UPSIDE LEARNING

Position Paper

Evaluation

Learning experience design requires evaluation, of several sorts. For one, a formative evaluation of the design is necessary to test and refine the approach. Summative evaluation determines whether you can stop iterating. There are equivalents for the learner, and that's a worthy topic too, but here we're talking about evaluating the design.

What's clear is that such evaluation is too rare and infrequently appropriate. Thus, here we explore evaluation: what it is, why it's important, why it's problematic, and a resolution.

What is it?

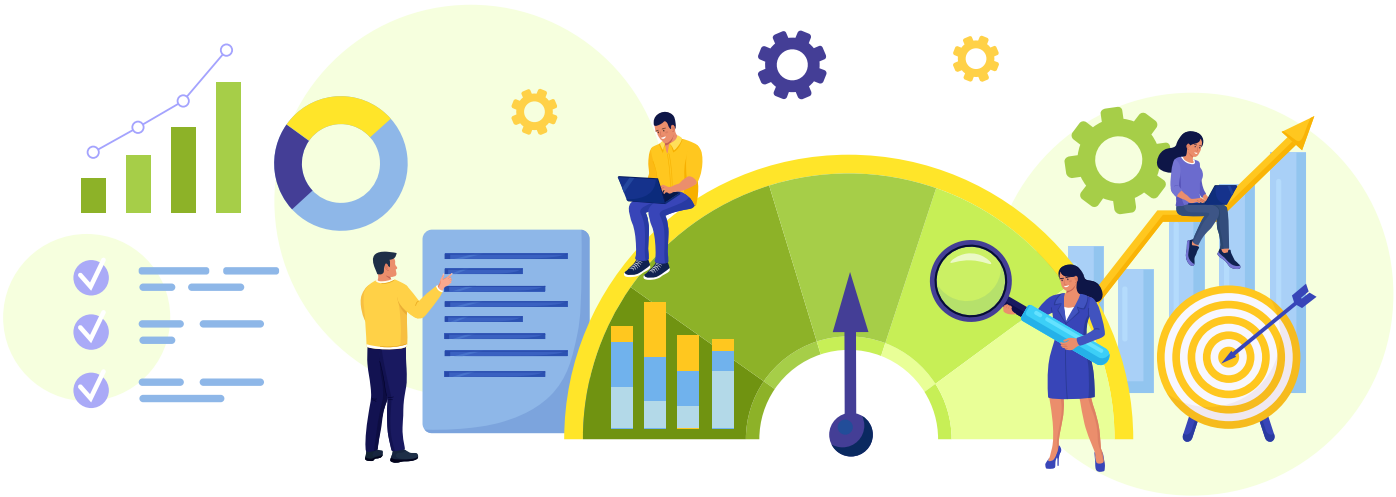
Evaluation is, quite simply, assessing the success of a learning design. We want to determine whether it's achieving our stated purpose. There are a number of things we can, and should, evaluate: the usability, the learning effectiveness, and the engagement.

We can do this in a variety of ways. For one, we can have a checklist of criteria, and compare the design against the criteria. Or, we can have experts review the design, and comment on areas to improve. We can also create a separate test of effectiveness, and then look to see whether our design leads to success. Finally, we can also look for independent measures, such as observation of learners, or impact on business outcomes. These aren't mutually exclusive, and we might use a mix on the way through development.

Another approach is subjective criteria: do learners think it was helpful? We can ask their supervisors or other observers whether their performance has changed. We can even ask consumers of their output, such as customer satisfaction.

It's also worth considering testing at different times. For one, we can see whether learners can perform immediately after the learning experience. We can also look to see if the learning has persisted later.

The goal is to determine whether the resulting design is achieving the intended outcomes.



Why is it important?

We invest heavily in learning. Michael Beer and his colleagues, in a Harvard Business School working paper, document that in just the US, 164 billion dollars were spent on training and education in 2012 alone. That's a significant outlay. What we should want to know is whether that investment is leading to any actual returns. In short, we should be looking for a suitable return-on-investment (ROI).

The only way to know whether we're investing wisely is to evaluate the learning. In fact, we should be regularly testing while under development, to ensure that our design actually matches the intent. Modern design approaches, whether Michael Allen's SAN Megan Torrance's LLAMA or David Merrill's Pebble in a Pond all have iterative evaluation.

In short, the only way to know whether our spending has generated value is to evaluate.

Why is it problematic?

There are multiple flaws in typical evaluation. First, it's not even always done! Then, it can be done superficially. It also might be done prematurely. We have good models for evaluation, from Kirkpatrick's Four Levels, Phillip's ROI, or Thalheimer's LTEM. However, they're too infrequently used. Beer, et al, in the afore-mentioned paper, cite a meta-analysis that shows only 10% of training programs are effective. That's a significant waste; 90% of the hundreds of billions each year is thrown away!

Why does this happen? One reason is that Kirkpatrick's model starts at level 1, just asking folks what they think. However, asking novices about the impact of the learning has about a .09 correlation with the actual value. That's essentially zero (with a rounding error). Kirkpatrick's design is intended to be used starting with level 4 (actually measuring the business impact), working backwards. Unfortunately, data suggests it doesn't happen that way. Instead, a lower percentage is used as you move up in levels.

Another problem is that meaningful evaluation requires going outside the boundaries of the L&D department; the business owner of the learning need is the one with access to the necessary data. Further, you should be looking at several points. At the end of the learning experience, you should ensure that it does lead to the ability (level 2). Then you should see whether it actually is achieving change in the workplace (level 3). Only then should you look to see if it's impacted the business.

Also, we should test usability first, then effectiveness. It used to be that we didn't really need to evaluate user experience, but that's changed with the move to learning experience design. That's not a bad thing. Engagement improves learning outcomes.

The process of evaluation actually starts before the design process. In the analysis process, you should have criteria that the learning is supposed to accomplish. This is based upon the desired performance, the gap between that and what is currently being experienced, and an identification that it is a skills or knowledge gap that is best addressed with a learning experience.

Our approach

We strongly believe that any request for learning should be driven by appropriate analysis. Too often it's the case that someone comes along and expects that a course will solve the problem, without sufficient insight about the core driver of the problem. Yet, that analysis is the difference between having an impact, and just spending money.

There are a number of intermediate steps that can be taken, and we know that evaluation needs to be flexible under typical circumstances. There are many real-world constraints that can push the evaluation decision one way or the other. It's just that, all else being equal, we prefer to be working on things that provide demonstrable value for money. We want documentable positive outcomes for you!

Our preference is to say "yes, and..." to see if we can go deeper to ensure that there's a clear goal we can impact. That doesn't mean we won't do work in other situations (we have to eat, too), but we'll push for trying to have a clear impact that the learning should occur. We're happy to work with you to do that right. Shall we?