

# AN OVERVIEW OF THE EVALUATION PROCESS FOR MINE TRAINERS

By Launa Mallett<sup>1</sup> and Dana Reinke<sup>2</sup>

## ABSTRACT

Interest in training evaluation in the mining industry was piqued with the promulgation of Title 30, Part 46, of the Code of Federal Regulations. Under Part 46.3, a training plan is considered to be approved if it contains certain information, including “The evaluation procedures used to determine the effectiveness of training.” The present paper is a broad overview of training evaluation and is intended to give trainers and decision-makers a framework for planning or assessing training evaluation strategies. It discusses questions to consider when starting an evaluation plan, Kirkpatrick’s model of evaluation categories, and various ways to collect data. It does not provide detailed instructions on how to develop evaluation methodologies, but reviews topics to consider when creating an evaluation plan.

Training evaluation is a term that has many different meanings: assessing the quality of a course, effectiveness of materials used, teaching style of an instructor, or the comfort of a classroom. An evaluation can be done informally over lunch or with highly structured data-gathering tools. It can produce results that are useful to trainers, program administrators, corporate decision-makers, or no one. The key to a worthwhile evaluation is clearly defining why the evaluation is being conducted. Once the purpose is defined, planning the evaluation strategy can begin. For an evaluation to be effective, it should be incorporated into the development of the training activity itself. The training evaluation worksheet at the end of this paper can be used to guide evaluation planning.

## EVALUATION PLANNING

### Planning a Training Evaluation

1. Who is interested in the evaluation results?
2. What questions do they want answered?
3. What resources are available?

The first step to developing an evaluation strategy is determining who will be interested in the results. Assessment of a typical mine safety or health training course could be important to the trainer, company managers, a labor organization, state or federal government agencies, or others. An evaluation designed only to meet the Mine Safety and Health Administration’s (MSHA) information needs may not contain results that company managers could use for future planning. One developed on the basis of a request from company managers may not help a trainer decide if the course is working. It usually isn’t practical to gather enough information to meet everyone’s information

needs, so the target audience for the evaluation must be clearly defined.

After identifying who will use the evaluation results, the next step is to determine what the interested parties hope to learn from the evaluation. The kinds of decisions to be made based on the results of the evaluation should guide the evaluation design. Is information wanted about the quality of the instructor, the usefulness of training materials, the appropriateness of the topics covered, the achievement of the students, or something else? How the results will be used should also be considered. Will the course, the trainer, or the training location be changed on the basis of the evaluation findings? Will the findings be used to convince someone of the value of the training?

An important determinant of the scope of training evaluation is the resources available. Some evaluation strategies are more resource intensive than others. The availability of personnel, time, dollars, equipment, and access to data will affect what can be accomplished (Dopyera and Pitone, 1987, p. 74). A balance between what would be interesting to know and what is practical to discover may have to be found. It is not resource efficient to gather and analyze more data than are required; neither is it productive to conduct an evaluation that fails to provide needed information. This is why it is important to define the reasons the evaluation is being conducted clearly before designing the strategy to be used. The questions in the box help focus the initial stages of evaluation planning.

<sup>1</sup>Sociologist.

<sup>2</sup>Sociologist, Pittsburgh Research Laboratory, National Institute for Occupational Safety and Health, Pittsburgh, PA.

## KIRKPATRICK'S EVALUATION MODEL

One way to think about what can be learned during an evaluation is to use the categories developed by Kirkpatrick. His four-level framework was first presented in a series of articles published around 1960 (Nichols, 2000) and go from level 1 (the easiest and least resource-intensive) to level 4 (the most difficult and expensive) (Kirkpatrick, 2001, pp. 122-132). Each level is described in table 1 and the text below. Over 40 years later, this framework is still being used to structure evaluations of training programs.

“As we move from level 1 to level 4, the evaluation process becomes more difficult and time-consuming, although it provides information of increasingly significant results” (Clark, 1997). The questions the evaluation needs to answer and the resources available for the task should determine which levels will be included.

- Level 1: Trainee reactions are the easiest kind of assessment data to gather. This is not to say they are not important. If trainees do not see value in the training, they are not likely to translate the objectives of the course into useful knowledge and skills. When trainees find a course uninteresting, they will be less motivated to learn the material being covered. Furthermore, quality instruction will be wasted in a training environment that is not conducive to learning. A classroom that is too hot, cold, noisy, or small can defeat the purpose of the class before it begins. While positive trainee reactions do not ensure that objectives are met, negative reactions guarantee a less-than-fully-successful transfer of knowledge and skills.

- Level 2: Measurements of learning are used to show whether trainees' knowledge and/or skills are changed by training. The best way to determine if changes are the result of specific training is to conduct an experiment in which the class is divided into two similar subgroups. Prior to training, both subgroups can be tested on the topics of interest either in writing or through observation. During this pretest, both groups should perform equally. Then only one of the subgroups is trained. After training, both subgroups are retested. If the trained group now performs better than the untrained group, the training can be identified as the cause of the improvement.

Often, however, it is not practical to leave a subgroup untrained, especially with regard to their safety and health. Less methodologically rigorous strategies can be used to assure that training is working. When the entire group is undergoing training, testing before and after the course can show any changes in knowledge or skills. While this cannot definitely prove that the

change resulted from the training rather than some other external factor, if another reasonable explanation is not available to account for the improvement, the success of the training can be inferred.

- Level 3: Unlike levels 1 and 2, measuring a change in behavior must be done outside the classroom and with sufficient time elapsed for knowledge and skills to have been tried out in the workplace. The most elaborate plan for level 3 evaluation would include an untrained subgroup as described for level 2 and detailed testing of both subgroups in their workplaces before and after the training. This type of evaluation is resource intensive and isn't practical for all training sessions. But less intensive strategies can yield valuable results.

[S]omething beats nothing, and I encourage trainers to at least do some evaluation of behavior, even if it isn't elaborate or scientific. Simply ask a few people: "Are you doing anything different on the job because you attended the training program?" If the answer is yes, ask, "Can you briefly describe what you are doing and how it is working out? If you are not doing anything different, can you tell me why? Did you learn anything that you can use on the job?" (Kirkpatrick, 2001, p. 128).

Another strategy is to talk with the trainees' supervisors about any behavioral changes they have observed since the training was completed. Level 3 evaluation can be difficult because it must be conducted months after the training has been completed. This highlights the importance of planning an evaluation strategy when planning the course. Time must be scheduled for the follow-up level 3 data collection so it won't interfere with future training activities and projects.

- Level 4: Determining how training affects the organization is the most difficult evaluation to perform. Level 4 evaluations should be conducted when the value of the training or the training program to the overall organization needs to be assessed. A relatively simple example is measuring changes in sales numbers after training salespeople in a new skill. An increase in sales can be compared to the cost of the training and a bottom-line return on investment calculated. Unfortunately, many topics aren't that easy to quantify. Even the relatively simple sales example can be complicated by a number of other variables. Unless a control group is used, as discussed above, the economics of the region, the introduction of a competing product, or the end of a fad could distort the data to an extent that the impact of training is difficult to calculate.

**Table 1.—Kirkpatrick's four-level evaluation scheme**

Level	Measurement focus	Questions addressed
1 - Reaction	Trainees' perceptions	What did trainees think of this training?
2 - Learning	Knowledge/skills gained	Was there an increase in knowledge or skill level?
3 - Behavior	Worksite implementation	Is new knowledge/skill being used on the job?
4 - Results	Impact on organization	What effect did the training have on the organization?

To conduct a level 4 evaluation, it is important to define clearly the tangible results to be measured, such as a decrease in accident frequency, an increase in use of personal protective equipment, a reduction in maintenance costs, or an increase in production per shift. Once the desired result is identified, a means to measure changes is needed. Next, factors other than training that could

influence the change should be identified so that they can be ruled out as the source of change, if possible. Finally, evidence that the training did cause the change being studied should be identified. “Most importantly, be satisfied with evidence, because proof is usually impossible to find” (Kirkpatrick, 2001, p. 129).

## GATHERING INFORMATION

It is good to acquire data from a number of sources to obtain a more complete view of how the training is perceived and its impact. There are many kinds and sources of data that can be used to evaluate training. An obvious source is the individuals being trained. Useful information can also be obtained from the supervisors of these trainees. Content or training experts can be asked to review course materials, assist with test development, or critique a training session. Company documents can contain indicators of change, such as maintenance costs, accident frequencies, or number of grievances filed.

The decisions made up to this point in the planning process guide the choice of data collection methods. Data can be obtained in a number of different ways, and table 2 lists some commonly used techniques and data sources. Table 2 also

describes when these methods would be used and what can be learned from each technique.

To be used successfully, each of the data collection methods listed in table 2 requires knowledge about its development and implementation. A more experienced program developer may be needed to assist a trainer with the integration of a particular method into a training course. Some methods, such as interpreting body language, require astute observational skills. Others, such as surveys and tests, do not require formal training to administer, but may require an experienced developer to construct questions that thoroughly address the training program and its effectiveness. The knowledge and skills of evaluation plan developers, trainers, and/or other observers who will gather the data must be considered when determining the best method or methods for gathering evaluation data.

**Table 2.—Training methods**

Method	When used	What can be learned
Questionnaire . . . . .	Before, during, or after training	<ul style="list-style-type: none"> <li>• Perceptions of trainees or supervisors</li> <li>• Opinions of content or training experts</li> <li>• Knowledge or skills</li> <li>• Transfer of training to job</li> <li>• Organizational impact</li> </ul>
Interviews . . . . .	Before, during, or after training	<ul style="list-style-type: none"> <li>• Perceptions of trainees or supervisors</li> <li>• Opinions of content or training experts</li> <li>• Knowledge or skills</li> <li>• Transfer of training to job</li> <li>• Organizational impact</li> </ul>
Facial expressions/body language . . .	During training	<ul style="list-style-type: none"> <li>• Perceptions of trainees</li> </ul>
Performance tests . . . . .	Before, during, or after training	<ul style="list-style-type: none"> <li>• Trainee skills</li> </ul>
Written tests . . . . .	Before, during, or after training	<ul style="list-style-type: none"> <li>• Trainee knowledge</li> </ul>
Workplace observations . . . . .	Before or after training	<ul style="list-style-type: none"> <li>• Trainee knowledge or skills</li> <li>• Transfer of training to job</li> </ul>
Games . . . . .	During training	<ul style="list-style-type: none"> <li>• Trainee knowledge or skills</li> </ul>
Group discussion . . . . .	Before, during, or after training	<ul style="list-style-type: none"> <li>• Perceptions of trainees or supervisors</li> <li>• Opinions of content or training experts</li> <li>• Trainee knowledge or skills</li> <li>• Transfer of training to job</li> </ul>
Analysis of statistics . . . . .	Before or after training	<ul style="list-style-type: none"> <li>• Organizational impact</li> </ul>

## WHAT TRAINING EVALUATION IS NOT

If the only result of interest is whether or not each trainee knows certain things or has acquired certain skills, then an evaluation of training is not needed. In this situation, it doesn't matter how each individual obtained the knowledge or skill, and, therefore, the effectiveness of a specific training activity is not important. A knowledge and/or skills test can be administered to

each person after training is completed to determine who is at an acceptable level of performance. A pretest is not needed unless there is an option that permits trainees to skip the training class if they can pass the test. When the question becomes how to train those who are not performing adequately so they can pass a future test, then evaluation of the training becomes important.

## SUMMARY

The only way to determine whether or not training is of value is to evaluate it. When objectives for the training are clearly defined, an evaluation plan can be designed to measure the training's effectiveness at achieving those goals. Sometimes company managers or outside organizations require evaluation data to assess a training program. Even when such outside influences are not present, it is in the best interests of a trainer to gather evaluation data routinely to assess course content,

delivery methods, and teaching skills. If a course is going to be repeated, evaluation can guide changes to improve future sessions. If the course will not be repeated, evaluation could focus on the skills of the instructor with results being used for professional development of that trainer. The important thing is to decide what can and should be learned during training evaluations and then design a strategy to meet that goal.

## REFERENCES

Clark, D. 1997. Instructional system development: Evaluation phase. Chapter 6. <http://www.nwlink.com/~donclark/hrd/sat6.html>.

Dopyera, J., and L. Pitone. 1987. Decision points in planning the evaluation of training. *More Evaluating Training Programs: A Collection of Articles from Training and Development Journal*, D.L. Kirkpatrick, comp. Alexandria, VA: American Society for Training and Development, pp. 74-77.

Kirkpatrick, D. 2001. The four-level evaluation process. Ch. 12 in *What Smart Trainers Know: The Secrets of Success from the World's Foremost Experts*, L.L. Ukens, ed. San Francisco: Jossey-Bass/Pfeiffer, pp. 122-132.

Nichols, F. 2000. Evaluating training: There is no "cookbook" approach. <http://home.att.net/~nickols/evaluate.htm>.

