HOW TO PLAN AND IMPLEMENT A SUPERVISORY TRAINING PROGRAM — PART 5

EVALUATING IN-HOUSE TRAINING PROGRAMS

BY DONALD L. KIRKPATRICK

When we complete a supervisory training program, we need some measure of the *"how did we do?"* effectiveness of the program. The answer to this question will give us important information regarding:

1. What kind of a contribution did it make to the organization?
2. Should the program be repeated or continued?
3. How can it be improved?

Nearly all training people agree that the program should be evaluated, but there is great confusion and disagreement about what to do and how to do it.

There are four stages or steps to consider in evaluating a supervisory training program. They are: Reaction, Learning, Behavior, and Results. Each of the four is important and can give valuable information. It is important to understand the meaning of each, its relationship with the others, and guidelines and procedures for doing it.

First, a clarification of the meaning of each is necessary.

1. *Reaction:* How do participants feel about the program? This is actually a measure of customer (participant) satisfaction. The purpose of the training program is to help participants. Measuring reaction is determining how satisfied they are with the product (training) they have received.

2. *Learning:* What knowledge and skills were learned? What attitudes were changed? All training programs are trying to increase knowledge and skills and / or change attitudes. How effectively did the training program accomplish these learning objectives?

3. *Behavior:* To what extent did on-the-job behavior of participants change as a result of the training program? It is obvious that the purpose of the training is to do more than improve knowledge, skills and attitudes in the classroom. On-the-job behavior change is an objective. How much and what type of change actually took place on the job because of the training program?

4. *Results:* What final results did the training program produce? Change in behavior was not the *final* objective we were trying to achieve. It is only a means to an end. The end result should include such things as improved productivity, better quality, lower costs, meeting deadlines, more competition, reduced accidents, improved morale, lower turnover, and ultimately, more profits or better service.

Guidelines

Some guidelines will help determine how to proceed.

In evaluating *Reaction* we want to get honest and also meaningful reactions. This feedback from participants is a first indication of the effectiveness of the program. And decisions are often made on the basis of reactions.

For example, the Management Institute (MI) of the University of Wisconsin-Extension is in the business of offering short practical institutes and conferences for management people from industry, business and government. The reaction of participants determines whether these programs continue to be successful. The job security of the MI faculty depends on reactions. Favorable reactions mean that people will come back to future programs and also encourage others to attend. Negative reactions will reduce future enrollments.

This situation is not unique to outside management-development programs. A very similar situation exists with in-house programs. If supervisors feel negatively about the program, they will tell someone, perhaps their boss. And their boss might tell higher-level management and soon decisions may be made to cut out the program because "it's no good." Supervisors might look for excuses for not attending programs they don't feel will be enjoyable. Or, they might attend with negative attitudes. All of these results would be detrimental to supervisory training.

Five guidelines will help to assure meaningful and honest reactions.

1. Determine what information you want to get. For example, you would probably want reactions to subject, leaders, schedules, and possibly meals.
2. Design a written comment sheet to get this information.
3. Use questions that can be tabulated and quantified instead of open-ended questions such as: *"What did you like best?" "What did you like least?"*
4. Encourage participants to write comments to explain and supplement the questions that will be tabulated.
5. Do not have the forms signed or otherwise identified. (This will assure honest answers.)

An example of a very simple form is provided in Figure 1. In using reaction sheets, be sure to allow enough time for participants to complete them. Also, be sure to collect them *before* participants leave the classroom. Don't have them take them along and send them back.

Figure 1.

REACTION SHEET

Please give us your honest reactions. They will help us determine the effectiveness of the program and offer suggestions for improvement.

1. How do you rate the subject?
   - ☐ Excellent
   - ☐ Very Good
   - ☐ Good
   - ☐ Fair
   - ☐ Poor

   COMMENTS:

2. How do you rate the leader?
   - ☐ Excellent
   - ☐ Very Good
   - ☐ Good
   - ☐ Fair
   - ☐ Poor

   COMMENTS:

3. How were the facilities?
   - ☐ Excellent
   - ☐ Very Good
   - ☐ Good
   - ☐ Fair
   - ☐ Poor

4. What would have improved the program?

Evaluating Learning

In evaluating learning, we want to determine what knowledge and skills were learned and what attitudes were changed. These are the three ingredients that can cause changes in on-the-job behavior.

We could ask the participants, *"What did you learn?"* But the responses would be subjective and wouldn't be reliable as a measure of learning. Instead, we should do it objectively by using the following guidelines:

1. Measure knowledge, skill and/or attitudes on a before (pretest) and after (posttest) basis.
2. Analyze the responses to compare pretest and posttest *total scores* as well as pretest and posttest responses to *each individual question.*
3. Where practical, use control groups that do not receive the training to compare with experimental (test) groups that receive the training.

If we are going to measure knowledge or attitude change, a paper-and-pencil test is appropriate. Standardized inventories can be purchased and used if the content of the test is related to the content of the training program. The other alternative is to develop your own test to cover the knowledge and attitudes that you are trying to teach. The test design can include any or all of the following approaches:

- "True" or "False"
- "Agree" or "Disagree" (or "Strongly Agree," "Agree," "Neutral," "Disagree," or "Strongly Disagree.")
- Multiple choice
- Fill in the blank

Where you are teaching skills instead of knowledge and attitudes, a performance test instead of a paper-and-pencil test is necessary. For example, if you are teaching "oral presentations," you will have participants make several presentations and have each one evaluated. The same approach should be used when teaching "written communication," "interviewing," "effective reading" and other skills.

On-the-Job Behavior

We must go out of the classroom to measure changes in on-the-job behavior. This becomes more difficult and time-consuming than measuring reaction and learning. We have to make such determinations as:

- When is the right time to evaluate? (Three weeks after the program? Three months after the program? One year after the program? Or when?)
- Where is the best source of in-

"The most difficult stage of evaluation is to determine what final results were accomplished because of the training program."

formation on the behavior of the participants? (The participant's boss? The participant's peers? The participant's subordinates? Or the participants themselves?)

Here are some guidelines that should help:

1. Measure behavior on a before-the-program and after-the-program basis. (This is better than only measuring it after the program and trying to determine what is being done differently.)

2. Allow enough time between the program and the after-the-program evaluation to allow for change in behavior. (It may be a good idea to evaluate at several different times, i.e., three months and six months.)

3. Use as many sources as practical. (Boss, peers, participants themselves and possibly subordinates.)

4. If practical, use a control group (not receiving the training) to compare with the experimental group that receives the training. This comparison will indicate changes in behavior caused by factors other than the training program.

These guidelines require a great deal of expertise and time to implement. This type of evaluation should only be done if the program is an extensive one and if it is going to be repeated in the future. Otherwise, it probably isn't worth it.

From a practical standpoint, a much simpler approach is possible. Even though it isn't as reliable, it can provide evidence as to whether or not behavior change is taking place. When using it, the limitations of its value must be recognized.

1. Decide what behavior changes you expect to take place.

2. Prepare interview questions to see whether these changes have taken place.

3. After the program, interview selected people (i.e., participants, bosses, peers, and/or subordinates) and try to find out: "What is the participant doing differently now than before the program?"

3. Quantify the responses to determine the impact of the total program on participant's behavior.

Note: If subordinates are going to be interviewed, *be sure* that the participants themselves do not object to the approach.

Determining Results

The most difficult stage of evaluation is to determine what final results were accomplished because of the training program. There are so many factors that influence results such as production, quality, costs, turnover, accidents, and profits that it is frequently impossible to prove that the training program caused the desired result. Some programs such as "work simplification" and "value analysis" can be evaluated in terms of results because the cost versus benefits can be quite readily measured. In training areas like leadership and communication, there is no way that tangible results can be directly related to the program.

Here are some guidelines for evaluating results:

1. Measure the conditions before the program and compare with the conditions after the program. Use tangible data.

2. Try to eliminate other factors that could have caused changes in the results. A control versus experimental group is one possibility. Another possible approach is indicated by the following example:

Example: Suppose you are trying to reduce turnover among new hires. You have decided that much of the turnover results from the poor job of orientation and training done by the supervisors. Therefore, you conduct a supervisory training program in which you teach the supervisors to orient and train new employees effectively. The program is conducted in April. Here are some figures:

Turnover Rates:

Jan.	Feb.	Mar.	Apr.
6%	5%	6%	7%
May	**June**	**July**	**Aug.**
2%	1%	2%	1%

At first glance, it seems that the program has been highly effective. However, the question must be asked, *"What else could have caused the reduction in the turnover rate?"*

One obvious possibility is that the turnover rates are seasonal. Therefore, we must look at last year. Here are the figures:

Turnover Rates:

THIS YEAR:

Jan.	Feb.	Mar.	Apr.
6%	5%	6%	7%
May	**June**	**July**	**Aug.**
2%	1%	2%	1%

LAST YEAR:

Jan.	Feb.	Mar.	Apr.
7%	6%	5%	7%
May	**June**	**July**	**Aug.**
6%	7%	5%	7%

The figures still look good and evidence is building up that the training program was effective in achieving the desired results. But we must continue to ask the question, *"What else could have caused the improvement?"* Some other things we'd have to consider are:

1. What was the employment situation in the community? How does it compare with last year's? What has happened to turnover figures in other organizations in the community and possibly the industry?
2. What else has my organization done that could have affected turnover? For example, has there been any change in pay, benefits or employment practices?

It is often impossible to *prove* that results (benefits) outweigh costs. Frequently we must be satisfied in finding *evidence* that desired results have been achieved.

Donald L. Kirkpatrick is professor of Management Development at the University of Wisconsin-Extension, Milwaukee. A past national president of ASTD (1975), he has authored numerous articles and books.