Seven Keys to Unlock the Four Levels of Evaluation

by Donald L. Kirkpatrick

If you are a professional trainer or performance improvement consultant, you are probably familiar with my "four levels" of evaluation for measuring the effectiveness of training or Human Performance Technology (HPT) programs. These four levels of evaluation include Reaction (Level 1), Learning (Level 2), Behavior (Level 3), and Results (Level 4). The purpose of this article is to highlight seven keys for implementing the four levels.

The Seven Keys

First Key: Analyze Your Resources

To do this, you must answer the following questions:
* Does your job consist of only one function—evaluating training programs—or does it include other and perhaps more important duties and responsibilities of planning the curriculum and teaching?
* How large a staff do you have for evaluation?
* How much of your budget can you spend on evaluating programs?
* How much help and cooperation can you get from other departments such as Human Resources or sales if you are evaluating sales training programs?
* How much support and help can you get from line managers if you are training their subordinates in programs such as Leadership Development for Supervisors?

The answers to these and similar questions will dictate what you can and cannot do in evaluating programs.

Second Key: Involve Your Managers

If you are going to be effective in evaluating programs, you need to have your managers' encouragement and support. If they have negative attitudes toward you or your programs, you will not be able to evaluate effectively. Here are some suggestions for getting them "on board":
1. Ask for their input in deciding on subject content. If you are training or consulting with supervisors (or their subordinates), ask them what knowledge, skills, and attitudes you should teach to help them do their job better. This can be done by a survey or personal contact. George Odiorne, in one of his books, made the following statement: "If you want people to support your decisions, give them a feeling of ownership." And this is what you can do by getting them involved in determining subject content for your programs.

2. Get your managers to establish an encouraging climate regarding the program. For example, ask them to do the following before sending their subordinates to a program:
   - Discuss the program with them and encourage them to attend.
   - Tell them to learn what they can and to think about how they can apply what they learn to their job.
   - Tell them that when they return from the program, you will discuss the program with them and help them decide what ideas can be put to use. You might even ask them, "Do you have any ideas for me to apply as your manager?"

   Do you realize the impact this could have on the attitude of supervisors and their motivation to listen, learn, and take practical ideas back to their jobs?

   A number of companies are doing this. Are you?

3. Ask for their helping evaluating the program. Levels 3 (Behavior) and 4 (Results) require this help. You can evaluate Levels 1 (Reaction) and 2 (Learning) without involving managers because you have control over these two levels. But Levels 3 and 4 are typically influenced by factors beyond those within your control. Here is where you must use influence.

For example, to evaluate Level 3, you need to find out what behavior change took place because of the program. I did such an evaluation of a three-day course I conducted at the University of Wisconsin, "Developing Supervisory Skills." Three months after the program, two research students and I conducted patterned interviews in the companies with those who had attended. On each of the subjects we taught, we asked participants about the extent to which they had changed their behavior as a result of attending the course. Then we asked their managers to describe the extent to which they had seen changes in participants after the course. We could not have done this without the cooperation of the managers.

Third Key: Start at Level 1 (Reaction) and Continue Through Levels 2, 3, and 4 as Resources Permit

Some organizations skip the first four levels and go directly to ROI. Others do not consider Reaction that important and go directly to one of the other levels. But don't do it!

Some trainers or HPT professionals refer to Level 1 as "happiness ratings" or "smile sheets," and I agree! That's exactly what they are. They measure the reaction of the participants to the program. But those trainers also claim that these evaluations are not of much value. I disagree. I consider Reaction to be a measure of customer satisfaction. Whether they pay to attend programs or not, they are your customers. As a professor at the University of Wisconsin Management Institute, I know that positive reaction kept us in business. Sometimes a high-level manager would attend one of our programs to decide if it was worth it to send their supervisors to our programs. And they better "smile" when they leave, or there would be no further business from that organization. We offered a special certificate for attending a 10-day program, divided into two five-day sessions. The programs were scheduled about six weeks apart to allow the participants to go back to their jobs and apply what they had learned. Whether they returned for the second half of the program depended on their reaction to the first five days.

In business, industry, and government, there is a slight difference. First, they may not pay for the program, and the existence of the program doesn't depend on their attendance. But you can be sure that they will be telling somebody—perhaps even their boss—whether they thought the program was worthwhile. And there is a good chance that the word will get around whether the program was worthwhile. Top management might even make decisions about the future of the program on the basis of what they hear.

And when the reactions are positive, the chances of learning are improved. The participants become more involved and interested. If participants are sitting in their places with a negative attitude, there is a good chance they are not eager to learn new knowledge or skills. And there will certainly be no positive change in attitudes. The desired result of the program is improved behavior, with positive results to follow. If changes in behavior result without measuring learning, there is no way to tell whether the change came from the training or from other sources. And if behavior change is not evaluated, there is no way to tell whether the results came from the training program or from other sources.

To look at it in a different way, the reason for training is to get better results. To get better results, behavior change needs to occur. And for behavior change to occur, the needed knowledge, skills, and attitudes must be taught in the training program. And to determine whether the results came from the training program, Levels 2, 3, and 4 must be evaluated.

A professional trainer from a well-known computer company asked me, "Is it all right to evaluate results without evaluating behavior?" I told her "no" because it is necessary to determine the reason improved results did or did not occur, and evaluating learning and change in behavior will provide the answer.

The first key is to "analyze your resources." Sometimes they are so small that all you can do is evaluate reaction. And this is important to do in all programs. Higher-level management realizes that the resources are small, and they would be happy to know that the participants reacted favorably to the program. This may be all they expect, and they may even be satisfied if the reactions are positive because they

have so much confidence in their supervisors and others who attend the training program.

If you have the resources to evaluate Level 2, do it. You have control over the content and teaching of the program, and positive learning results will tell higher-level management that you have done your job.

Evaluating Level 3 depends on your success in implementing the second key, which is involving your managers. The same is true for evaluating Level 4. In evaluating these levels with multiple variables, you must rely on "influence." If you can get managers involved and helpful, your limited resources may be enough.

In summary, start with Level 1 and proceed through the other three levels in order as resources and cooperative managers permit. Do not skip any levels, and—most important—do not jump to ROI without evaluating the four levels.

Fourth Key: Evaluate Reaction

Here are the guidelines for evaluating Reaction:
1. Decide what you want to find out—make a list of items to which you want the reaction of the participants (i.e., subject content, leader's effectiveness, schedule, audiovisual aids, handouts, case studies, facilities, meals, etc.).
2. Design a form that will quantify reaction. The most common form consists of a five point scale: either Excellent, Very Good, Good, Fair, and Poor; or Strongly Agree, Agree, Neutral, Disagree, and Strongly Disagree. The objective is to get as much information as possible in the shortest period of time. Participants are not eager to spend time writing the answers to questions.
3. Provide the opportunity for written comments. End your reaction sheet with the question, "What would have improved the program?"
4. Get 100% immediate response. When the participants leave the program, have them put their completed reaction sheet on the back table. Do not tell them to fill out the form and send it back, or do not ask them to email their reactions. If you do either of these, you will not get enough responses to represent the entire group.
5. Be sure you get "honest" answers. Tell participants you want their honest answers and do not ask them to sign the form.
6. Establish an acceptable standard for their combined reaction and tabulate the forms to see if you achieved or exceeded that standard. For example, on a five-point scale, you establish a standard of 4.2 as being acceptable. Tabulate the forms (5 for Excellent or Strongly Agree, 4 for Very Good or Agree, etc.), and compare the result to your standard. If the reaction does not meet your standard, decide what changes to make, such as a new approach, new handouts or audiovisual aids, a new leader, or suggestions for the leader based on answers to the question, "What would have improved the program?"

Fifth Key: Evaluate Learning

Here are the guidelines for evaluating Learning:
1. Measure before and after knowledge, skills, and attitudes.
2. Use a form the participants can complete for evaluating knowledge and attitude change.
3. Use a performance test for evaluating skills.
4. Get 100% response.
5. For knowledge and attitudes, design a test that measures what you want them to know and the attitudes you want them to have at the end of the program. A pretest (before you conduct the program) will measure what they know and their attitudes before the learning begins. A posttest will measure their knowledge and attitudes after the program is over. The difference will be what they have learned.
6. A question that usually arises about the pretest and posttest is whether the same form can be used or if "Form A" and "Form B" should be developed. There are too many problems when you try to develop a "Form A" and "Form B" that will cover the same knowledge and attitudes. So use the same form.

Sixth Key: Evaluate Behavior

Here are the guidelines for evaluating Behavior:
1. Measure on a before-and-after basis if practical. If this is not practical, the alternative is to measure after the program and ask, "What changes in behavior have occurred since you attended the program?"
2. Allow time for behavior change to take place. This time is a difficult determination because change in behavior may occur immediately after the program, or not until 3 months or 6 months after the program, or maybe never. The best compromise seems to be 3 months after the program.
3. Use a patterned interview or written survey asking the same questions of all respondents. One important question to include is, "Do you plan to change your behavior in the future?"
4. Decide who will be polled. For example, the following options are possible:
    *   The participants
    *   The bosses of the participant
    *   The subordinates of participants
    *   Others who would know the participants' behavior
    The answer, of course, depends on your resources, including time. However, there is a caution about the third option. Most participants do not like to get criticism (sometimes called "suggestions") from subordinates. If you are familiar with 360-degree feedback, you will know what I mean.
5. Based on the fact that some participants have not changed their behavior but did answer positively the question, "Do you plan to change your behavior in the future?" repeat the research after 3 more months.

Seventh Key: Evaluate Results

Here are the guidelines for evaluating Results:
1. Measure on a before-and-after basis.
2. Allow time for results to develop—perhaps 6 months or a year.
3. Repeat at appropriate times.
4. Use a control group if practical. A "control" group is individuals who did not attend the program. An "experimental" group is the participants. Unless you are from an organization that has many resources for comparing the two groups, this is difficult and probably cost prohibitive. If you have the resources, this approach is an effective way of "proving" or at least getting "strong evidence" that the results came from the program and not from other sources.

A Final Word

Incribing the last four "keys," I have obviously not answered the question, "How do I do it?"

I have not included possible forms for measuring Reaction, instructions on how to design a test for measuring Learning, questions for the patterned interview, or a survey for measuring Behavior. Nor have I gone into detail on how to measure Results.

This will require further research on your part using the many books and resources available from the growing network of evaluation professionals. ◢

References

Kirkpatrick, D.L. (2001). *Managing change effectively.* Woburn, MA: Butterworth-Heinemann.

Kirkpatrick, D.L. (2001). *Developing supervisors and team leaders.* Woburn, MA: Butterworth-Heinemann.

Kirkpatrick, D.L., & Kirkpatrick, J.D. (2006). *Evaluating training programs: the four levels* (3rd ed.). San Francisco: Berrett-Koehler.

Kirkpatrick, D.L., & Kirkpatrick, J.D. (2005). *Transferring learning to behavior.* San Francisco: Berrett-Koehler.

Kirkpatrick, D.L. (2006). *How to conduct productive meetings.* Alexandria, VA: American Society for Training and Development (ASTD).

Kirkpatrick, D.L. (2006). *Improving employee performance through appraisal and coaching* (2nd ed.). New York: AMACOM.

Odiorne, G. (1981). *The change resistors.* New York: Prentice-Hall.

Donald L. Kirkpatrick holds BA, MA, and PhD degrees from the University of Wisconsin in Madison. A member of *Training Magazine*'s Hall of Fame, he is the author of seven management books, including the 3rd edition of *Evaluating Training Programs: The Four Levels* (2006), which has become the basis for evaluation all over the world. His other books include: *Improving Employee Performance Through Appraisal and Coaching* 2nd edition (2006); *Developing Supervisors and Team Leaders* (2002); *How To Conduct Productive Meetings* (2006); *Managing Change Effectively* (2002); and, with son, Jim, *Transferring Learning To Behavior* (2005). A regular speaker at national conferences and an international consultant, he has presented programs to many U.S. organizations and those in many foreign countries, including Singapore, Korea, Argentina, Brazil, Saudi Arabia, Malaysia, Greece, Netherlands, Spain, Australia, and India. At the Management Institute of the University of Wisconsin, Kirkpatrick taught human relations topics to managers at all levels. In industry, he served as Human Resources Manager of Bendix Products Aerospace Division and Training Director for International Minerals and Chemical Corp., where he developed a Performance Appraisal Program. He is a past national president of the American Society for Training and Development, where he received the 2003 award for Lifetime Achievement in Workplace Learning and Performance. He may be reached at www.donaldkirkpatrick.com.