research design: An interactive approach*. Thousand Oaks, CA: Sage.
- McLinden, D. J., & Trochim, W. (1998). Getting to parallel: Assessing the return on expectations of training. *Performance Improvement*, 37, 21–26.
- Michalski, G. M. (2000). Learning to work in web time: Evaluating time-to-market instruction at Nortel Networks. In Z. L. Berge (Ed.), *Sustaining distance training: integrating learning technologies into the fabric of the enterprise* (pp. 199–217). San Francisco: Jossey-Bass.
- 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.
- Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis* (2nd ed.). Thousand Oaks, CA: Sage.
- Patton, M. Q. (1996). A world larger than formative and summative. *Evaluation Practice*, 17, 131–144.
- Patton, M. Q. (1997). *Utilization-focused evaluation: The new century text* (3rd ed.). Thousand Oaks, CA: Sage.
- Phillips, J. J. (1996). ROI: The search for best practices. *Training & Development*, 50, 42–47.
- 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. (2000). The learning dimension of evaluation use. In V. J. Caracelli, & H. Preskill (Eds.), *The expanding scope of evaluation use*. New Directions for Evaluation, 88, (pp. 25–37). San Francisco: Jossey-Bass.
- Richards, T. J., & Richards, L. (1994). Using computers in qualitative research. In N. K. Denzin & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 445–462). Thousand Oaks, CA: Sage.
- Sacks, M. (1994). *On-the-job learning in the software industry: Corporate culture and the acquisition of knowledge*. Westport, CT: Quorum.
- Shulha, L. M., & Cousins, J. B. (1997). Evaluation use: Theory, research, and practice since 1986. *Evaluation Practice*, 18, 195–208.
- Stake, R. E. (1994). Case studies. In N. K. Denzin, & Y. S. Lincoln (Eds.), *Handbook of qualitative research* (pp. 236–247). Thousand Oaks, CA: Sage.
- Stake, R. E. (1995). *The art of case study research*. Thousand Oaks, CA: Sage.
- Stewart, T. A. (1999). *Intellectual capital: The new wealth of organizations*. New York: Currency Doubleday.
- Tesoro, F. (1998). Implementing an ROI measurement process at Dell computer. *Performance Improvement Quarterly*, 11, 103–114.
- Trochim, W. (Ed.). (1989). Concept mapping for evaluation and planning [Special Issue]. *Evaluation and Program Planning*, 12.
- Twitchell, S., Holton, E., III, & Trott, J. (2000). Technical training evaluation practices in the United States. *Performance Improvement Quarterly*, 13, 84–109.
- Winch, R. G. (1998). *Telecommunication transmission systems* (2nd ed.). New York: McGraw-Hill.