



## Kirkpatrick – Making the Four Levels Work

Julie Alexander, CIT



16902 El Camino Real, Ste. 2B

Houston, TX 77058

(800) 480-1128

[www.tdshou.com](http://www.tdshou.com)

Almost everyone involved in the training or Learning & Development world knows about Donald L. Kirkpatrick's Four Levels of Evaluation, developed in the 1950's and further refined in the late 1990's. And most of us have used at least the first level to look at our participants' satisfaction with the learning events we create. However, these first level "smile sheets," as they are often called, are just one part of the process of determining the success of a training intervention, and effectively using levels 2 through 4 becomes increasingly challenging. In addition, many of us find ourselves in a position today where management insists that we justify the cost of our existence and their continued support of our training initiatives. In order to do this, we need to take a fresh look at Kirkpatrick's Four Levels and determine how to best leverage them to create and deliver training programs that not only turn out *happy* participants, but ones who are able to make changes in their behavior that positively impact the expectations of management, and ultimately the bottom lines of their organizations.

## Before Evaluation

### Managing expectations

So where do we start? Long before we begin to create questions for any type of evaluation, there are a number of steps we must take if we expect to be able to create effective assessments at any level once the training is complete. Donald Kirkpatrick, in his book "Evaluating Training Programs: The Four Levels," states that there are ten factors that "should be carefully considered when planning and implementing an effective training program" (pg. 3). This paper will not deal with all ten of these factors, but does include several of them. The first thing we need to do is define the need. To start, ask the question, what are the expectations of the stakeholders of the program? There is an element of performance consulting involved in this part of the process, and trainers need to be able to ask the right questions of the appropriate people in order to be able to create effective interventions. As any experienced performance consultant or trainer will tell you, it isn't always wise to take at face value a statement by management of "we must have training." You will need to ask questions of managers and/or supervisors, as well as employees, to determine the true cause of the performance gap. Is the issue something that will even be helped by a training intervention? There are three basic types of "gaps" or issues that can cause performance to be sub-standard. They can be referred to as Will, Hill, or Skill.

- **Will** – If it is determined that the employees have all of the support and/or tools they need to do the job right, and they really do know how to do it (in other words, they could do it right if their lives depended on it), then the issue is one of motivation. In most cases, training isn't going to help with this kind of gap.
- **Hill** – If the employees want to do a good job, and they know how, but still somehow can't make it happen, then there is probably some lack in support by supervisors or management, some of which might include not having the right

tools, time, or feedback. No amount of training, no matter how perfect it is, will help with this kind of gap.

- **Skill** – If the employees simply can't do it right, even though they want to and have all the tools and support needed, training is probably the right response. If this is the kind of gap that you find, and you are able to clearly define what the employees should be doing, what they are doing, and what the difference is, then you can probably move forward to begin the process of creating a training intervention.

### **Enlisting management support**

The success of your training will depend greatly on the commitment of supervisors and management to the process. Research today tells us that the learning event itself is not the most important element of training when it comes to actually achieving results. It's still essential to have a solid, well-written and conducted learning intervention and evaluation process. However, no matter how good that intervention is, you will not realize the full potential of performance improvement without support from management both before and after the event itself. This would include managers or supervisors meeting with employees before they attend the training to establish expectations for what will be accomplished. In addition, managers and supervisors would need to provide coaching and support after the training to help the employee put their new knowledge and skills into action and achieve the intended performance improvements. Building these management components into the actual implementation plan of your training intervention will help ensure a better final outcome.

### **Writing objectives (with evaluation in mind)**

Once you have management onboard, it's time to drill down to the specifics of what needs to be trained and write the objectives. This is one of the most critical elements of creating effective training that meets the expectations of stakeholders. Creating well-constructed objectives will not only make it easier to design the rest of your course and help the participants understand what they are expected to learn, it will also make creating effective evaluation instruments that much easier. According to many current publications, the best objectives will be stated in the form of a performance, and have three parts:

1. What the participant will be able to do differently after the training (whether it is an actual performance, or in some tangible way proving that they have gained specific, important knowledge),
2. The conditions under which the performance will be expected to take place, and
3. The standards by which success or failure will be measured.

As you can see, the way each objective is constructed will give you a head start in creating evaluations that will measure the participants' success in achieving what is

expected of them. It is good to begin developing the evaluation design immediately after creating the objectives. Once you have created clear, well-written objectives and the evaluation instruments to test your participants' success, the content of your intervention will nearly write itself.

### Level 1: Reaction

So, let's start looking at the actual Four Levels themselves. We are most familiar with Level 1, often referred to as the "smile sheet." Because these evaluations can't really tell us whether the participant actually learned anything, some organizations feel that they aren't necessary. However, it would be counter-productive to exclude the level 1 evaluation entirely. In their book, "Transferring Learning to Behavior," the Kirkpatrick's (Donald and his son, Jim) state that "it is unrealistic to expect learning to occur if steps haven't been taken to create a positive learning atmosphere," (pg. 58). And the level 1 evaluation is what tells you whether or not you have succeeded in doing so.

In order to set the stage for positive level 1 evaluations, training should be designed following solid Adult Learning principles, providing plenty of opportunities for participant interaction and practice, and including appropriate facilitator feedback. For instructor-led training, the facilitator's knowledge of the subject matter is obviously important, and so is his or her skill at presenting the training. Most of us have attended a course with a trainer who was a subject matter expert at the course topic but was either deadly boring or spoke over the heads of the students. These are obviously traits that need to be discovered, and a Level 1 evaluation is a good tool for doing so.

Getting honest responses is critical and most organizations allow participants to remain anonymous when providing feedback. Whether you allow this or not, it's important to create a form that will enable you to quantify your participants' reactions. Providing a rating scale with five or six choices ranging from poor to excellent is one way of achieving this. The book "Evaluating Training Programs" provides a number of examples. It's important to decide in advance what numerical values will be acceptable and at what point they will trigger changes in the program. It will also be important to actually take the step to make changes if the ratings reach your pre-established trigger points.

The ratings you gather from these types of questions only provide part of the picture. Be sure to review written comments as well and do your best not to dismiss them out of hand when they seem extreme. Give each one careful consideration and if the suggestion would make the program better, try to incorporate as much of it as you can.

And in all cases, do whatever you can to obtain 100% immediate response. Keep the form short enough that people will stay and complete it after class. One suggestion is to make it a trade – if you have a certificate to present to the participants, they give you their evaluation and *then* you give them their certificate.

## Level 2: Learning

This is the level at which we begin to use the learning objectives established earlier. Level 2 evaluations measure the acquisition of new knowledge and skills, *not* actual on-the-job performance. If the objectives were written well, creating evaluation questions or exercises to test the participants' ability to meet the objectives should be fairly easy. At this level we want to determine the participants' level of success at gaining the intended knowledge, skills, and even attitudes in some cases.

It is obvious that this is a crucial step that must be followed before there can be any change in lasting behavior. A Level 2 evaluation could be as simple as a multiple choice test based on the information presented in the training. In many cases, this can be combined with actual demonstrations of performance to allow for a more comprehensive test of gained skills. There are times when management will want to see documentation of the increase in learning – a before and after picture if you will. One way to show an increase in learning is to use a control group of similar makeup as the one to be trained and administer a test to each group after the training takes place. A process that is easier for most organizations is to provide a pre-training evaluation to the target audience that can be compared to another that is administered after the training.

There are numerous publications available to assist with creating effective test questions. This paper will not spend much time on this topic; however, here are a few things that stand out as important elements.

- Provide clear and concise instructions for your test, and make sure everyone has the same instructions.
- Don't take statements word-for-word from the participant's training materials. This only tests their ability to memorize, not their understanding of the topic.
- Use True/False questions sparingly – the participant always has a 50/50 chance of getting it right so it doesn't always tell you whether they learned anything or not.
- Be sure that each item presents only one question.
- In multiple choice questions, use items with only one correct or clearly best answer.
- Don't use "trick" questions – in most cases they don't tell you what you want to know and frequently just frustrate the participant.
- Avoid using "none of the above" and "all of the above" statements in multiple choice questions. If you can't think of enough valid alternatives, you may not be writing the right kind of question.

- Have subject matter experts review the test before administering it to ensure that you have well-worded questions and correct answers.

As we said earlier, you can have the best Level 1 and 2 evaluations in the world, with very high participant scores on both, and it still doesn't guarantee that there will be lasting, positive behavior changes once they return to work. This is where the Level 3 evaluation comes in.

### **Level 3: Behavior**

How do you know whether the high-scoring participants in your class will actually behave differently once they return to work? In order to measure at this level, you will need to have done some work before the training began. If it is possible to have a control group for which job performance has been measured, this should be a bit easier. If that isn't possible, measurements of performance for the intended audience of your training need to be made before implementing the training. These measurements should be specific enough to target the intended training content. For instance, if the training is intended to remedy a problem with product quality, measuring the amount of product produced isn't really going to tell you what you need to know. Be sure the measurements match the intention of the training intervention.

Another important point is to allow enough time for behavior change to actually take place. It's impossible to calculate exactly when performance will change, so you will need to wait. Depending on the situation, this waiting period could be anywhere from a few weeks to a few months. Because different people may change their behavior at different rates, it is a good idea to repeat the evaluation several times, if you can. One important element to remember during this time, however, is the critical part that is played by management and supervisors in whether or not the new behavior is embraced by the employees. Part of the implementation plan created by the trainer should include, if possible, meetings with managers and supervisors to help them encourage their people to use the new knowledge and skills they have acquired.

Behavior level evaluations can take the form of questionnaires, interviews, and observation of the employee doing their job. The Kirkpatrick books provide additional examples and recommendations. The bottom line here is that there can be no positive business results unless behavior changes, not only immediately after training, but that it remains consistently better than before the training took place. If behavior has indeed improved, then you can move on to Level 4.

### **Level 4: Results**

As stated in "Evaluating Training Programs," (pg. 63), many people believe that level 4 is "the most important" of all the levels. Some feel that it is also the most difficult to achieve, but it doesn't have to be. Here again, it is vital to have taken measurements of

business impact prior to any training taking place. Without these measurements, there is nothing to compare the new status with. As with level 3, you need to allow time to pass in order for results to be achieved and measured, and you will probably need to measure results more than once. Management will need to decide how important it is to have proof of benefit as there is always a cost involved in measuring, especially at this level. It may be necessary to settle for evidence that there is a probable return rather than solid, quantifiable proof of the return on investment.


In order to make an impact at this level of evaluation, the training program needs to be focused on the business benefit from the beginning. Objectives, content, activities, and earlier evaluations should all be written in a way that they drive the learning toward behaviors that are geared to making a specific, measurable, business impact. For instance, take a case where there is a problem with higher than acceptable off-spec product, and it is determined that a course on quality control will provide employees with the knowledge, skills, and improved attitude needed to reduce the quality errors. The amount of off-spec product is obviously one measure that must be performed in advance, but there may be others. It would also be necessary to observe and evaluate the employees' actions to determine where the problem is occurring, and this may result in yet more measures that should be made. Objectives must be written to address the specific actions or lack of actions on the part of the employees that are resulting in the problems.

So, the success of Level 4 evaluations depends on the accuracy and detail of the measures taken before the training, objectives and content that focus on impacting these measures, and the ability to repeat these same measures after training has occurred and enough time has passed to allow for improvement. Evaluation Levels 1 through 3 should also have been completed, and when they are done right they can actually give you a good idea as to how your Results evaluation will turn out.

### **Conclusion: Return on Expectations**

Remember that the Four Levels do not, by themselves, ensure that a training event will be successful. Engaging stakeholders, managers, and supervisors from the start to define expectations and to obtain a commitment to support the employees in putting their new skills into practice will be critical to the success of any training intervention.

By now it should be clear that the Four Levels of Evaluation build on each other. If Levels 1 and 2 have not been done before conducting a Level 3 evaluation, it will be impossible to show that training was the deciding factor in the behavior change. Likewise, without the Level 3, any positive results discovered in a Level 4 evaluation could be attributed to other factors. In many cases it will be difficult, if not impossible to absolutely “*prove* that a particular training event led to a specific bottom-line contribution.” (Implementing the Four Levels, pg. 124) However, as the Kirkpatricks suggest, by evaluating as much of the four levels as possible, in sequence, you should “be



## Kirkpatrick – Making the Four Levels Work

able to *build a compelling chain of evidence as to the value of learning to the bottom line*” (Implementing the Four Levels, pg. 123). And if you have done a good job of defining your stakeholders’ expectations at the start, showing a Return on Expectations will be easy.

## References and Further Reading

Kirkpatrick, Donald L. *Evaluating Training Programs: The Four Levels*. San Francisco: Berrett-Koehler Publishers, Inc., 1994.

Kirkpatrick, Donald L. and Kirkpatrick, James D. *Transferring Learning to Behavior: Using the Four Levels to Improve Performance*. San Francisco: Berrett-Koehler Publishers, Inc., 2005.

Kirkpatrick, Donald L. and Kirkpatrick, James D. *Implementing the Four Levels: A Practical Guide for Effective Evaluation of Training Programs*. San Francisco: Berrett-Koehler Publishers, Inc., 2007.

Basarab, Sr., David J., Root, Darrell K. *The Training Evaluation Process*. Norwell: Kluwer Academic Publishers, 1993.

Mager, Robert F. *Measuring Instructional Results or Got A Match? How to find out if your instructional objectives have been achieved*. Atlanta: CEP Press, 1997.

Tuckman, Bruce Wayne. *Evaluating Instructional Programs*. Newton: Allyn and Bacon, Inc., 1985.

Deming, Basil S. *Evaluating job-Related Training: A Guide for Training the Trainer*. Washington, D.C. American Society for Training and Development, and Englewood Cliffs. Prentice-Hall, Inc., 1982.

Brinkerhoff, Robert O. and Apking, Anne M. *High Impact Learning: Strategies for Leveraging Business Results From Training*. New York. Basic Books, 2001.