Level 1 Evaluations: Do They Have a Role in Organization Learning Evaluation Strategy?

by: Ken Phillips

PHILLIPS ASSOCIATES

Level 1 Evaluations: Do They Have a Role in Organizational Learning Strategy?

By Kenneth R. Phillips

   President, Phillips Associates

> *"When designed well and used properly, a Level 1 evaluation can produce valuable data regarding whether or not a learning program is on track to meet key business objectives plus serve as an early warning to issues that might undermine program effectiveness."*

Level 1 evaluations are a long-time staple of the Workplace Learning and Performance (WLP) field and are used extensively. According to a 2009 ASTD research study titled "The Value of Evaluation: Making Training Evaluations More Effective" 92% of organizations evaluate some learning programs at Level 1 of the Kirkpatrick/Phillips five-level evaluation model (Reaction, Learning, Behavior Change, Business Results and ROI). However, when asked about the value Level 1 evaluation data has for their organization, only 36% of study respondents indicated it had high or very high value. This raises the question: "If so few organizations find Level 1 evaluation data to be of value, should they be included as part of an organization's learning evaluation strategy?" The answer is a definite "Yes, but…"

When designed well and used properly, a Level 1 evaluation can produce valuable data regarding whether or not a learning program is on track to meet key business objectives plus serve as an early warning to issues that might undermine program effectiveness. Unfortunately, as evidenced above, most Level 1 evaluations aren't designed well and thus produce data that has little perceived value. This raises a second question: "What's the difference between a Level 1 evaluation that produces valued data and one that doesn't?" The answer, quite simply, is careful and informed design. However, the unfortunate truth is that many learning and development professionals are uninformed in the art and science of evaluation design. While many advances in the field of evaluation design have been made in recent years, few of these have found their way into the hands of learning professionals who continue to follow design principles formulated 50 or more years ago, and since outdated.

In this article, we'll examine four different questions to include on a Level 1 evaluation that produce data with high perceived value for both learning and development professionals and business executive stakeholders alike, and will earn Level 1 evaluations the right to be included in an organization's learning strategy. Does it seem too good to be true? It's not.

"Predictive questions forecast the likely outcomes to be achieved by a learning program and begin to answer the questions business executive stakeholders have about the value of their learning investment."

Traditionally, Level 1 evaluations are used to collect data around three major topic areas: the quality of the learning program; the effectiveness of the facilitator, in the case of classroom and synchronous online programs; and the conduciveness of the learning environment, in the case of classroom training. While information about each of these topics has value for us as learning and development professionals, they are of little interest or value to our business executive stakeholders. They want to know whether people who attended a learning program are performing better on the job and whether business results have improved. In short, they want know if they are getting value from their learning investment. However, collecting data that addresses this issue requires asking different questions from those found on the traditional Level 1 evaluation form. Two examples are asking predictive questions that forecast participant learning, intent to apply what was learned back-on-the-job and likely impact on business results, and asking learning program relevancy questions.

Predictive questions forecast likely outcomes

Predictive questions forecast the likely outcomes to be achieved by a learning program and begin to answer the questions business executive stakeholders have about the value of their learning investment. While the data collected from predictive questions isn't proof that certain program outcomes are going to be achieved, it is a best estimate prediction of the outcomes. A familiar analogy is the local weather service predicting the track and intensity of an upcoming storm based on various computer forecasting models. And as we all can attest, these predictions aren't always correct, but they are often enough so that we consider weather forecasts to be credible. The same holds true in learning – the data we collect from predictive questions on a Level 1 evaluation may not always result in an accurate forecast, but it will often enough to be viewed as credible by our business executive stakeholders.

"By computing the difference between participant knowledge before and after attending the program, we can forecast whether or not learning took place."

An example of a predictive question used to forecast whether or not participant learning occurred during a training program is to create parallel Likert scale items asking participants to indicate how much knowledge they had about the material taught both before and after attending the program (See the example below). By computing the difference between participant knowledge before and after attending the program, we can forecast whether or not learning took place. Of course, we can't say with complete certainty that participant learning occurred, assuming the difference score is positive, because we haven't measured actual learning. However, we can say that all signs point in that direction. In addition, if the data suggest that learning didn't take place – there is no positive difference score, it serves as an early warning that either the learning program needs some adjustment or the wrong participants are attending the program.

Example:

How much did you know about developing individual performance objectives *before* attending this seminar?

No Knowledge   Thorough Knowledge

1	2	3	4	5	6	7

How much do you know about developing individual performance objectives *after* attending this seminar?

No Knowledge   Thorough Knowledge

1	2	3	4	5	6	7

An example of a predictive question used to forecast participant on-the-job behavior change is to create two items asking participants how likely they are to apply back on the job what was learned in a program, and what obstacles if any might prevent them from applying what they learned (See the example below). Obtaining a high score on the first question combined with either few or no obstacles identified in the second question enables us to forecast that there is a high likelihood participants are going to apply what they learned.

Again we can't say this with complete certainty because we haven't measured participant actual on the job behavior change, but the needle is pointing in the right direction.

On the other hand, obtaining a low score on question one combined with a number of obstacles identified in question two is compelling evidence that the learning program is in jeopardy of achieving its intended objectives. To pinpoint the source of the problem, the obstacles should be summarized into groups of like-minded items and then analyzed for patterns or themes. These patterns or themes will identify the specific problems associated with the program and provide a starting point for taking corrective action. It's also likely that some, if not most, of these problems are going to focus on participant work environment issues such as lack of support for using what was learned or not having an opportunity to apply what was learned. Moreover, because work environment issues fall under the jurisdiction of business executive stakeholders, you'll need their input and commitment to mitigate or eliminate them. An effective way to gain this commitment is to present the stakeholder with a "business case" created from the quantitative data collected with question one and from the themes and patterns identified from summarizing the obstacles in question two. Having a business case to present a stakeholder places you in a stronger position to gain his or her support and commitment for solving work environment issues.

Example:

How likely are you to use the skills and behaviors you learned in this seminar back on the job?

Not at all Likely   Very Likely

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

What obstacles, if any, are likely to prevent you from applying these skills and behaviors back on the job? _____

An example of a predictive question used to forecast a learning program's likely impact on business results is to create two items asking participants how likely any of the key business

"An effective way to gain this commitment is to present the stakeholder with a "business case" created from the quantitative data collected with question one and from the themes and patterns identified from summarizing the obstacles in question two."

metrics (financial, operational or HR) tracked by their department are to improve as a result of them applying what they learned in the program, and how confident they are in their response (See the example below). Receiving a high score on the first question, while encouraging, is also likely to be biased by the natural optimism many participants feel immediately after attending a learning program. The second question corrects for this by asking participants to indicate how confident they are in their response to the first question. Multiplying the response choice on question one by the confidence percentage from question two results in a more conservative figure or error adjustment. Summarizing the adjusted numbers and dividing by the number of participants results in a single number that is a forecast of the extent to which the learning program is likely to improve business results. A high average adjusted score forecasts a likely improvement in a key business metric and a low average adjusted score serves as an early warning that participants either don't see how the learning program connects to the business metrics tracked by their department, or they see the connection, but don't think the program is going to have any affect on the metrics. While these are quite different issues with each requiring a different solution, the program is in jeopardy until they are resolved. *Note:* if you know the particular business metric the learning program is connected to (e.g. employee turnover), question one should be made more specific. Also, these two questions should not be included on a Level 1 evaluation if the program content doesn't have a clear connection to any of the business metrics being tracked by the department or organization.

Example:

How likely to improve are any of the key business metrics tracked by your department as a result of you applying the knowledge and skills you learned in this program?

Not at all Likely   Extremely Likely

| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

How confident are you in your response to the previous question where 0 = no confidence and 100 = extremely confident? _____

> *"Summarizing the adjusted numbers and dividing by the number of participants results in a single number that is a forecast of the extent to which the learning program is likely to improve business results."*

Relevancy questions predict learning

A fourth question to add to a Level 1 evaluation to ensure it has strategic importance is to include at least one item asking participants how relevant the learning program was to them and their job (See the example below). According to research conducted by Neil Rackham, author of *SPIN Selling and Major Account Sales Strategy*, and reported in *Training* magazine, relevancy questions have a strong positive correlation with participant learning. In fact, according to Rackham, they have a higher correlation with learning than an item actually evaluating participant learning. Relevancy questions also can be used to evaluate specific topics in a multi-topic program such as in leadership training. Again, while a high relevancy score isn't proof that participant learning occurred, when it is combined with a high positive difference score from the predictive question forecasting participant learning discussed earlier, a highly credible forecast can be made regarding participant learning. On the other hand, low scores on both questions provide an early warning that the program is in jeopardy and corrective action is required.

Example:

How would you rate the overall relevance of this session to you and your job?

Not at all Relevant   Very Relevant

1	2	3	4	5	6	7

In summary, Level 1 evaluations, while ubiquitous, often miss the mark with business executives because they traditionally capture data that is of little value or interest to them. However, it doesn't have to be this way. Including predictive and relevancy questions such as those described above can place Level 1 evaluations and the data they collect on the same level as other learning evaluation data that is viewed as strategically important. And after all, if you're going to spend time and effort capturing Level 1 evaluation data, shouldn't it be regarded as worthy of being included in an organization's learning evaluation strategy?

"According to research conducted by Neil Rackham, author of SPIN Selling and Major Account Sales Strategy, and reported in Training magazine, relevancy questions have a strong positive correlation with participant learning."

For more information, contact:
Ken Phillips
Phillips Associates
34137 N. Wooded Glen Drive
Grayslake, IL 60030
(847) 231-6068
www.phillipsassociates.com
ken@phillipsassociates.com

Phillips Associates