The Level 2 of Training Evaluation

LBAT Model
A Case of Training the HEIs Staff in Pakistan

M. Awab, M&E Specialist, TESP-HEC, Islamabad

Abstract

The case study is about the experience of the TESP M&E Unit of the evaluation of the trainings of HEIs staff in Pakistan. The case study explains how the M&E Unit developed the Learning Based Assessment of Training (LBAT) Model to analyze the *pre-existing learning*, the *net learning* and the *learning lag*. The Model produced surprisingly interesting results for improved accountability and planning for future trainings.

Background

My recent assignment was with the World Bank funded Tertiary Education Support Project (TESP), as Monitoring & Evaluation (M&E) Specialist. One of the important tasks of the Project was to build the capacity of the senior management of higher education institutions (HEIs) in Pakistan. The target was to train 1600 participants, ranging from vice chancellors to rectors, registrars, treasures and directors etc.

The Project started late in 2011. The end of the project was in December, 2015. The trainings could not be kicked off until March 2015. There was hardly any time left for a formal training needs assessment (TNA). The Project team decided to offer the trainings through two implementation partners (IPs), selected from the public sector institutions, having repute and experience in the field of training.

Each IP had to conduct the training on 8 modules such as Leadership and Change Management, Human Resource Management, and Quality Management etc. In view of the big number of the staff members to be trained (800 per IP), for each module the participants were divided into three batches. Each batch consisted of around 35 participants.

Training Evaluation: The SOPs

It was decided that the trainings will be evaluated on the first two levels: i.e. 'Reaction' and 'Learning', using the Kirkpatrick 4-level Model of training evaluation. As for the remaining two levels, the project life did not allow sufficient time.

The IPs were supposed to design the pre & post-training tests to assess the learning of the participants. The tests were to be shared with the M&E team of the TESP so that they could ensure that the tests were designed in accordance with the learning objectives. Once the tests were approved by the M&E team of the TESP, the IPs could use them for evaluation. The results of the test were to be made a part of the training report. The payment to the IP was conditional to the acceptance of the report.

The Challenge

What the IPs submitted as the analyses of the pre & post-training tests was found to be vague and telling only half of the truth. The IP-1 submitted the following analysis:

Figure 1: Pre & Post-Tests Analysis by IP-1
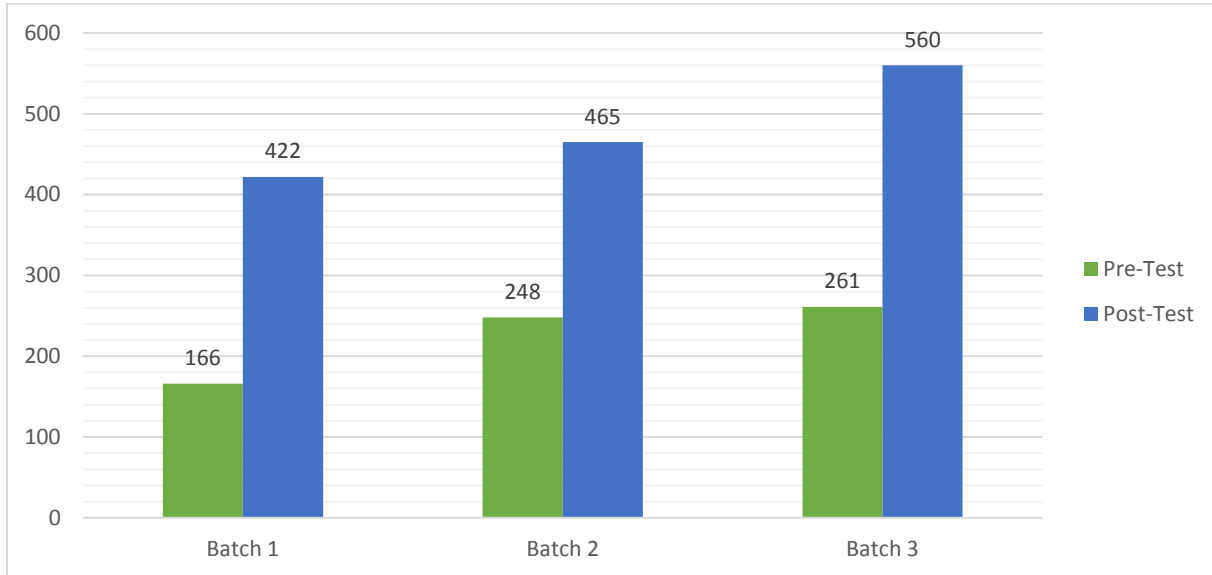
The IP-1 administered a 25 marks test and summed up the total scores of the participants in each batch and compared the scores of pre-test with the post-tests. The results looked more than good. In explanation of the chart, the IP-1 wrote: "…the figure show[s] an aggregate increase of 154.21%, 87.5% and 114.56% in the marks scored by participants in the 1st, 2nd and 3rd batch respectively."

The IP-2 also came up with a similar design to present the pre & post-test analysis.

Figure 2: Pre & Post-Tests Analysis by IP-2
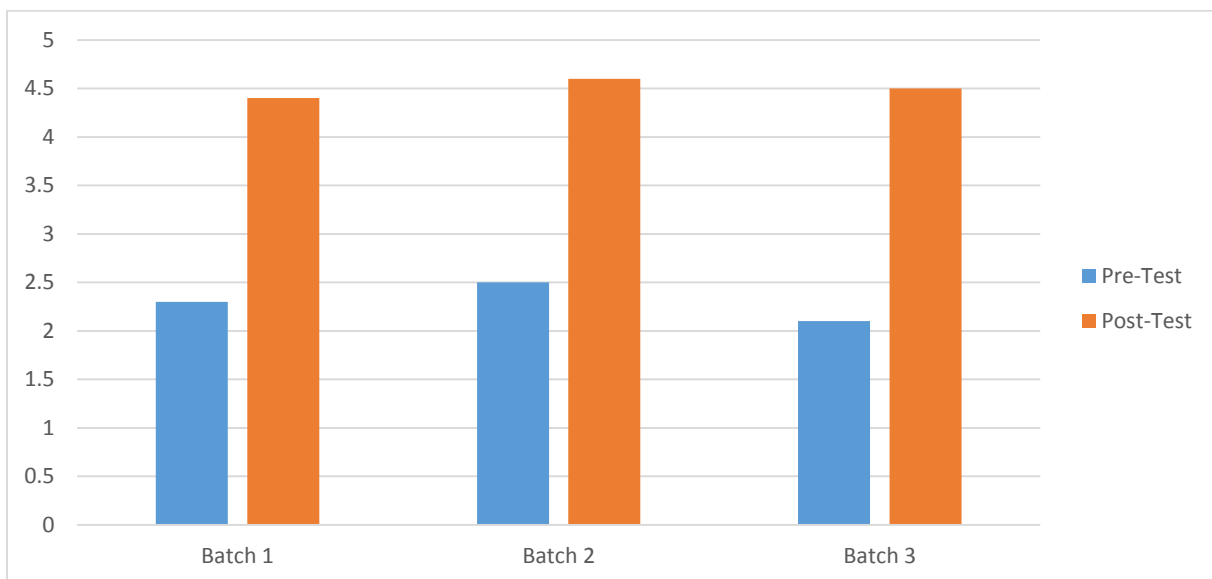
In explanation of the chart, the IP-2 wrote: "…the correct answers ratio stood at 2.3 in the first batch, 2.5 in the second batch and 2.1 in third batch. However, the post training results showed remarkable improvements that is 4.4 in first batch, 4.6 in second batch and 4.5 in the third batch respectively."

From the analysis of pre & post-tests, we expected to find out the answers to the following questions:

1. Keeping the learning objectives of the training in mind, what was the level of **pre-existing knowledge** of the participants? Simply put, how much of the training contents they knew already?
2. What was the **net learning** of the participants? How much of the learning could be attributed to the IP?
3. Was there a **learning lag**? If yes, how much? How much more the participants needed to learn to fully achieve the learning objectives?

In both the analyses presented above, answers to some of the above-cited questions were either vague or missing.

Learning Based Assessment of Training (LBAT) Model

The existing analyses of pre & post-tests did not provide a clear picture of the pre-existing learning, the net learning and the learning lag. This made us to develop a model that could clearly articulate these three areas of learning.

First of all, the three assessment criteria and the methods of their measurement were defined in the following manner:

1. **Pre-Existing Learning:** Pre-test Score.
2. **Net Learning:** Post-test score minus pre-test score. This is the score attributable to the training.
3. **Learning Lag:** 100 minus post-test score. This shows the lag between the total learning and the desired learning.

Note: All the above are calculated in terms of aggregate percentages.

Presentation of Results

The calculations could best be presented in the form of stacked bars. The three learning areas could be color coded; orange for pre-existing learning, green for net learning and red for learning lag.

The Model was shared with the IPs and the results were surprisingly interesting. The following is one of the analysis presented by an IP on the LBAT Model:

Figure 3: Pre & Post-Tests Analysis on LBAT Model

	Batch 1	Batch 2	Batch 3
■ Learning Lag	29.65%	30.36%	21.43%
■ Net Learning	40.48%	28.79%	42.86%
■ Pre-Existing Learning	29.87%	40.85%	35.71%
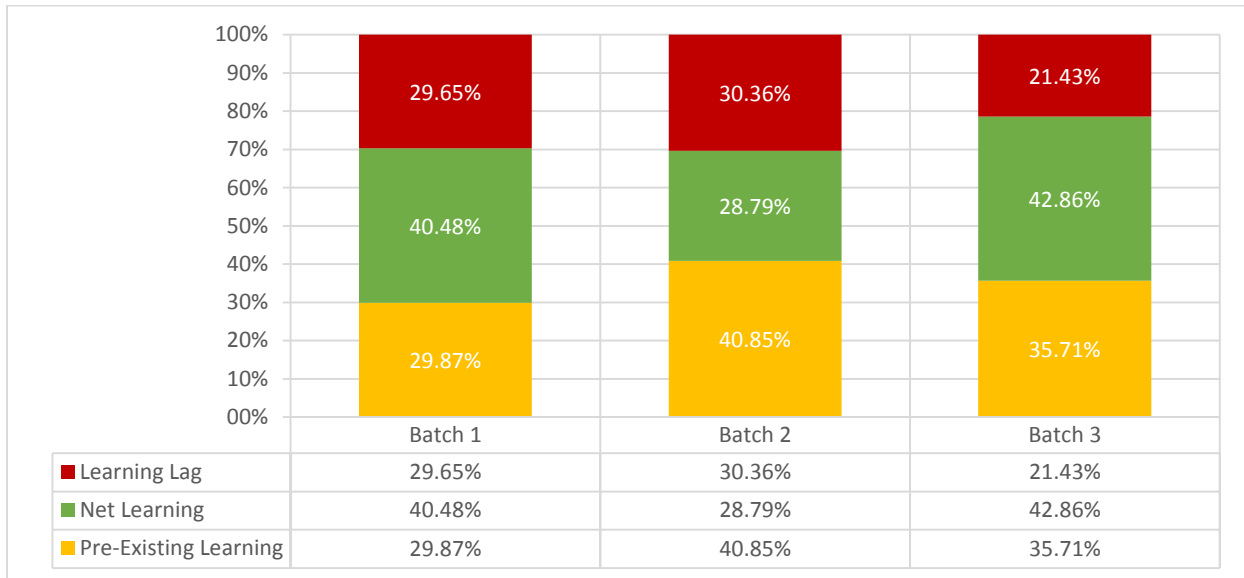
Interpreting the LBAT Model through Green Band

The green band, which signifies the net learning in the LBAT Model, has a central role in describing the success or failure of the training. There are two key things to keep in mind while looking at the green band; its position and its health. Both are explained in the following.

The position of the green band

Generally, the bottom of the green band, on the spectrum of 0-100 percent, tells us how familiar the participants were with the training contents. The top of the green band informs us how difficult it was for the participants to learn from the training.

The ideal position for the green band is to rest its bottom above 10-30 percent. If the green band bottom is below 10 percent, that would mean that the training contents were new and unfamiliar to the participants. And if the green band bottom is above 30 percent, that may mean that the training contents were already too familiar to the participants.

The health of the green band

The second interesting thing about the LBAT Model is the health of the green band. The fatter the better! The benchmark in this regard is 50%. Anything less than that would mean a fatter orange band or a fatter red band. The orange band fatter than 30% is not good. And for red band, any fats are absolutely bad!

Some Salient Feature of the LBAT Model
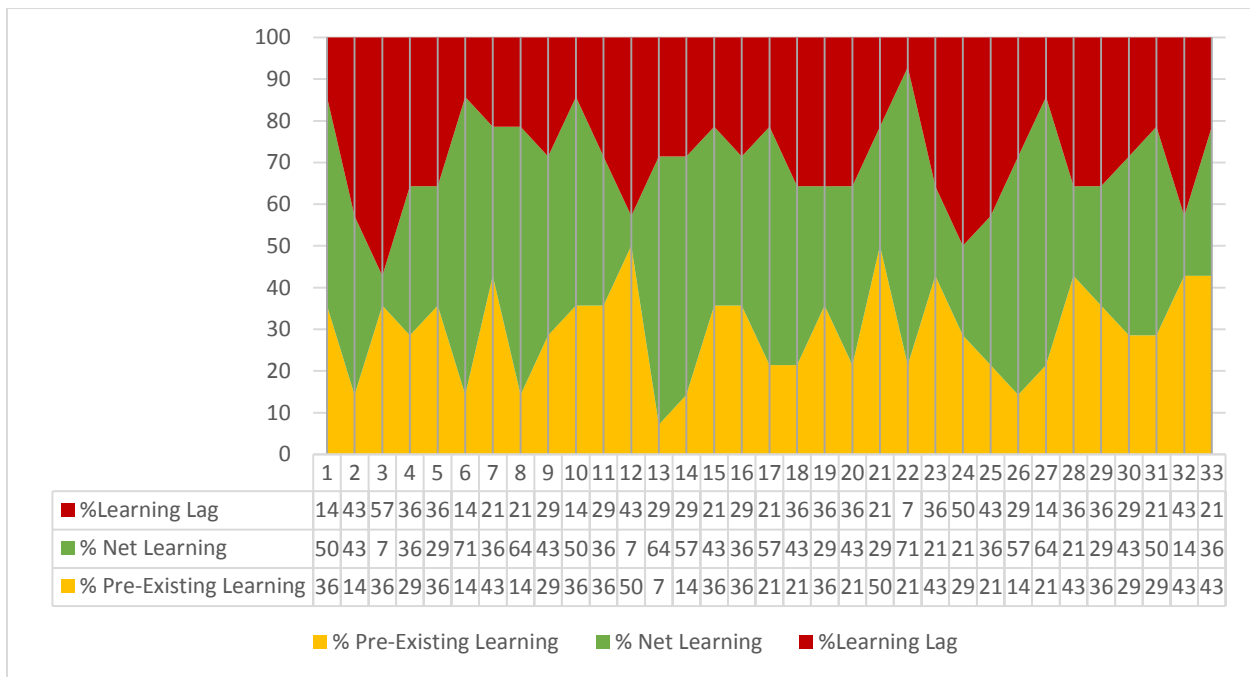
i-   Cream in; Cream Out

The LBAT Model only performs quantitative and visual analysis of the data fed into it, regardless of its quality. Cream in; cream out. If the data is collected through a test valid enough to measure

the learning of the participants against the learning objectives, it will produce the best results. If the test is invalid, the analysis too will be invalid. Similarly, some training participants may fill up the test carelessly. A normality check may help to identify the outliers. The outliers should be removed from the data to reach the best results.
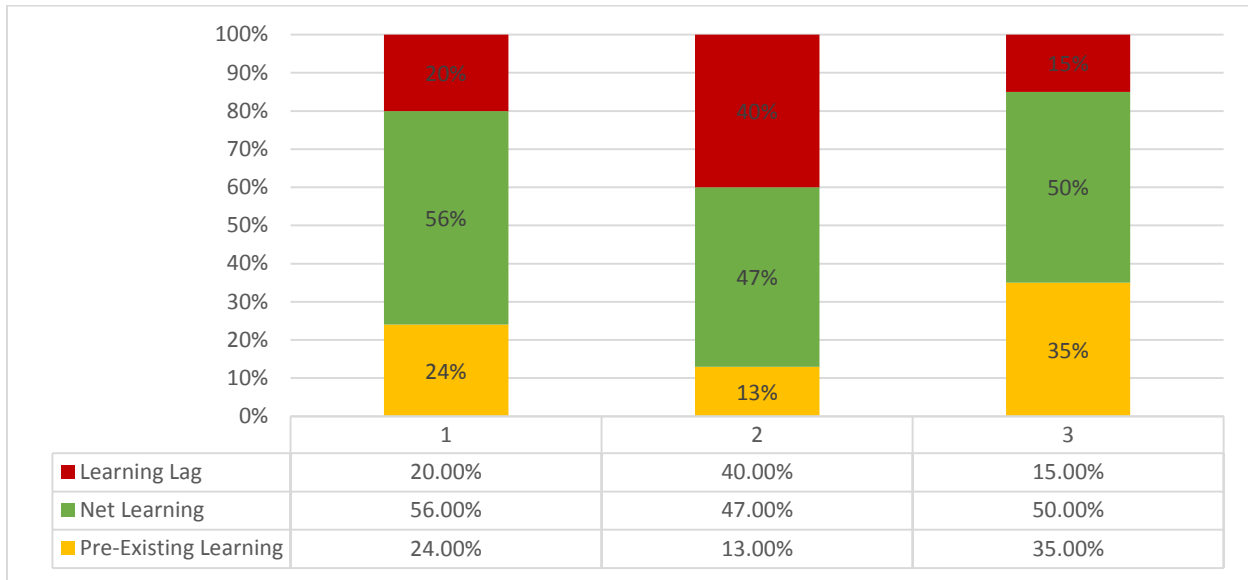
ii- Flexibility in Analysis

The Model is flexible enough to accommodate different levels of analyses, comparing a training batch with another or comparing individual participants with each other or comparing the learning lag in one area with another area etc. In the example illustrated above in figure 3, the unit of analysis is a training batch, compared with other batches. The unit of analysis can be changed to individuals in a single batch, and their levels of pre-existing learning, net learning and learning lags can be compared, as shown in figure 4.

Figure 4: Comparison of Individual Learning in a Training Batch

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32	33
%Learning Lag	14	43	57	36	36	14	21	21	29	14	29	43	29	29	21	29	21	36	36	36	21	7	36	50	43	29	14	36	36	29	21	43	21
% Net Learning	50	43	7	36	29	71	36	64	43	50	36	7	64	57	43	36	57	43	29	43	29	71	21	21	36	57	64	21	29	43	50	14	36
% Pre-Existing Learning	36	14	36	29	36	14	43	14	29	36	36	50	7	14	36	36	21	21	36	21	50	21	43	29	21	14	21	43	36	29	29	43	43

■ % Pre-Existing Learning   ■ % Net Learning   ■ %Learning Lag

Similarly, the LBAT Model can inform us about the area where the learning lag lies the most. As shown in figure 5, the participants' scores against different learning objectives are visualized. Here the scores of the training participants have been compared in three different areas of training on Project Proposal Writing i.e. (1) Results Framework, (2) Developing Indicators and (3) Writing the Financial Proposal. The analysis reveals that the most learning lag is in the area of *Developing Indicators*.

Figure 5: Comparison of Learning Areas in a Training

	1	2	3
■ Learning Lag	20.00%	40.00%	15.00%
■ Net Learning	56.00%	47.00%	50.00%
■ Pre-Existing Learning	24.00%	13.00%	35.00%

iii-   Attribution and Appreciation

The LBAT Model calculates the net learning that is attributable to the training. Thus the trainer can be appreciated for his contribution in the learning of the participants.

iv-   Accountability

The Model also calculates the learning lag between the total learning and the desired learning. Thus, the trainer can be held accountable for not being able to fully impart the desired learning objectives.

v-   Planning for Improved Quality

The results of the LBAT Model, can also help in planning for future trainings. If the red band, representing the learning lag, is too broad, further investigations could be made to reduce it by changing the teaching methodology, the curriculum or the trainer. Similarly, if the orange band, signifying the pre-existing learning is too broad, maybe the difficulty level of the training contents needs to be reviewed and revised. Further detail of possible meanings of variation in different bands is given below.

Figure 6: Variation in Different Bands and their Possible Meanings

Bands	Too Broad	Too Narrow
Pre-Existing Learning	• The training contents were already too familiar to the participants.	• The training contents were rather new and unfamiliar to the participants.
Net Learning	• The training contents were too easy to grasp. • The trainer was successful in imparting the learning.	• The training contents were too difficult to grasp. • The trainer failed to impart the intended learning.
Learning Lag	• The training contents were too difficult to grasp. • The trainer failed to impart the intended learning.	• The training contents were too easy to grasp. • The trainer was successful in imparting the learning.

It is important to note here that different levels of the health of the bands may have different stories to tell about the training contents and the trainer. However, none of these stories can be taken for sure. This is because these stories are not meant to provide definite answers. Rather, they raise questions for further investigation and looking into both the training contents and the trainer. And if both are alright, maybe there is a need to look into the test or the data fed into the analysis. Sometimes, the test may not be valid or the data may contain outliers, distorting the whole analysis.