Learning Measurement: It’s Not How Much You Train, But How Well

BY JEFFREY BERK

A few weeks ago I was talking to a colleague, Toni Hodges, a consultant specializing in learning measurement for organizations. We were discussing the difference between what I referred to as “activity measures” vs. “performance measures.” Toni aptly recast my terms into “how much you train vs. how well you train.” The terminology stuck with me. Let’s explore this very critical issue: how can we focus training measurement on how well instead of how much?

Quality and impact matter more in today’s world than quantity and “butts in seats.” Astute users of metrics will be the first to understand the difference, so those who produce the metrics should be making sure the right things are being measured.

Activity measures vs. performance measures

It seems sensible to start with the difference between activity measures and performance measures. Activity measures tell you how much you have trained. Performance measures tell you how well you have trained. Examples of activity measures include average number of students per class or facility utilization. A more comprehensive list of activity measures is included in Sidebar 1 on page 2.

Performance measures tell you how well you have trained. “How well” is reflected by such things as time to job impact, change in strategic results isolated to the training, or instructor performance. Some other common performance measures are shown in Sidebar 2 on page 3.

A key consideration in designing e-Learning measures is to decide what to measure and report. Measurement can be expensive and time-consuming. This has strategic implications for budgeting as well as for continuous improvement of e-Learning. Read this article to identify not only the measures, but also a system for applying them in a way that does not break the bank!

Continued on next page
MANAGEMENT strategies

Activity measures tend to be easier to collect than performance measures. Most learning management systems can gather such data, although it is debatable how well they perform at reporting the data for useful analysis. Activity measures are primarily data components of the training registration process. Because training needs to be coordinat-ed and scheduled, the information is readily available for “data mining.” Most organizations do a fairly decent job of reporting some activity measures that are easy to obtain such as number of classes run, or number of people trained.

Performance measures are more challenging to obtain. Some feel that certain performance measures are nearly impossible to obtain. Satisfaction-type measures such as instructor performance or courseware quality can be obtained through end-of-class evaluations, often referred to as “smile sheets.” Knowledge transfer measures can be obtained through pre- and post-test scores but the testing exercise can be a significant drain on a company’s resources, often making testing impractical to use for training across the board. Furthermore, measuring the impact of training on the job and on specific business results baffles many managers. The question is not only how to measure it, but how to measure it without spending more on the measurement exercise than on the training itself. Finally, even though Return on Investment (ROI) is often thought of as the Holy Grail of training measures, for many training groups this is nothing more than a distant dream.

Nonetheless, performance measures are critical. If you don’t know which programs had the greatest impact on the job and the company’s business objectives then your measurement system has some significant shortcomings.

Call to action: Performance metrics needed

The new millennium brought a completely different workplace — almost overnight. In the late nineties the economy was booming, the money was flowing, and it seemed that training had a blank checkbook to experiment with all kinds of new programs. It was an exciting time, e-Learning bloomed during this period and perception of its value grew.

Everyone knows what happened next. The bubble burst. The dot-com hey day

SIDEBAR 1 Activity measures

1) Fill rates by class, course, program, curriculum, location, or learning methodology
2) Student cancellation rates by class, course, program, curriculum, location, or learning methodology
3) Min, Max, Average participants enrollments in class, course, program, curriculum, location, or learning methodology
4) Enrolled vs. attended ratio by class, course, program, curriculum, location, or learning methodology
5) Student activity per class, course, program, curriculum, location, or learning methodology (% students attended, % students no-show, % students cancelled)
6) Class cancellation rates (% cancelled)
7) Completion rates by learning methodology (% of e-Learning offered vs. completed, % of Instructor-led Training (ILT) offered vs. completed)
8) Instructor utilization rates (available vs. used ratio)
9) Location utilization rates (available vs. used ratio)
10) Courseware utilization rates (available vs. used ratio)
11) Courseware Aging Report (% of courseware that is <1 yr old, 1-2 yrs old, etc.)
12) Student profile by class, course, program, curriculum, or location (% senior executives, % managers, % staff)
13) Response rates to evaluations (post-event, follow-up, manager, by course, class, location, program, curriculum, or learning methodology)
14) Financial summary (cost-per-course, cost-per-student, cost/benefit ratio)
and exuberance ended. The economy dried up and senior management adopted a “back to basics” attitude about doing business. In many companies, massive layoffs became the order of the day and departments were asked to do more with less. Finally, all budgets came under much closer scrutiny. As a result, just as in previous economic downturns, training departments could no longer use activity measures to justify expenditures or as the basis for budgets. The wonder is that, three years into the downturn, many groups continue to struggle with the issue of what to do instead.

Let's take an extreme and overly simplistic example. The CEO of a company is trying to control spending and costs. She asks the Chief Learning Officer for a training budget. That budget is prepared and submitted. The CEO wants some evidence to justify not cutting the budget this year. Reasonable business decisions dictate that tools that help the business perform better and improve the bottom line are good investments. The Chief Learning Officer tells the CEO that 80% of the workforce received training in the prior year, that nearly 300 classes were taught and that anecdotal evidence (smile sheets) suggest people really liked the training. Unfortunately, these are not sufficient reasons to support continued expenditures.

Why was the CEO not compelled to approve the budget by the metrics the Chief Learning Officer gathered? The facts offered are quantitative, they can be graphed, and they can be subjected to statistical analysis. The problem is that they do not relate to the business objectives of the company. Did the training help increase revenues? Did it help decrease costs? What kind of impact did it have on the average employee’s job performance relative to the salary paid to the employee? None of that data was collected, measured or presented.

What do you do?

**Choose offense rather than defense**

Being on the offensive is the worst situation to be in. The way to avoid this situation is to showcase the value of training as a strategic tool for improving job performance. Through scaleable, replicable, and practical measurement solutions, show how e-Learning improves the human capital of the organization.

Rather than continuing to provide the same activity metrics from years past that don’t work in today’s business world, the best-practice Chief Learning Officer will search for a proactive solution rather than a reactive one. Key ingredients of this solution include taxonomy, process, methodology and technology. All of these ingredients are absolutely critical to putting in place a cost-effective and compelling learning measurement solution.

**Finding the right model**

The first step in the process is to research learning measurement taxonomies and models. After all, why start with a blank sheet of paper when others have solved this problem before you were faced with it? In doing your research you are likely to run into many models with strong advocates. However, in my opinion, the one approach that has best stood the test of time is the Learning Levels model by Dr. Donald Kirkpatrick. Widely adopted since its creation in the 1950’s, and supported by professional groups such as the American Society of Training and Development (ASTD), many organizations today base their learning metrics on the Kirkpatrick model.

Many readers may already be familiar with Kirkpatrick’s four Levels of Learning Evaluation. However, let’s review them briefly for those who have not seen them before.

**Level One: Reaction.** This answers the question “Did they like it?” If done right, Level One can include a set of performance measures for all the pieces that comprise the satisfaction component of a learning program.

**Level Two: Learning.** This answers the question “Did they learn?” Since e-Learning is in the business of transferring knowledge and skills to individuals, it is only fitting to measure performance in this area. Just because someone had a satisfying experience does not mean they learned new skills or knowledge.

**Level Three: Behavior.** This answers the question “Did they use it?” A learner could have learned significant new knowledge or skills, without ever applying them on the job. This is a key performance metric.

**Level Four: Results.** This answers the question “Did it impact the bottom line?” This is probably the most important measure since it looks at the real business results delivered as a result of e-Learning.

The Kirkpatrick model is nice, but without a process to measure these levels it might not be practical. Enter Dr. Jack Phillips. Phillips is another individual you’ll run across when doing research for your measurement solution. He has made two major contributions to the field of learning measurement that will help you in your learning measurement solution. First, he built a process to measure Kirkpatrick’s four levels. Second, he added a fifth Level — Return on Investment (ROI). The ROI Process is used around the world as a tool to help learning organizations measure Kirkpatrick’s four levels, and Phillips’ fifth. You can see a brief overview of the Phillips ROI
But you’re not out of the woods yet, not by a long shot. If you read the books by Drs. Kirkpatrick and Phillips or attend certifications on the ROI Process endorsed by ASTD, you’ll quickly conclude that applying the ROI Process is a lot of work. It seems that nothing worth doing is easy. The Learning Levels have been embraced by so many because they cover the range of performance metrics from satisfaction to linkage to business results. The ROI Process has been accepted because it is rigorous, conservative and solid. So putting what is right into practice is the next step, but if it takes a significant investment of financial, physical and human resources it is not something you can easily sell to management. Asking for more money to measure training than the cost of training itself would not be prudent. So what can be done?

Enter technology. The new millennium, if nothing else, has taught us how to leverage technology so we can do more with less. The next section will discuss how to use technology as a practical enabler to implement industry-accepted measurement models.

Technology wrapped around methodology

Any measurement process aims at collecting data, storing it, processing it and reporting it. That is what needs to happen to measure Kirkpatrick’s Four Levels, Phillips’ Fifth Level and to work your way through the Phillips ROI Process.

With this in mind, why not leverage technology to streamline the steps so you don’t have to make significant outlays of resources for your measurement solution? Let’s start with data collection. There are a host of inexpensive Web-based data collection tools. These tools are easy to use and allow you to collect data from learners, instructors, managers, or any other stakeholder at any point in time. Leveraging the Internet to collect data saves the costs of paper processing. It is worth investigating and leveraging where practical and feasible. Even if paper is necessary, scanning technologies can easily move quantitative and qualitative data into databases for centralized data storage and processing.

In former times, “data storage” meant a metal file cabinet in a basement closet where so many end-of-class evaluations went as their final resting place. As a step up, simple tools like spreadsheet applications are inexpensive and powerful technologies to collect online data for efficient processing and analysis. More sophisticated solutions include using more powerful local relational databases such as Access, or using enterprise relational databases such as SQL or Oracle.

Queries. Analysis and interpretation of the collected information is the most critical piece in all of measurement. The key to success is to have flexible tools that support building queries that can then be automated or standardized. On-Line Analytical Processing (OLAP) tools such as Cognos or Microsoft Analysis Services are very powerful for querying large amounts of data. Do not, however, throw technology into the hands of functional users who are untrained on learning measurement. Allowing a line-of-business person to write their own OLAP query that compares learning data to business data can be counter-productive — if not downright dangerous — if they do not know what they are doing.

Here is an example of the problem. I recently read an article about a sales department that reported an ROI on training of over 10,000%! At first glance that is phenomenal. However a deeper
look revealed that there was no attempt to isolate the revenue increase from training. It just so happened that revenues increased at the same time sales training occurred. The data showed this correlation, and the writer of the article attributed the sales rise to the training. What about the economy, the competition, technology, people, process and a host of other factors that could have accounted for the sales increase? This was a classic result of an unsophisticated query and a naïve interpretation.

If you have skilled people writing appropriate queries that a novice user can then run reports against, you can leverage OLAP tools appropriately. The resulting reports will be targeted at the right level of person for the analysis.

**Reports.** This leads to reporting. In most organizations there is a need for four primary types of reports. Your measurement solution needs to be able to provide reports for each. These reports include tactical, aggregate, executive, and value analysis.

Let’s start with tactical reports. These are for the staff personnel in the training department who need to manage learning investments on a day-to-day basis. Examples of tactical reports include class evaluation summaries, verbatim respondent comments, and the actual evaluations themselves. These are used to spot problems before they get any bigger.

Next are the aggregate reports. Middle managers in a learning group use these reports. They aggregate tactical data for monitoring and quality control. For example the person responsible for managing the sales curriculum needs to quickly view all the course titles and determine which are most effective. The person managing the e-Learning content needs to easily see how effective it is compared to traditional instructor-led training.

Then executives in the learning group need reports. They need to see comparisons internally and externally. For example how did employees in each line-of-business that funds the training see the training impact their jobs? How did the key indicators of training compare to external benchmarks? And what exceptions occur on a day-to-day basis? Executives need summary, exception-based data that is comparable internally and externally.

Finally there is value analysis. This packages it all together. It is your balanced scorecard that you can present to management when the budget is up for renewal or when you need to fund a new program. It’s the measures you have an appointment with the CEO on a quarterly basis to review. It shows e-Learning vs. Instructor-led Training (ILT) so the CEO can see how that e-Learning investment is doing. Value analysis sums it all up. It presents all the measures in an easy-to-generate and easy-to-interpret scorecard. Value analysis allows a Chief Learning Officer to sit down with a C-level person or a line-of-business manager to cover how training impacted the job, how human capital was improved through training, and, yes, the financial return on training in hard- and soft-dollars. Metrics such as benefit-to-cost ratio, payback period and ROI percentage are in the value analysis.

Examples of some of the reports mentioned above are shown in Figure 2, below, and in Figure 3 on page 6. Figure 2 shows how actual data can be compared to performance goals and external benchmarks for each level of learning for continuous monitoring of the Learning Levels. Figure 3 shows the balanced scorecard that can be presented to management to showcase value in learning investments.

**A solution to work 100% of the time**

We’ve talked about how to take the Kirkpatrick and Phillips models and wrap technology around them for your measurement solution. But you still need to ensure that your solution has flexibility. Most of the time you can get by with reasonably accurate indicators to manage your business. These don’t have to be in-depth and dead-on accurate, but they must provide enough detail to allow people to make intelligent decisions. Designing the right data collection instruments targeted at the right stakeholders at the right time is important when building a model that is easy to use and that provides reasonable outputs. Recognize, though, that you might want to go deeper when the program is strategic or if it costs a lot. Your solution should then allow for more in-depth measurement to occur.

---

**Learning Levels Score Card**

<table>
<thead>
<tr>
<th>Time Period for Reporting: January 1, 2003 to May 29, 2003</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Level One - Satisfaction</strong></td>
</tr>
<tr>
<td>My Average [n=2,300]</td>
</tr>
<tr>
<td>My Goal</td>
</tr>
<tr>
<td>MTM Benchmark [n=341,934]</td>
</tr>
<tr>
<td>3 4 5 6 7</td>
</tr>
<tr>
<td>5.89</td>
</tr>
<tr>
<td>6.50</td>
</tr>
<tr>
<td>6.54</td>
</tr>
<tr>
<td><strong>Level Two - Learning Effectiveness</strong></td>
</tr>
<tr>
<td>My Average [n=2,953]</td>
</tr>
<tr>
<td>My Goal</td>
</tr>
<tr>
<td>MTM Benchmark [n=140,449]</td>
</tr>
<tr>
<td>3 4 5 6 7</td>
</tr>
<tr>
<td>5.47</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>6.24</td>
</tr>
<tr>
<td><strong>Level Three - Job Impact</strong></td>
</tr>
<tr>
<td>My Average [n=2,967]</td>
</tr>
<tr>
<td>My Goal</td>
</tr>
<tr>
<td>MTM Benchmark [n=83,526]</td>
</tr>
<tr>
<td>3 4 5 6 7</td>
</tr>
<tr>
<td>5.21</td>
</tr>
<tr>
<td>5.50</td>
</tr>
<tr>
<td>5.74</td>
</tr>
<tr>
<td><strong>Level Four - Business Results</strong></td>
</tr>
<tr>
<td>My Average [n=2,809]</td>
</tr>
<tr>
<td>My Goal</td>
</tr>
<tr>
<td>MTM Benchmark [n=109,674]</td>
</tr>
<tr>
<td>3 4 5 6 7</td>
</tr>
<tr>
<td>5.37</td>
</tr>
<tr>
<td>5.80</td>
</tr>
<tr>
<td>5.50</td>
</tr>
<tr>
<td><strong>Level Five - ROI</strong></td>
</tr>
<tr>
<td>My Average [n=2,986]</td>
</tr>
<tr>
<td>My Goal</td>
</tr>
<tr>
<td>MTM Benchmark [n=91,252]</td>
</tr>
<tr>
<td>3 4 5 6 7</td>
</tr>
<tr>
<td>5.90</td>
</tr>
<tr>
<td>6.00</td>
</tr>
<tr>
<td>6.33</td>
</tr>
</tbody>
</table>

**FIGURE 2** Learning Levels scorecard
Because of this need for flexibility, I suggest a set of three models that can help provide measurement solutions with 100% coverage. These models are the learner-based, manager-based, and analyst-based solutions. You get good breadth in measuring all five Levels of Learning in each of these models. But the depth increases as you move from learner-based to manager-based to analyst-based. The cost and complexity of the measurement solutions increase with depth too, so be careful of that.

**Learner-based**

This measurement model captures data from training participants at two distinct points during the learning process. The first point is directly after the learning intervention (Post Event) where the main measurement focus is on Kirkpatrick’s Levels One and Two in order to gauge satisfaction and learning effectiveness. Because there is a high response rate to these data instruments it is also critical to capture indicators for the advanced levels Three through Five. These indicators are, in effect, forecasting or predicting the future impact the training will have on the participant and the organization.

A second data collection point is a follow-up survey conducted a period of time after the participant has been back on the job. This survey is meant to true up the forecast and predictive indicators of Levels Three through Five by gathering more realistic estimates after the participant is back on the job.

The approach is low-cost if one leverages standard data collection instruments and if one utilizes technology and automation to capture, process and report the collected data. Thus this model can be used for all of your training, whether classroom-delivered ILT or e-Learning, to yield continuous measurements.

**Manager-based**

This model has the same data collection points as the learner-based solution but adds a manager-based dimension. The learner’s manager is another important data point. Managers can be sent an evaluation instrument timed to coincide with the participant receiving the follow-up survey. The manager survey focuses on Levels Three through Five of the Kirkpatrick and Phillips models. This provides estimates of job impact, business results and ROI from the manager’s perspective. The manager survey also asks “support” type questions to understand the on-the-job environment where the participant applied the training.

Due to the increased effort it takes to conduct and analyze manager surveys, the cost and time to measure at this level is higher than the learner-based model. With automation and technology to facilitate the dissemination, collection, processing, and reporting of the data, the cost and time can be minimized and this model could be used on a continuous basis for every training event a participant attends. More realistically, it will be used on a periodic basis for more strategic programs where manager data is more relevant.

**Analyst-based**

This model uses significantly more comprehensive post-event, follow-up and manager surveys. It also uses other analytical tactics that go beyond surveying. For example, to analytically measure Level 2 a detailed test is designed and...
administered to participants. Due to the time commitment of conducting a significantly detailed data collection and analytical exercise the analyst-based approach might only used for about 5% of all training programs in an organization. Typically these programs would be the more strategic or visible, and would have the budget to afford a more costly and time-consuming measurement exercise.

See Figure 4, right, for a view of how the learner, manager and analyst based models relate to each other.

Concluding thoughts

Activity measures are necessary to manage learning activities and will tell you how much you trained. But the workplace has changed and the demand for more comprehensive measurement tools heightens the need for better use of performance metrics that tell you how well you trained.

Through appropriate methodology that is made scaleable, practical, and replicable by leveraging technology, organizations can provide performance metrics in a cost-effective manner. Organizations should always keep in mind that the most perfectly accurate quantitative metrics are not required in order for performance measurement to be worthwhile. In fact, a recent study published in Harvard Business Review (May 2003) found that senior managers make decisions based on instinctive factors, not on the highly accurate and very costly data from highly-paid number-crunchers.

Use reasonable data based on indicators that are meant to predict and estimate your key performance metrics. This will save you significant cost and significant time, and it will accomplish the objective of providing the right performance measures to management for their decision-making.

FIGURE 4 Learning measurement methodology

San Francisco, California • November 12 – 14, 2003 Pre-Conference Workshops November 11, 2003

The program is divided into three core practice areas that mirror the three main roles that e-Learning producers are engaged in – Designer | Developer | Manager. The Program Committee, comprised of volunteer Guild Members, has identified the most critical topics that address these areas and those topics will be the focus of all our General Sessions, Think Tank Sessions, Concurrent Sessions and optional Pre-Conference Workshops.

Every session is designed to allow time to hear from the experts, debate and discuss with your peers and explore topics to their fullest. In addition to the structured time you will spend learning in the sessions, you will also have ample opportunity to network with the other participants in morning discussion groups, over lunch, and during social activities. The size and format of the conference will help you develop important professional connections during the conference that you can continue in the months and years to follow.

About the Guild

The eLearning Guild™

The eLearning Guild™ is a Community of Practice for designers, developers, and managers of e-Learning. Through this member-driven community, we provide high-quality learning opportunities, networking services, resources, and publications. Community members represent a diverse group of instructional designers, content developers, Web developers, project managers, contractors, consultants, and managers and directors of training and learning services — all of whom share a common interest in e-Learning design, development, and management.

The eLearning Developers’ Journal™

The Guild publishes the only online “e-Journal” in the e-Learning industry that is focused on delivering real world “how to” information. The Journal is published weekly and features articles written by both industry experts and members who work everyday in environments just like yours. As an active member, you will have unlimited access to the Journal archive.

People Connecting With People

The Guild provides a variety of online member networking tools including an Information Exchange and a Job Board. These services enable members to discuss topics of importance, to ask others to help them find information they need, and to provide leads to other members.

Become a member today! Join online at www.eLearningGuild.com!

Resources, Resources, Resources

The Guild hosts the e-Learning industry’s most comprehensive resource knowledge database. Currently there are over 3,800 resources available. Members have access to all of these resources and they can also post resources at any time!

Guild Research

The Guild has an ongoing industry research service that conducts surveys on a variety of topics each year. These topics are identified by the Research Advisory Committee. The data collected is available for all Members.

It’s About Leadership

The Guild draws leadership from an amazing Advisory Board made up of individuals who provide insight and guidance to help ensure that the Guild serves its constituency well. We are honored to have their active engagement and participation. The Guild has also established two committees made up of active members who help steer its events program and research efforts.

Discounts, Discounts, Discounts

Guild Members receive discounts on all Guild conferences and on other selected products and services. Your Guild membership will save you 20% off the list price of Guild events!

Join today at www.eLearningGuild.com!

AUTHOR CONTACT

Jeffrey A. Berk is Vice President of Products and Strategy for KnowledgeAdvisors, a corporate learning business intelligence firm. In his current role, Jeffrey oversees the operations and strategic direction for KnowledgeAdvisors. Berk also teaches a graduate school course on continuous improvement through business process benchmarking at Loyola University in Chicago and Northwestern University in Evanston, IL. He designed and developed the course curriculum.

Prior to joining KnowledgeAdvisors Berk worked for nine years with Andersen as Manager of Benchmarking Services. Jeffrey wrote the Andersen Benchmarking Methodology and spearheaded the development of quantitative and qualitative benchmarking tools to measure the effectiveness of client business processes.

Jeffrey has participated in analysis, integration and measurement engagements in the U.S., Europe and Australia. He has spoken on these topics and has been quoted and published in periodicals such as T+D Magazine, Training Magazine, CLO Magazine, Workforce Magazine, Fortune, IndustryWeek, and The CPA Journal.

Jeffrey is a CPA and began his professional career first working as an auditor before moving to consulting at Andersen and then to KnowledgeAdvisors. He has an MBA from the University of Chicago and degrees in business and accounting from the University of Kansas. Contact Jeffrey by email at jberk@knowledgeadvisors.com or by telephone at 312.423.8599.

Additional information on the topics covered in this article is also listed in the Guild Resource Directory.