

HERE'S A FRAMEWORK DESIGNED TO HELP YOU
IDENTIFY EVALUATION ALTERNATIVES . . .

EVALUATING TRAINING

BY KAREN S. BRETHER
AND GEARY A. RUMMLER

Most discussions of training evaluation we hear are not very satisfying. Nothing is ever resolved because each discussant seems to be talking about evaluating a different dimension of the training. When people can't agree on *what* they are trying to evaluate and why, they actually won't agree on *how* to evaluate.

In this paper we would like to present a framework for viewing evaluation alternatives and deciding what evaluation is appropriate. Included also are some guidelines for conducting evaluation studies.

Evaluation and the General System Model

A number of aspects of training might be evaluated. The range of reasonable alternatives is suggested by a general systems view of training, which shows the relationship between the training function and the organization it supposedly services. Figure 1 (p. 16) illustrates the key components of an ideal training system, consisting of the

receiving system (in this case, the jobs or organization) and the processing system (the training function). The specific systems components are:

1. The inputs into the system (students or trainees).
2. The processing system, which converts inputs into outputs. Depending on the system in question, the processing system might be an instructional lesson, a classroom, a course, or a training department.
3. The outputs of the processing system (trained, or educated, students or trainees).
4. The receiving system, which is the area or unit into which the outputs immediately go. For a company apprentice program, the receiving system is the job. The processing system and the receiving system are always subsystems of some larger system (e.g., the school, the agency manpower planning system).
5. The mission goal, or stated goal, of the receiving system. This might be "All claims honored within X days with Y errors" where the receiving system is the claims office of an insurance com-

pany and the processing system is the training course for claims representatives.

6. The evaluation of the accomplishment of the stated mission goal (e.g., percentage of claims paid correctly within the stated time). This evaluation consists of measuring the output of the receiving system and matching that output against the stated criteria for the mission.

7. The evaluation of the quality and quantity of the outputs of the processing system (e.g., degree of mastery at the conclusion of the Claims Representative course). This evaluation requires measurement of the processing system outputs and their comparison with the product criteria.

8. The feedback to the processing system on the outputs of the processing system and attainment of the mission goal. Based on this feedback, adjustments can be made in the processing system itself, in the criteria for product of the system, or in both. For example, feedback might indicate that even though Claims Representatives did exceptionally well in their class,

Claims Representatives were deficient in knowledge that they were thought to have acquired in that class. Based on more specific data, it might be necessary to alter the content of the course, to raise the performance standard, or to do both.

The training function has these systems characteristics:

1. Its output is the input to another part of the system. It does not function in isolation. It must contribute to the larger, total system. If it does not contribute, then it will cease to function. Also, any attempts to maximize its output or effectiveness will be neutralized by the need for the total system to optimize all the subsystems. (Specifically, this is done through budget allocations.)

2. It responds to data. It must be correct. It must adapt or die.

3. It is controlled by the evaluation criteria, as it adapts. If it is evaluated on the basis of headcount and popularity, all adjustments will be made accordingly. If it is evaluated on its performance and contribution to the organization, it will correct toward that goal.

System Sophistication

Figure 2 (p. 16) illustrates three degrees of system sophistication, which point up the relative effectiveness of training organizations. The simple input-output training systems are characterized by those training departments that do no evaluation and pride themselves on having unlimited budgets. These are the training organizations that disappear in strenuous economic times because they are unable to show any apparent value to the organization when asked. They always cry "foul," explaining that they were never required to justify their existence before.

A more sophisticated system, the guided system, is still deficient, though somewhat better off. It evaluates its output. If it evaluates according to performance criteria ("Can the trainees do what we set out to have them do?"), then it may very well be an effective training organization. But if it chooses to evaluate (or gets lulled into evaluating) according to the



Karen Brethower.

"Evaluation is frequently thought to be too difficult to do in the real world. Sometimes it is done in a way that undermines its utility."

criteria of the amount of training activity (e.g., bodies trained per budget dollar) and its popularity (e.g., "Did you enjoy the training? Do you think you will find it useful?"), then it is a little better off than the less sophisticated ballistic system. That is, its own internal evaluations do not measure what the organization needs.

Thus, the only truly effective training system is one based on the



Geary Rummler

"Most of the training evaluation today is classified as action research. That is, the research occurs in a real situation with all the constraints imposed by the organization going about getting its work done."

adaptive system model. Without the receiving system in the loop, there is no way of assessing whether the products of this training function (however good they might be by behavior-change standards) are of any value to the organization.

Thus, a general system view of the training function or of a particular training course or experience exposes a number of dimensions which might be evaluated. Figure 3 combines four of the possible evaluation alternatives with the general systems model of training.

An evaluator might evaluate at any one of these four levels to determine whether the training is having the desired effect. The training action in response to the evaluation can be of two major types. He/she can:

1. Decide to continue or discontinue the training (summative evaluation); or

2. Decide to continue the training as is, or to revise any aspect of the training system until it meets the criteria (formative evaluation).

If the evaluation information is to be used to revise the training it must be much more detailed as to specific aspects that worked or failed.

Evaluation Matrix

From the four levels of evaluation we have identified, we can form the evaluation matrix shown in Figure 4 (p. 16). For each level we can ask:

1. What question(s) do we want answered?
2. What might we measure to answer those questions?
3. What are the dimensions of learning or performance we are trying to measure?
4. What are the sources of the data to help measure?
5. What are alternative ways of gathering data for measurement?
6. What are the evaluation criteria we want to apply to each question?

Figure 5 (p. 16) is a partially complete matrix for one type of teaching activity, a management workshop. In the first column, alternative explanations of the "why not" have been made explicit

for the workshop, e.g., I — “unhappiness” might emanate from the concepts not being relevant, inappropriate workshop (WS) design, or pre-workshop setting of expectations. The entries in each cell would vary depending on the course and the situation.

There are several practical applications of the evaluation matrix. The first and most obvious one is as a guide to systematically determine *what* and *how* to evaluate

training at the four levels.

However, a second application (which should be obvious to those readers who engage in performance analysis or front-end analysis) is a test for the existence of a real problem which can be corrected by training (or anything else). For example, if someone in your organization has approached you to develop a report writing course and you are reviewing how such a course might be evaluated prior to

designing it, you might quickly see how you would evaluate at levels I and II. However, if you sat down with the party requesting the training to discuss how you might measure levels III and IV, chances are you would find them unable to help with IV because in fact there is no relationship between the perceived deficiencies of report writing and job performance. Or perhaps there is. The point is, thinking through the evaluation matrix prior to beginning the training effort is one method to approximate a front-end analysis.

The third practical application of the matrix is as a vehicle to gain organization support of a training effort. For example, you are approached to develop a training course on a “quality control” matter. Again, levels I and II should offer you no trouble. A discussion of evaluating III and IV with the management requesting the training will point up the need for management support of the trainees using the new skills or concepts on the job. The matrix makes it clear that all the training in the world around quality will be to no avail unless management takes the necessary steps to assure that use of these concepts will be supported on the job. In fact evaluation results that show passing grades at level I and II and failing grades at III and IV are an indictment of management, not the trainees or the trainers.

And that leads to application four — the use of the matrix to negotiate the degree of evaluation required. Frequently there is a mismatch between the expected or desired degree of evaluation by the requesting party and the training group. The matrix can help clarify how much (to what degree) evaluation is necessary or worth investing in.

Thus far, we have presented a model for training evaluation and some propositions related to that model. The propositions are:

1. The training system be viewed as a subsystem of the total organization.
2. There are four potential levels of evaluation.

I. Do trainees like the train-

