

# evaluation doesn't have to be difficult

---

Robert P. Olson  
Marjorie F. Fruin

The need for evaluation is seldom debated. The fact that evaluation doesn't need to rigidly adhere to the experimental model has also been generally accepted. In addition, models incorporating subjective opinions, skills learned, and even the group process have been suggested as more practical approaches to evaluation of program offerings.<sup>1</sup>

## Need for Different Evaluation

Although most program offerings are educational, they usually don't fit the typical classroom situation. This "quasi-educational" situation has been described as being more personalized, more intense, and often involving several teachers or change agents.<sup>2</sup> These programs need a different basis for evaluation.

Despite the fact that evaluation need not be experimental research, and that learning experiences are usually different from the classroom situation, it's still common to find program evaluation based entirely on a change in scores of participants on a pre/post-test. These scores are analyzed to see if there's a "significant difference."

There's usually a difference, for two reasons. First, almost any educational offering will bring about some learning for most people, and their scores will be higher on the post-test. Also, the participants saw the questions on the pre-test, and many are going to listen for the answers during the presentation. Second, unless the test is highly reliable (and often this isn't the case), scores may increase due solely to measurement error.

If the difference between the pre- and post-test scores is significant, that's great. But what does a t-test tell a program planner about what was learned by which participant? What does a higher group average on a post-test really say?

---

*Robert P. Olson:* Extension Specialist, Food Service Management, Department of Food Science and Nutrition, and *Marjorie F. Fruin:* Doctoral Candidate, Educational Psychology, University of Minnesota—St. Paul. Accepted for publication: November, 1978.

All it really says is that some of the participants' scores are higher than others, and the average score is higher than it was a few hours ago. It might be higher for the reasons previously mentioned; it's probably unreliable.

This is regrettable. Other measuring instruments could be developed that would be able to evaluate a unique program offering and furnish more information to the planners and funding agencies.

This article describes two approaches that can be adapted to many types of programs and information on a variety of program characteristics. They can be used in addition to a pre/post-test if desired, provided the evaluation doesn't become too lengthy. The types of information these formats provide are different from the typical knowledge-based pre/post-test, and they could strengthen the insights into what has actually been achieved.

## Measuring Learning

### Activity Items

There are two major points in constructing one of the formats we used to assess the learning of participants at a workshop. First, the evaluation form needs a collection of knowledge or activity items, like those listed on the left in Example 1.

#### Example 1. Measuring learning.

Please read the following statements\* and mark the column that best describes what you feel you know about each item.

(Place an X in <i>Only One</i> column)	Already knew before workshop	Know now	Need more help
Implement a feedback system	_____	_____	_____
Redesign or simplify a job	_____	_____	_____
Conduct a performance review	_____	_____	_____

\*The original questionnaire contained a total of eight statements.

These items can be drawn from program objectives or goals, although in most cases they'll have to be broken down and made more specific. The items can also be obtained from the important points made by the speakers and/or small group leaders in their presentations or problem-solving activities. Sometimes the planning committee can identify topics that could be evaluation items. In any event, the statements should be clear, concise, and preferably deal with a fairly small area of learning.

Because each item is content specific, it's possible to identify where learning occurred, and which areas need more attention. The latter information can be of great help to program planners, and becomes a kind of formative, or ongoing, type of evaluation.

*State of Knowledge*

Second, we need to separate what participants already know about the area when the program started, what they actually found out during the program, and what they felt they still don't understand or can't do. This information is collected in the three columns on the right in Example 1. Respondents must check only *one* of the columns, and this statement should be emphasized, as shown in Example 1.

By examining the comparative frequencies of the "already knew" versus the "know now" versus the "need more help," it's possible to identify workshop effectiveness, whether these participants might desire further information, or if the workshop was presented at too low a level for those present.

If statistical analysis is desirable, nonparametric techniques are appropriate for these kind of data.<sup>3</sup> Many of these have been computerized for analysis of larger samples (for example, the chi-square test); they're essential for samples of 10 or less. Percentages are an effective way to examine results for larger groups.

---

**. . . evaluation doesn't need to rigidly adhere to the experimental model . . . models incorporating subjective opinions, skills learned, and even the group process have been suggested as more practical approaches to evaluation . . . .**

---

**Measuring Behavioral Change**

To identify behavioral change of participants requires a follow-up study after the program ends. Names and addresses must be collected while the program is in session, with an indication that the individual would be willing to be contacted for such a follow-up. This can be included on the registration form if the on-site evaluation is anonymous.

If there's time, money, and a commitment for a follow-up questionnaire, the problem arises as to what kinds of questions will identify change. Again, these questions should be related to the content and goals of the program.

Questions to measure behavioral change must show what the individual does now that he/she didn't do before. For example, suppose that one of the recurring themes in a workshop on motivation of food service employees was

the necessity for communication. Suppose also that the participants were managers of food service employees. This need for communication could take many forms, but it certainly would involve the managers talking to the employees. For this workshop, a behavioral change might center around how often or how much the managers talk with their employees. It could also involve what they talk about.

The items shown in Example 2 were used to evaluate behavioral change following such a workshop.

The respondent was asked to report whether the number of contacts and the length of time of the contacts had increased or decreased since the workshop. They were also asked to check a list of the kinds of things that might have been talked about. We were interested in both individual contacts and group contacts (such as more formal meetings), since both would show a change in behavior on the part of the manager.

**Example 2. Measuring behavioral change.**

---

During the past six months, the FREQUENCY of my contacts and the DURATION of my contacts with (individuals, groups)\* in my work unit has:

(Place an X opposite one phrase in EACH column)

<u>FREQUENCY</u>		<u>DURATION</u>
a. _____	increased a great deal	a. _____
b. _____	increased moderately	b. _____
c. _____	not changed	c. _____
d. _____	decreased some	d. _____
e. _____	decreased a great deal	e. _____

During the past six months, the following\*\* has been *newly* incorporated into my contacts with the work unit group or individuals in the group.

(Place an X in front of ALL that apply)

- a. \_\_\_\_\_ nothing
- b. \_\_\_\_\_ shared my personal problems/frustrations/joys with employees
- c. \_\_\_\_\_ greeted each individual personally at start of day
- d. \_\_\_\_\_ learned about employees' job-related problems
- e. \_\_\_\_\_ involved employees in problem solving

\* Respondents filled out duplicate sets of items, one asking about groups and one asking about individuals.

\*\* The original questionnaire contained a total of 17 items.

---

Of course, respondents may not tell the whole truth and nothing but the truth, but it's definitely not practical in most cases to send out trained observers to assess behavioral change, both before and after a program. This is one reason why assessment of behavioral change is seldom tried.

One of the main points of this example is that the workshop dealt with motivating employees, and we chose to measure the manager's behavioral change in only one area of motivation: communication. We could have chosen others. This is a decision to be made by program planners, supervisors, and the like. In this case, communication was one of the areas that had been stressed in the workshop. Note that the items used to measure behavioral change didn't ask the participants whether they'd been trying to motivate their employees. Rather, the questions asked for specific actions taken by the manager that demonstrated use of information presented at the workshop.

## Summary

There's always a place in program evaluation for subjective opinions, that is, liking the offerings, finding speakers interesting, or feeling the audio visuals were helpful. These types of questions help to identify "distractors" from a program's effectiveness. They don't, however, tell you if the participants learned anything.

The pre/post-test can sometimes identify learning, but the information it supplies is limited.

Two additional formats have been suggested. One is a different approach to measuring what was learned, and the other a mail questionnaire to assess changes in behavior following a program. Either or both could be used in conjunction with subjective opinions or pre/post-testing, depending on what kinds of information the evaluation is expected to furnish. Either or both can strengthen the program evaluation.

## Footnotes

1. David Logsdon, "A Practical Look at Evaluation," *Journal of Extension*, XIII (March/April, 1975), 31-38.
2. Gerald G. Udell, "Yes, a Change Agent *Can* Evaluate!" *Journal of Extension*, XIII (September/October, 1975), 14-21.
3. Sidney Siegel, *Nonparametric Statistics for the Behavioral Sciences* (New York: McGraw-Hill, 1956).