ABSTRACT
Using experiences from an assessment of an outdoor management training program, this paper examines issues in training evaluation that are not often addressed. We present a five-step model suggesting key questions that must be answered before the appropriate evaluation approach can be determined. We conclude by warning that the clarion call for results-oriented training evaluation may be misguided, given the multiple and subtle purposes that training serves in the life of organizations.

ISSUES IN TRAINING PROGRAM EVALUATION
We identified five issues in training evaluation. To illustrate our discussion, we use the example of an outdoor management skills training (OMST) program in which we have been involved over the past two years.

Work or “Perk”. The first issue is the extent to which the training program is truly an attempt to improve work performance or is in reality a perquisite for job performance already judged successful. Such objectives for the training program are very seldom spoken. To the extent that the true purpose of training is to reward good performers or to renew sagging spirits at company expense, an extensive performance-based training evaluation is misguided. A simple reactions measure, or “smiles sheet,” may be all that is really necessary.

OMST programs are necessarily off-site, and thus are potential candidates for “perks” rather than serious training. Participants in our OMST feel that it definitely is not a perk. A typical reaction is: “I’m glad I came, but I wouldn’t want to do it again.” For this reason, we decided that our evaluation study had to go beyond the mere reaction level.

Substance vs. Symbol. Considerable attention has been paid of late to the concept of corporate or organizational culture: shared values, rites, myths, symbols, and ceremonies. All organizational practices, including training, serve both to do (the practical side) and to say things (the symbolic side). Thus, training can be considered suggestive of the basic underlying culture or the “expressive side” of an organization.

Off-site training programs like OMST are particularly likely to carry ritualistic or symbolic overtones. For example, selection to attend an OMST program can be seen as a symbol of the organization’s confidence in the individuals chosen. Program participants in our OMST case developed an informal network of program graduates that net occasionally, wore distinctive pins to work, and provided quick access to decision makers in other functional areas of the company. Completion of the OMST program essentially symbolized the opening of the gate to this extremely helpful informal network of prior program graduates. In fact, access to this network appeared to be one of the most beneficial outcomes of the program.

External vs. Internal. A related issue is the extent to which desired training program outcomes are internal or external to the program itself. Authors have railed against the use of internal evaluation measures because they assume that all training must be designed to affect some external criterion. However, this is not always the case. For example, some training may be designed primarily to instill company pride, or perhaps inculcate company history and/or philosophy.

OMST is again a potential example of this subtle expectation for training outcomes. Organizations sending employees to such training may be attempting to emphasize some unique aspects of corporate philosophy, such as the importance of individual risk-taking, or the balance between the need for individual excellence and group teamwork. Again, these changes may be unlikely to manifest themselves in behavior or results back on the job, so internal measures (e.g., diary entries of attitudes toward taking risks over the course of the program) may be appropriate.

Behavior vs. Outcomes. Once it has been determined that the purpose of training is truly work related and that the outcomes desired are more substantive than symbolic and more external than internal, the question becomes how to measure the changes back on the job. Basically, the issue here is whether to use a more “objective” quantitative criterion or a more “subjective” ratings criterion. It is not true, as is often assumed that results measures are superior to behavior measures.

Improvement in participant skills are first order outcomes in OMST; results are second order outcomes. Many other potential causes exist for results besides managers’ skill levels. Further, in two years only 37 managers out of the total work force of over 6,000 have been trained in our OMST program. How great an impact on measurable organizational results can this small group be expected to have?

Self vs. Other Ratings. Given the limitations of ratings from any single source, we suggested to our OMST client a criterion involving combined ratings from multiple sources: self, supervisor, subordinates, and peers. We were acutely aware, in particular of the limitations of self-ratings when used by themselves. However, the client was more concerned about the time, cost, and organizational intrusiveness of collecting ratings from sources other than program participants.

CONCLUSION
Training evaluation research is not as straightforward as “textbook” discussions would lead one to believe. Many subtle issues are not typically addressed; in fact, some of the issues are such that even the organizational sponsors of a training program may be unaware of their existence. Our concern here is that the clarion call for increased accountability in the training function may drive out all training that cannot demonstrate quantitative, measurable improvements in performance or results. Such an outcome would be unfortunate indeed. Future training evaluation research needs to approach the problem from a broader perspective, possibly using ideograph qualitative research techniques, to capture important intangible changes that often accompany management training programs.