Phillips Associates
*on* M&E
*2nd in the series*

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

"When designed well and used correctly, a Level 1 evaluation can produce valuable data regarding whether or not a learning program is on track to meet key business objectives…"

Level 1 evaluations, a long-time staple of the Learning and Development (L&D) field, are used extensively. According to a 2015 ATD research study titled "Evaluating Learning: Getting to Measurements That Matter" 88% of organizations evaluate some learning programs at Level 1 of the 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 44% of study respondents indicated it had high or very high value. This low percentage 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 correctly, 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, which 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 frankly, 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 occurred 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the hands of learning professionals who continue to follow design principles formulated 50 or more years ago, and since outdated.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, I'll examine four different questions that if used on a Level 1 evaluation will produce data with high perceived value for both L&D professionals and business executives alike, and will earn Level 1 evaluations the right to be included in an organization's learning strategy. Sound too good to be true? It's not.

Phillips Associates

"…business executives… want to know whether people who attended a learning program are are applying what they learned and whether business results have improved."

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-based and synchronous online sessions; and the conduciveness of the learning environment, in the case of classroom training. While information about each of these topics has value for L&D professionals, they are of little interest or importance to business executives. They want to know whether people who attended a learning program are applying what they learned and whether business results have improved. In short, they want to see if they are getting value from their learning investment. However, collecting data that address 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

"While the data collected from predictive questions isn't proof that specific program outcomes are inevitable, it is a prediction that the results are likely."

Predictive questions forecast the anticipated results to be achieved by a learning program and begin to answer the questions business executives have about the value of their learning investment. While the data collected from predictive questions isn't proof that specific program outcomes are inevitable, it is a prediction that the results are likely. A familiar analogy is the local weather service predicting the track and intensity of an upcoming storm based on various computer forecasting models. Moreover, 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 is 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 business executives.

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 in participant

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

knowledge before and after the program, you can forecast whether or not learning took place. Of course, you can't say with complete certainty that participant learning occurred, assuming the difference score is positive because you haven't measured whether actual learning occurred. However, you can say that all signs point in that direction. Also, if the data suggest that learning didn't take place – there is no positive difference score -- it serves as an early warning that 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 they learned in the program, and what obstacles if any might prevent them from using 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 you to predict that there is a high likelihood the participants are going to apply what they learned. Again you can't say this with complete certainty because you haven't measured participant actual on the job behavior change, but the needle is pointing in the right direction.

On the other hand, a low score on question one combined with a host of obstacles identified in the second question is compelling evidence that

"Building a credible business case to present places you in a stronger position to gain executive support and commitment to solving the work environment issues."

the learning program is in jeopardy of achieving its intended objectives. To make sense of the training obstacles captured, summarize them into groups of like-minded items and then analyze the groups for themes or patterns. This analysis will help pinpoint 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 the obstacles, are going to focus on 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 the business executives you are supporting, you can't change them on your own; you'll need their input and commitment to mitigate or eliminate them. An effective way to gain this commitment is to present the executive with a "business case" created from the quantitative data collected with question one and from the obstacle themes and patterns identified in the second question. Building a credible business case to present places you in a stronger position to gain executive support and commitment to solving the work environment issues.

Example:

| NOT AT ALL LIKELY | | | | | | VERY LIKELY |
|---|---|---|---|---|---|---|
| 1 | 2 | 3 | 4 | 5 | 6 | 7 |

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

*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 metrics (financial, operational or HR) tracked by their department are to improve as a result of them applying

"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. "

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 the second question 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 forecasts the extent to which the learning program is likely to improve business results.

A high average adjusted score forecasts an expected improvement in a critical business metric whereas a low average adjusted rating serves as an early warning that participants don't see a learning program business metric connection, or they see the link but don't think the program is going to have any effect on the metrics. While these are entirely different issues with each requiring a different solution, until resolved, the program is in jeopardy. (Note: if there is a particular business metric the learning program is targeted to address (e.g., employee turnover), this information also should be included in question one.) 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 department business metrics.

Example:

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

| NOT AT ALL LIKELY | | | | | | VERY 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?* _____

"…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?"

## 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 relationship with learning than an item written to measure participant learning.

Relevancy questions also can be used to assess 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 combined with a high positive difference score from the predictive question forecasting participant learning discussed earlier, a highly credible forecast regarding participant learning is possible. 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 strategically important level as other learning evaluation data. 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?

202006020

# Share Phillips Associates *with your colleagues*


## Corporate Workshops

OFFER THESE WORKSHOPS TO YOUR ENTIRE LEARNING & DEVELOPMENT TEAM

### Mastering M&E
2-Day Workshop

Provide your L&D team with the latest guidelines and hands-on techniques for creating valid, scientifically sound Level 1, 2, 3, and 4 evaluations that produce data perceived by business executives as both credible and valuable.

### Boost Training Transfer using Predictive Learning Analytics
2-Day Workshop

Equip your L&D team with a systematic, credible and repeatable process for maximizing the value of your learning investments by boosting training transfer.

### Crack the Code of Test Question Design
1-Day Workshop

Equip your L&D team with practical tips and specific techniques for creating quizzes and tests that actually measure something.

### Survey Magic: Capturing Level 3 Evaluation Data
1-Day Workshop

Equip your L&D team with a five-step process for creating Level 3 surveys that capture on the job behavior change.

## Presentations for Professional Meetings & Industry Events

IDEAL FOR CORPORATE L&D TEAMS AND INTERNAL LEARNING CONFERENCES

Ken Phillips is available to present on the following topics. All include the valuable, "how-to" tips and hands-on measurement and evaluation techniques that L&D professionals crave—and can't find anywhere else! All topics can be delivered as 75-90 minute programs or webinars.

- **Power up your Level 1 Evaluations** and Gain Surprisingly Useful, Valued Data
- **Take Your Level 2s Up a Notch:** The Magic of Well-written Multiple Choice Test Questions
- **Capture Elusive Level 3 Data:** The Secrets of Survey Design
- **Business Results Made Visible:** Design Proof Positive Level 4 Evaluations
- **Boost Training Transfer Using Predictive Learning Analytics™ (PLA)**
- **Going The Distance: Making Sense Out of Level 1–4 Evaluation Data**

**Contact Ken Phillips at 847.231.6068 or ken@phillipsassociates.com**

**Ken Phillips, CPLP,** delivers all programs and workshops in his signature style: professional, engaging, and approachable.

Ken is founder and CEO of Phillips Associates, and the creator and chief architect of the Predictive Learning Analytics™ (PLA) learning evaluation methodology. He has more than 30 years experience designing learning instruments and assessments and has authored more that a dozen published learning instruments. He regularly speaks to Association for Talent Development (ATD) groups, university classes, and corporate L&D groups. Since 2008, he has spoken at the annual ATD International Conference on topics related to measurement and evaluation of learning.

Phillips Associates

34137 N. Wooded Glen Drive | Grayslake, IL 60030

**847.231.6068 or ken@phillipsassociates.com**