

Original Article

The Voices of Evaluation

NANCY L. BENNETT, PhD
Department of Continuing Education
Harvard Medical School
Boston, MA

Abstract: *Evaluation in continuing medical education (CME) is a traditional part of planning but is often not approached with the same level of sophistication as other parts of the teaching/learning transaction. An understanding of the models and schema in the literature allows broader and more reasoned thinking about effective evaluation. This article briefly describes and compares several models and schema that are important to adult education and CME. The contributions of qualitative and quantitative evaluation approaches in CME point to the need for the field to create a new model to extend our current thinking.*

Key Words: Continuing education, continuing medical education (CME), evaluation, program evaluation

Evaluation makes clear the path that learning takes in the transaction between the planner, teacher, and learner. It judges, describes, defines, values, shapes opinion, and directs attention. Evaluation is the part of planning that assesses the product of program development to judge whether the planned goals match what learners take from the activity. Most planners, teachers, and learners agree that evaluation is an essential component of learning for everyone involved, yet we often tend to approach it with less intensity than other parts of the teaching/learning transaction. Knowing about the wide range of evaluation options means that we can make deliberate choices about how we evaluate, choices that fit the context and the task. Evaluation can help us decide what to focus on in changing physician

behaviors. This article compares several schools of thought in the field of evaluation as a starting point for that discussion.

Writing about the underpinnings of evaluation is difficult because there is no "right" or "best" approach. The planner must understand evaluation—both the process and politics—to match a theory base with a specific situation. Understanding models provides a base for thinking more rigorously about how to approach evaluation and the role of evaluation in learning. Balancing the proportions of measurement and judgment may not always be obvious. No single evaluation model or schema fits all kinds of learning or provides the best process to answer questions. A sense of the field may help to expand the ways in which we think about evaluation and the places in which continuing medical education (CME) makes special contributions to the discussion. This article proceeds along a continuum from quantitative models to qualitative models to describe some examples of important reference sources. It assumes that many models draw upon both approaches in varying proportions and that both contribute to

Reprint requests: Nancy L. Bennett, PhD, Department of Continuing Education, Harvard Medical School, 641 Huntington Avenue, Boston, MA 02115.

Evaluation in CME

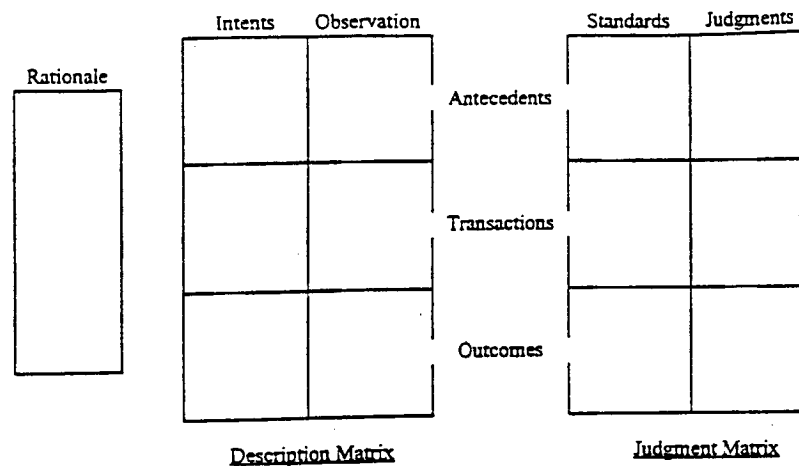


Figure 1 Stake's Countenance Model of Evaluation.

curriculum development. Is one part taking too long or not getting across important distinctions between concepts? Is another section not long enough to allow learners to develop their thinking? The summative role seeks to improve use or recognition of a product outside the planning group. Why did a segment of the intended audience not attend the course? How can you teach a given technique with constraints on equipment and practice time? Those in CME have tended to use formative evaluation to look at a program during its presentation for midcourse corrections and summative evaluation to assess the course on its completion.

While using objectives or goals as the organizing principle for a model of evaluation is a simple concept that practitioners find helpful, other authors recommend moving beyond objectives to a variety of other organizing principles. Stufflebeam^{7,8} used decision making as the central theme in evaluation. He classified four decision types in a matrix of ends and means with intended and actual outcomes, assuming an ongoing process within the context of the project. The CIPP model judges relative values of

competing alternatives. Four types of evaluation address four types of decisions:

- Context—determine the objectives, assess needs, problems, and opportunities (intended ends);
- Input—determine project design or how to use resources (intended means);
- Process—implement decisions to control project objectives (actual means); and
- Product—measure and interpret those things attained during and at the end of a project (actual ends).

The CIPP model defines and applies descriptive and judgmental information about the worth of a project as described by its goals, structure, process, and product. Context evaluation occurs all the time with the other three types of decision information implemented as needed. The model's limits include its assumptions of rationality in decision making and ignore the multiplicity of interests and interpretations by assuming that

Evaluation in CME

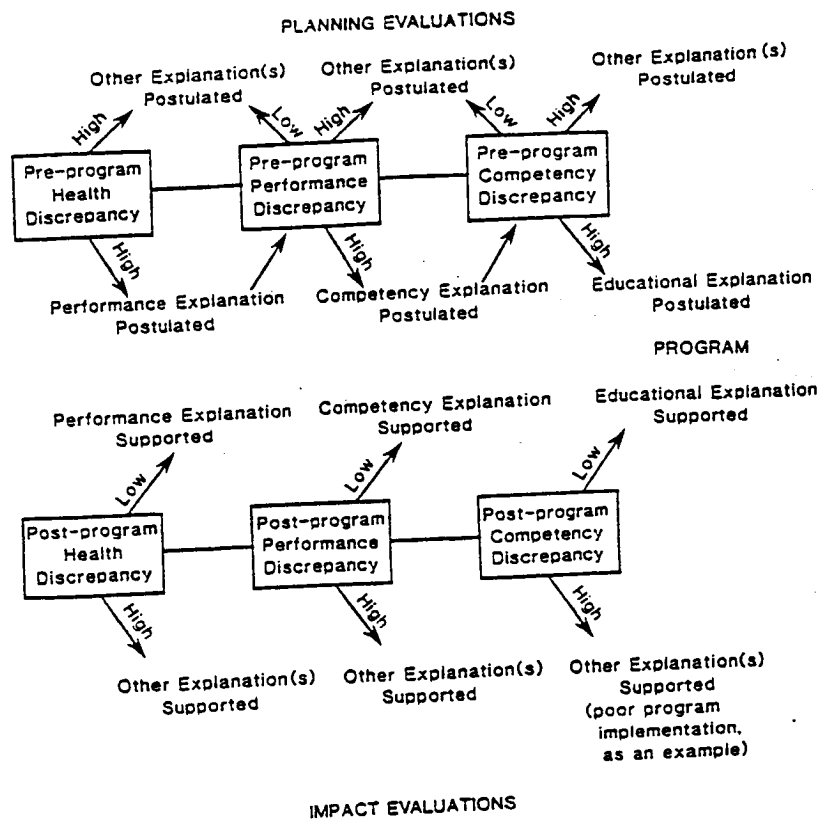


Figure 2 Model of discrepancy analysis in planning, monitoring, and evaluating the CME programs.

conceptualize moving beyond participants' perceptions of a course. Mazmanian et al. suggested one application of a schema like Dixon's.¹⁷

Assuming that evaluation can be a marker of quality, the authors translate the four levels into an evaluation-based approach to accreditation with methods of evaluation that are cumulative and progressive. Level 1 describes methods of evaluation as attendance data and opinion surveys; level 2 includes written tests of knowledge, skills, attitudes; level 3 includes chart-stimulated recall, objectively structured clinical evaluation; and level 4 incorporates peer review of process and patient outcomes.

The complexity of the evaluation methodology is reflected in the types of activities that a sponsor may wish to include in a range of learning options.

Fox proposed a discrepancy analysis model for evaluation specific to CME. The three stages are to (1) establish criteria or standards for acceptable patient health status, (2) define physician performance and competency, and (3) describe and measure the discrepancy. The model explains and predicts in a manner that accounts for the complexity of variables in looking at patient health. Further, it characterizes one way of looking at accountability.¹⁸

10. Irby DM, Bland CJ, Meurer LN. Quantitative and qualitative research: an e-mail exchange. *Acad Med* 1996; 71:687-692.
11. Sechrest L, Babcock J, Smith B. An invitation to methodological pluralism. *Eval Practice* 1993; 14:227-235.
12. Lincoln YS, Guba EG. RSVP: we are pleased to accept your invitation. *Eval Practice* 1994; 15:179-192.
13. Greene JC. Qualitative program evaluation. In: Denzin NK, Lincoln YS, eds. *Handbook of qualitative research*. Thousand Oaks, CA: Sage, 1994.
14. Kirkpatrick DL. *Evaluating training programs: the four levels*. San Francisco: Barrett-Koehler, 1994.
15. Phillips J. Return on investment—beyond the four levels! In: Holton E, ed. 1995 conference proceedings of the Academy of Human Resource Development. Baton Rouge, LA: Academy of Human Resource Development, 1995.
16. Dixon J. Evaluation criteria in studies of continuing education in the health professions. *Veal in the Health Professions* 1978; 1:47-65.
17. Mazmanian PE, Mazmanian PM, Moore MA. Evaluation-based accreditation. *J Cont Educ Health Prof* 1996; 16:112-116.
18. Fox RD. Discrepancy analysis in continuing medical education: a conceptual model. *Mobius* 1983; 3(3):37-44.
19. Houle CO. *Continuing learning in the professions*. San Francisco: Jossey-Bass, 1980.
20. Mazmanian PE. Personal communication, June 1997.