Learning Effectiveness Measurement: A New Approach for Measuring and Managing Learning to Achieve Business Results

Dean R. Spitzer

The problem and the solution. One of the major challenges facing human resource development is how to achieve organizational impact through the use of learning interventions. So much has been written about the importance of achieving results from learning, but even practitioners who advocate results-oriented measurements tend to default back to low-level (attendance, satisfaction, and learning) measures in actual practice. This article describes the Learning Effectiveness Measurement (LEM) methodology, which was developed to address the challenge of providing a credible results-oriented learning measurement that would not only help evaluate learning interventions but also increase their effectiveness. The article begins by identifying some of the major limitations of existing learning measurement approaches, and then, it describes the LEM methodology in detail, provides an example of LEM in use, and concludes by showing how LEM addresses the limitations discussed at the beginning of the article.

Keywords: *HRD evaluation; Learning Effectiveness Measurement (LEM); business results; predictive measurement; causal chain*

Human resource development (HRD) is in the midst of a crisis today that some contend threatens its very survival. Simply stated, the crisis is the failure to show that investing in training and employee development produces demonstrable business results. Despite the more than $300 billion American companies spend annually on training, there are little or no data to show any positive impact of training on business results (Spitzer & Conway, 2002). In fact, most companies do not even try to measure the impact.

Advances in Developing Human Resources Vol. 7, No. 1 February 2005 55-70
DOI: 10.1177/1523422304272167

It should be disconcerting to all in the HRD field that according to estimates from the American Society for Training and Development (ASTD) and others, only 3% of all training courses are evaluated for business impact, whereas 88.9% are measured by happiness indexes (ASTD, 1995). With only a slight degree of exaggeration, Rutgers University Professor Cary Cherniss concluded that the American industry is spending billions and billions of dollars on training programs and doing no measurement of their effectiveness (Armour, 1998).

For a long time, training and development was accepted as something that was inherently good, and measures of success were the number of programs, number of participants, number of course days, training costs, end-of-course satisfaction survey ratings, and sometimes learning test scores (Spitzer, 1999). Slowly, this situation is changing. Today, there is a new call for accountability in the business of learning (van Adelsberg & Trolley, 1999). Organizational stakeholders are increasingly asking for proof instead of assumption—or at least strong evidence—of bottom-line effectiveness. Measuring effectiveness, business value, and return on investment (ROI) are becoming hot topics, especially because upper management is more frequently scrutinizing funding for training. When asked what benefits their organizations get from their investments in learning, a question increasingly being asked in business, government, and nonprofit organizations around the world, few HRD professionals have very good answers. Even more troubling is that few know how to increase the impact. IBM faced the same issues when, about 5 years ago, the author was asked hypothetically by a key executive, "What would happen if we cut IBM's training budget in half?" I answered, "It depends on which half we would cut!" The real problem underlying that exchange was that no one really knew the impact of learning. This situation led to the development of the Learning Effectiveness Measurement (LEM) approach, which is the main subject of this article.

It is sad that the technology of training measurements has not advanced very far. Although almost everyone believes that there must be a causal relationship between training and business results, few have been able to find one. Kirkpatrick's (1959, 1998) half-century-old four-level model gives sparse attention to Level 4 (business results) measurement and provides little insight into how to measure it. Furthermore, Phillips's (2003) five-level model has done little more than add another level, for ROI calculation, to Kirkpatrick's existing four levels.

The reader should note that the focus of this article is on for-profit organizations, and therefore, the term *business* is used; however, virtually everything presented in this article can be easily used in any government or nonprofit organization by simply focusing on the organizational impacts of that type of organization.

Although I share the concerns expressed by many about the limitations of the four-level model, the five-level model, and the current ROI craze, I have many more concerns about the current state of learning measurement that I believe are even more basic hindrances to progress in HRD measurement and evaluation. In the next section, I will share those concerns, and later, I will show how the LEM methodology addresses them.

Barriers to Progress in HRD Measurement

In this section, I will review some of the barriers that are impeding progress in HRD measurement. At the end of this article, I will return to this subject and show how the LEM methodology addresses them.

1. *Excessive emphasis on proving the value of learning rather than improving it*. In the HRD field today, the most common reason to measure is to justify and prove. This focus on summative rather than formative measurement plays right into the hands of the executives who insist that HRD professionals justify their existence. This, in turn, leads to an array of dysfunctional behaviors, including the ROI game and trying to isolate the impact of learning (rather than working synergistically with other business stakeholders to make a real difference). The emphasis on proving, rather than understanding and improving, has led to the use of methods that often distort reality and can prove almost anything. At IBM, I am often asked "How do I determine the ROI of this program?" My quasi-facetious response is "What ROI do you want?"
2. *Learning measurement tends to be retrospective*. The focus on summative evaluation has led to the tendency to measure retrospectively, as if that measurement is something that should be done only after a training program is completed rather than using measurement data to manage the program toward success. In contrast, as you will see, measurement is most powerful when used early and often.
3. *Confusion about what constitutes learning effectiveness*. *Effectiveness* means producing the desired result. So what is the desired result of learning? Common claims of success include the following: Two thousand employees attended the class, the average postcourse rating was 94% favorable, 87% of employees passed the end-of-course test, 75% of employees completed the end-of-course action plan, 80% of employees say that they are implementing at least part of their action plans, and so forth. But what do these claims really mean, and how do they demonstrate effectiveness? The problem with most of the popular measures of learning success is that all of them can occur without any positive impact on the organization. In fact, if the wrong skills are being taught and used, a success at Kirkpatrick's (1959, 1998) Levels 1 (reaction), 2 (learning), or 3 (application) can actually be a dismal failure at Level 4 (results). It should be clear that learning effectiveness is only relevant if it is aligned with what is important to the organization. Learning is only effective if it contributes to improved performance and organizational value.
4. *Reluctance to measure learning programs*. When given the option, most HRD practitioners will prefer not to measure learning programs. There

are many reasons for this reluctance, including the lack of strong measurement skills and the sense that "no news is good news." It is not completely their fault. Few HRD professionals have ever taken more than a basic statistics course, and there is very little consulting support for learning measurement in most organizations. Many practitioners perceive measurement as a political accountability mechanism just waiting to be used against them. For many, measurement is viewed, at best, as a necessary evil. At worst, it is seen as a menacing force that is greeted with about the same enthusiasm as a root canal. For instance, when I ask participants in my workshops to list terms they typically associate with measurement, the list commonly includes emotionally charged terms such as *judgment*, *fear*, *anxiety*, *threat*, and *insecurity*. This is one reason that might explain the seemingly illogical fixation with happiness indexes: They are easy to develop and administer, relatively safe to use, and virtually certain to produce positive results. Clearly, if trainers felt confident about the power of their interventions, they would seek more meaningful ways of keeping score.

5. *A vicious cycle of low expectations*. The focus on low-level outcomes has led to a vicious cycle of low expectations. Low expectations lead to weak interventions, which lead to poor outcomes, which reinforce low expectations. In contrast, as discussed later, when people feel good about their performance potential, they aim for high targets and want as much information as possible about how the intervention is performing so that positive action can be taken. They realize that win or lose, measurement is the key to improvement.
6. *Training logic takes precedence over business logic*. Most HRD practitioners use what I call "training logic," which is based on learning principles and tends to revolve around instructional design. Oftentimes, there is a total disconnect between HRD practitioners and management. For example, if HRD professionals do not understand the organization's strategy and business model (how revenue and profits are generated in the business), it is very difficult to trace the impact of any intervention to the bottom line, or even close to it. That is why the causal chain, which is discussed later, is such an important tool for transitioning HRD professionals from training logic to business logic.
7. *Inability to identify meaningful organizational measures*. Few HRD measurement studies use measures that are integral to the success of the organization. Instead of using the wealth of operational performance measures that could help track more relevant outcomes in their organizations, HRD practitioners are engaged in an endless, irrational reinvention of forms, questionnaires, and multiple-choice tests specifically aimed at measuring satisfaction and trivial learning outcomes. Too often, poor proxies of operational measures are used, such as surveys that ask for estimates of business impact. That is why ROI, based on subjective estimations of impact, became so popular. When actual organizational measures are used, they tend to be measures of convenience (measures readily available) rather than measures of strategic importance to the organization. Why does it seem so difficult to identify more relevant measures and collect more meaningful data? This is related to the failure to use business logic, as mentioned above. Unless one understands the

business, it is difficult to identify the measures that are most crucial to business success.

The LEM Methodology

LEM was developed to address the weaknesses of the existing approach to learning measurement being used at IBM and for use with IBM customers. The methodology was created based on a business need but was designed consistent with research and best practice. Its validity is based on numerous applications within IBM and with IBM customers.

In contrast to a traditional learning measurement, LEM was developed to guide the design of more effective (and powerful) interventions as well as track their effectiveness. The methodology development was guided by the principles discussed earlier, "What you measure is what you get!" Whatever we measure will inevitably drive the kind of intervention we design. In short, LEM was developed to better align learning with business results up front, not just to measure learning effectiveness after the fact.

There are five phases of LEM, as shown below:

1. *Predictive measurement* should be done before an intervention is selected or designed to help make the best learning investments, target the highest performance improvement opportunities, and provide crucial data to increase the effectiveness of intervention design. Predictive measurement can be viewed as part of the front-end analysis of a proposed intervention.
2. *Baseline measurement*, which should be done anytime before the intervention is implemented, is used to identify preimplementation measurements and target values for each measure.
3. *Formative measurement* should be done during intervention design to make sure that predictive measurement data are implemented in the intervention design and implementation plan so that maximum effectiveness can be realized.
4. *In-process measurement* should be done during the implementation of the intervention to track intervention effectiveness during deployment and enable corrective actions needed in a timely manner.
5. *Retrospective measurement* should be done after the intervention is fully implemented to collect postintervention data and provide input for final evaluative decision making.

One of the most important distinctions between the five LEM phases is the relative leverage of the measurement—its ability to influence results. Figure 1 shows the alignment of the LEM phases with the phases of learning and also shows the relative leverage of each phase. Clearly, the predictive phase has the most leverage, whereas the retrospective measurement has the least.

The following presents the LEM phases with particular focus on predictive measurement, which is the cornerstone of the methodology.

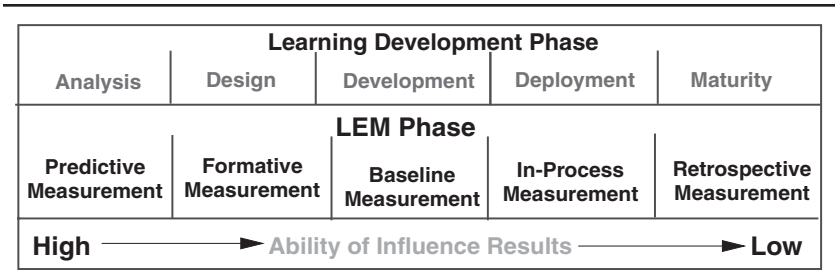

FIGURE 1: The Leverage of Measurement

Predictive Measurement

LEM is predicated on the belief that measurement should be used actively (in fact, proactively) to lead one to the kind of results desired. Similar to a road map, it should be used to select one's destination and navigate him or her to it and not just used to confirm that one has arrived or not! This requires a transformation of the traditional view of learning measurement from being predominantly retrospective to being predominantly predictive. Predictive measurement asks "What should happen?" whereas retrospective measurement asks "What already happened?"

Most people are somewhat skeptical about any discussion of measurement that relates to preintervention analysis because they may think of measurement as something that happens later and seldom think of analysis as a form of measurement. Because "what you measure is what you get," it is very difficult to expect learning that was not specifically designed to achieve business or organizational impact to get it. Measurement is crucial for establishing expectations, or targets, for impact.

Predictive measurement occurs early in the learning planning cycle, typically at the beginning of the analysis stage, and helps decision makers make the best learning design and investment decisions. Waiting to measure until after the intervention is selected and designed risks the loss of valuable measurement that can help drive results.

To achieve results, learning must be aligned with results. However, many times, desired results are poorly defined. Too often, the desired organizational impact, if it is ever defined, is disconnected from the learning objectives. Therefore, one of the biggest challenges in learning is to bridge the gap between learning interventions and real organizational impact. To do so, a systematic process is needed to trace the chain of causality between typical learning measures and business measures. To measure and attain business results through learning interventions, it is crucial to understand the intermediate relationships, even if the understanding is initially incom-

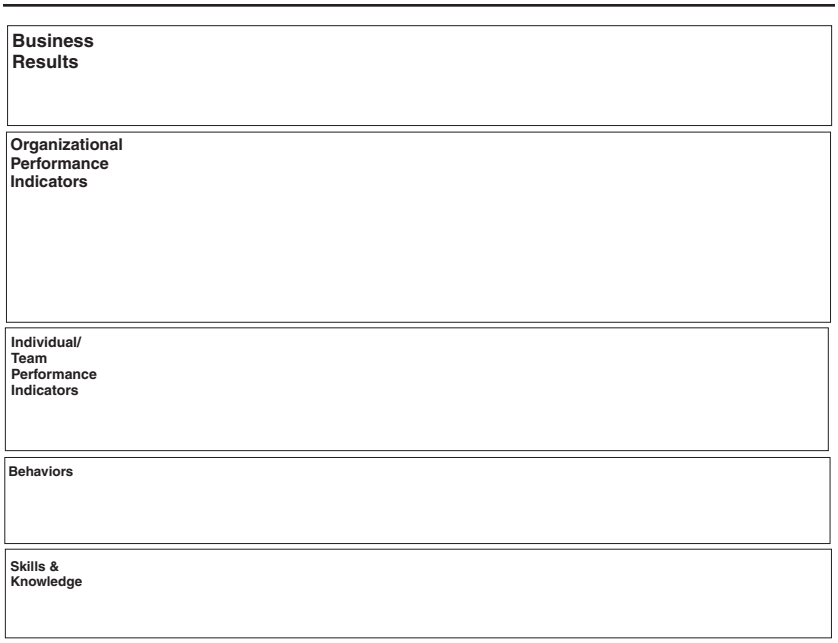

FIGURE 2: Causal Chain Worksheet

plete. In LEM, this is done through the use of qualitative causal chain diagrams. This approach is consistent with Kaplan and Norton's (2004) use of strategy maps that depict hypothesized causal relationships between business objectives and between business drivers and business outcomes.

Causal chains are typically relationship diagrams that can be used to trace the impact of learning through a chain of causes and effects—from skills, knowledge, and attitudes to behavior to individual and team performance to organizational performance—culminating with financial business results. The causal chain represents a map composed of a hierarchy of indicators linking learning with business results. This causal understanding has long been the missing link in learning that attempts to achieve business impact. Not only does this causal logic help identify measures that can be used for tracking business impact, but more importantly, it provides critical linkages for driving that impact. One thing is certain: A learning intervention is unlikely to ever have an organizational impact if it does not connect the critical links in a causal chain. To develop causal chains, a five-tier worksheet can be used, an example of which is depicted in Figure 2.

Using this template, and through an interactive, iterative process, a causal chain diagram can be developed. Figure 3 depicts an actual causal chain, which will be discussed as the LEM example later.

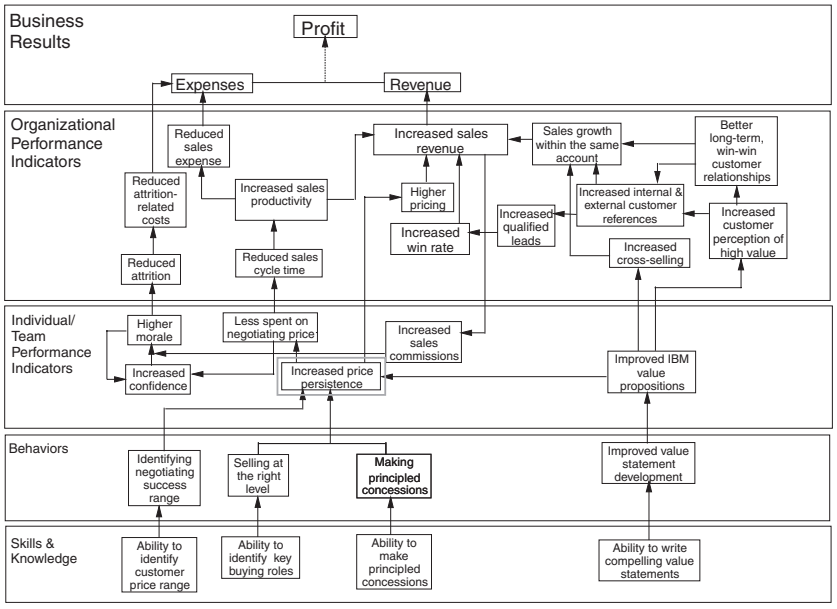

FIGURE 3: Example of Causal Chain Analysis

Causal chains begin with desired business results (derived from the priorities of business leaders). From those business results, relevant organizational performance indicators or potential measures (usually existing operating metrics) can be identified. The prioritized organizational performance indicators (linked to desired business results) provide a clear indication of what is important from an individual or team performance perspective. Once the individual or team measures of success are defined, the key behavior gaps can be determined. Then, the knowledge and skill needed to affect the behavior change can be defined. In other words, this is to ask what knowledge and skills are needed to enable the behavior change necessary to drive individual or team performance improvement that will drive organizational performance metrics (linked with desired business results) in the right direction. However, the significance from a measurement perspective is that this process defines knowledge and skill priorities (Level 2), behavioral measures (Level 3), and individual and team measures and organizational performance measures (Level 4), which become the basis for a measurement plan that can not only track impact but also help drive it.

Causal chains do not have to be fancy. It is not the diagram that is important; it is the understanding that it is developed through the process of developing the causal chains. This causal understanding has long been the missing link in learning to achieve a business impact. As one moves up the chain

of causality, one will be able to demonstrate the business logic of the intervention to the client and the organization.

One of the major factors discouraging efforts to link learning interventions with business results has been the perceived need to isolate the effects of training from other causal influences. In rigorous research, this might be warranted. However, one of the keys to high-leverage interventions is the synergy among intervention components. Business logic is based on the concept of contributory causality—the realization that organizational results are caused by the confluence of multiple influences—and that no function or intervention is solely responsible for any business result. In an applied business or government context, no function should be required to rigorously isolate its impact or the impact of an intervention from other potential contributing factors. Doing so may tend to undermine collaboration and the potential for synergy.

The process of developing a causal chain is as important as the product, the causal chain itself. The dialogue among stakeholders that is generated through the development of causal chains is beneficial in many ways, including providing a common language, which learning consultants, education sponsors, and executives can use to collaborate for achieving higher impact learning interventions. The more learning people work with business logic, the more attuned they may become to the business and business drivers. This will ultimately have a profound impact on the business literacy of learning professionals in an organization.

It is important to emphasize that causal chains are a set of hypotheses and not established facts, at least not until they are validated, which can be done based on regression analysis with actual data collected. Whether explicitly validated, causal chains are always subject to refining with time as causal relationships are verified or discounted through empirical testing or discarded as business requirements change with time. This iterative process may lead to new insights in learning and its interaction with other business drivers. As they are refined, causal chains also become reusable intellectual capital for the units that develop them and for the organization as a whole.

Causal chains also help identify key leverage points (indicators that appear to have the greatest impact potential) and critical factors for increasing the success of learning interventions. These key leverage points and critical success factors become the most crucial intersection point between LEM and intervention design.

Based on the causal chain, the next step in predictive measurement is to identify measures to target and track. Everything on a causal chain does not have to be measured, but the most important indicators to business impact should be measured. These become the basis for the measurement plan. Although the focus of the LEM is on organizational performance and business impact measurement, measures at all levels can be included in an LEM

measurement plan. In selecting the best measures, emphasis is placed on alignment between the proposed interaction and desired results, ease of measurement, and accessibility of data.

Baseline Measurement

The next phase in LEM is baseline measurement, which can actually be done any time before the intervention is deployed. One of the biggest mistakes made in learning measurement is the failure to collect baseline (preintervention) data. Without baseline data, no before-and-after comparisons can be made, and it is impossible to know if there has been any improvement. Furthermore, without baseline data, credible learning measurement targets cannot be established. Without such vantage points, learning professionals are often designing interventions without knowing how much improvement they want and in what areas. This is like shooting in the dark—not a very good idea—either with a gun or a learning program.

If relevant data are being continuously tracked in the organization, baseline data should be relatively easy to collect. Unfortunately, if data collection is time consuming, most learning professionals and their clients are reluctant to invest their scarce resources in collecting measurement data. This is a serious mistake that has come back to haunt most learning functions that fail to track the effectiveness of their learning interventions, and without baseline data, no meaningful measurement comparisons can be made.

Formative Measurement

Formative measurement involves the review of the intervention design to ensure that the intervention and its implementation plan are powerful enough to achieve the desired results identified in the predictive measurement phase, especially with respect to the critical success factors. The use of the term *formative* in LEM is slightly different from how the term is often used. In this case, formative measurement occurs during the design and development process. Improvements are still encouraged after deployment, but those in-process improvements are handled in the next phase.

To achieve business impact, learning programs must overcome the many barriers that prevent the potential chain of causality from being realized. Even if a learning program is well designed in the traditional sense, target learners might not participate in the learning program; learning might not occur; learning might not result in behavior change; behavior change might not improve employee performance; even if employee performance improves, it might not impact organizational performance; and organiza-

tional performance improvement might not be reflected in the financial results of the business or government agency. Only when these barriers to impact are overcome can real results from learning be achieved. Focusing on behavior change, performance improvement, and business measures, rather than just learning measures, significantly changes the rules of the game for intervention designers as well as those involved in learning measurement (perhaps the same people). Intervention design is no longer just a matter of disseminating knowledge and developing skills. The challenge is to develop effective instruction that is wrapped in a package that might include nonlearning components as well. As we know, it is virtually impossible to make a significant impact on business measures with instruction alone. LEM makes a case for more potent, systemic interventions.

In-Process Measurement

It is vital for measurement data to be collected during the implementation of an intervention. Just because an intervention has been carefully designed to meet the learning and performance requirements does not mean that it will be as effective as anticipated once it is deployed. The frequency of in-process measurement depends on the criticality of the intervention. Obviously, the more critical the intervention, the more important continuous measurement will be. These data should provide timely feedback on how well the intervention is working, on possible problems that might require some corrective actions, and on opportunities to further enhance the intervention or, in rare instances, discontinue it.

In-process measurement is high leverage because it typically requires little effort and provides valuable, ongoing feedback. Traditional measurement typically occurs too late for such decisions and remedial action. To maximize effectiveness, measurements should occur throughout the intervention lifecycle, from initial conception to the end of deployment. Bear in mind that without measurement, it is impossible to manage anything and that the earlier we start measuring, the more leverage we can get from the measurement.

Retrospective Measurement

Retrospective measurement is the final measurement that occurs at the end of the intervention deployment. This does not necessarily mean that the intervention is terminated and taken out of service; it means that it has reached a point at which the intervention is sufficiently mature that it can be retrospectively measured. Retrospective measurement can also be considered the last data point for in-process measurement. The primary purpose of

measurement at this time should be making final judgments about the intervention, including the calculation of ROI, if desired. However, retrospective measurement is typically too late to enable changes in the existing intervention; although, the data can be used to inform decisions about future interventions.

An LEM Example

The causal chain depicted in Figure 3 is an actual IBM example. The IBM Software Group was concerned with the trend toward cutthroat cost cutting (Bank, 2004), long before this became a software-industry-wide problem. The crux of the problem is that customers (who have been used to getting free or low-cost software bundles on their PCs) do not realize why enterprise software is so costly and do not always appreciate the software research and development costs that must be amortized (similar to drugs in the pharmaceutical industry).

In a proactive effort to address the issue, LEM was used to develop a solution. Key stakeholders worked with the author to develop a causal chain. It was determined that the key driver of increased profits (the business result) was higher pricing (the key organizational performance indicator) and that the key driver of higher pricing was increased price persistence during negotiations (the key individual and team performance indicator). At the same time as maintaining the price during the negotiations, it was important to also provide customers with improved value propositions (another important individual and team performance indicator) to justify the higher prices.

Drilling down further, the stakeholders were able to identify the critical behaviors that would drive the desired performance improvement. The most important behavior was making principled concessions, which means that a price concession must be accompanied by a concession on the part of the customer. This was something our sales teams did not do well. The problem was exacerbated by a lack of cross-brand collaboration, which was identified as a nontraining critical success factor.

Once the key behaviors were identified, it was easy to identify the skills and knowledge needed to enable those behaviors. What is more, the causal chain had produced amazing consensus among the stakeholders and high commitment to the intervention. The causal chain also identified the key measures of business results (gross profit), organizational performance (pricing—the price of the software negotiated by the sales teams that were trained), individual and team performance (price persistence—the difference between the price at the beginning of the negotiation and the price at the end of the negotiation), behavior (making principled concessions—records of principled concessions offered by the negotiation team and accepted by the customer), and skill (the ability to make principled con-

TABLE 1: Measurement Data for the Learning Effectiveness Measurement Example

Measure	Predictive Phase	Baseline Phase	In-Process Phase	Retrospective Phase
Profit (% profit change)	11%	0%	6%, 8%	9%
Pricing (% pricing change)	15%	0%	8%, 9%	10%
Price persistence (% average reduction in beginning vs. end price concessions)	–18%	0%	–8%, –12%	–14%
Making principled concessions (number of principled concessions / number of total concessions)	.50	.10	.20, .35	.40
Ability to make principled concessions (average course assessment score)	80%	0%	70%, 75%	85%
Reaction to course (average follow-up survey score)	90%	Not applicable	85%, 90%	90%
Course participation (% of target audience participating)	95%	0%	30%, 60%	90%

cessions as measured by instructor ratings of the trainees during training program). There were many others that could have been used. It is interesting that causal chains provide a wealth of relevant measures from which to select, as opposed to the normal dearth of relevant business and organizational performance measures.

Data on these measures were then collected through the major LEM phases (baseline, in-process, and retrospective). The predictive data were the target values for each measure. Formative measurement simply related to ensuring that the intervention appeared to be potent enough to hit the target. Table 1 provides a view of the data collected. Percentages, rather than absolute figures, are provided because of IBM confidentiality restrictions. Profit increases could be easily monetized if desired.

The measures being tracked are provided in the left-hand column. The other columns provide the data for four of the five phases of LEM. There are no data for the formative phase because that phase is a purely qualitative one, during which the intervention is matched with the predictive phase requirements. Note that the baseline phase data for profit, pricing, and price persistence are listed as 0% because there will be no change for the baseline in percentage.

The data indicated that there is clear improvement in the desired direction but that it is not always consistent with the goal established by the pre-

dictive phase. One important aspect of the LEM design is that we want to encourage ambitious goals. The fact remains that there is highly significant improvement on all performance measures. You will also notice that there are two in-process phase data points. LEM encourages multiple in-process measurements, which allow us to track the intervention during deployment. If the data are below expectations, corrective actions can be taken. In this case, most of the shortfalls in the in-process data can be explained by the ramping up of participation in the training.

Using mostly existing data sources, data were collected at each level of training measurement: business results (profit, pricing, and price persistence), behavior (making principled concessions), learning (ability to make principled concessions), reaction (reaction toward the course), and attendance (course participation). In this case, similar to most implementations of the LEM in practice, the key to effectiveness was the use of the predictive measurement for goal setting and the design of the intervention and the use of the in-process measurement to ensure that the intervention was on track and not the use of measurement after the fact.

The Advantages of LEM

This section discusses LEM advantages by presenting how it addresses each of the barriers discussed earlier.

1. LEM is unlikely to be misused for justification purposes. The entire LEM methodology was carefully crafted for increasing the effectiveness of learning, not for justifying it. If justification is the objective, there are many other easier ways of doing it.
2. LEM uses five phases of measurement that span the entire learning cycle, from initial requirements gathering to design and development to deployment to end of lifecycle. As such, LEM can, and should, be done continuously, not just retrospectively when the deployment of a learning intervention is completed. LEM is truly an end-to-end learning measurement methodology.
3. LEM focuses on targeting actual business and organizational indicators of effectiveness rather than just the typical pseudo-effectiveness indicators.
4. Because LEM is seamlessly integrated with requirements gathering, front-end analysis, and design, it should be viewed as an extension of what HRD practitioners are already doing. This should reduce the reluctance to measure because measurement is no longer an activity artificially separated from design and implementation.
5. LEM avoids the vicious cycle of low expectations by encouraging HRD professionals to set ambitious goals and providing the tools for realizing them.
6. LEM encourages the use of business logic through the development of causal chains, which identify relevant intermediate impacts needed to achieve the ultimate business or organizational impacts and link them with learning outcomes. Consequently, the learning logic becomes much

more relevant because it is viewed in the context of performance and desired business results.

7. LEM measures both financial and nonfinancial outcomes using a combination of qualitative and quantitative measurement methods. It puts quantitative measurement in its proper perspective as something that is only meaningful if the proper context and qualitative linkages have been established. LEM also tends to encourage the use of more business-oriented measurement methods and to deemphasize the old ones (e.g., satisfaction surveys and tests). LEM uses measures that are meaningful to all stakeholders. LEM identifies and uses real performance and business measures rather than depending on pseudomeasures, inadequate proxies, and traditional learning and satisfaction tools.

LEM is more than just a conventional learning measurement methodology. It is an approach to managing the learning and performance improvement process and achieving the desired organizational impact. The real power of LEM lies in its ability to drive the learning, performance, and business outcomes, not just assess the extent to which outcomes were, or were not, achieved.

However, to be successful, LEM will benefit from different working relationships than those that have traditionally existed between HRD professionals and line management. The emphasis on aligning learning with business results and on continuous measurement will make new partnerships desirable, if not essential. Clearly, HRD practitioners using LEM methodology will have to build new relationships with their clients. This new relationship will be much more business and performance focused and will provide access to those who own the metrics that will be used to measure individual and team performance and organizational impact. Under this paradigm, HRD practitioners can no longer live in their training worlds and operate according to their training logic alone.

Conclusion

LEM methodology has been presented as a practical, results-oriented approach to HRD measurement and evaluation. Additional use and refinement of the methodology is certainly called for. However, based on initial use, LEM has been shown to be a methodology that can truly change the context in which HRD measurement is done by making it more proactive, more positive, and with a higher leverage.

As mentioned earlier, the biggest problem is not with the technical methodologies; it is with the measurement mind-set that exists in the HRD field. Another function the LEM methodology may have is to change the mind-set as well as improve the technical methods. As an HRD practitioner, my vision is for measurement to be used proactively and enthusiastically by both researchers and practitioners. In particular, those who have used LEM have embraced it eagerly with the same excitement that I have seen with employees managing the quality of their own work using Statistical Process

Control or with athletes tracking their progress to make appropriate mid-course corrections. After all, without effective measurement, it is very difficult to win the game of learning, business, or any other aspect of life. I believe that concepts such as proactivity, participation, and partnership are more important to the future success of HRD measurement and results than any technical issues, such as sampling, statistics, control groups, and ROI calculations.

References

American Society for Training and Development (ASTD). (1995). *ASTD benchmarking forum report*. Alexandria, VA: Author.

Armour, S. (1998, October 7). Big lesson: Billions wasted on job-skills training. *USA Today*, p. 12.

Bank, D. (2004, June 21). In cutthroat software industry, vendors slash pricing to bone. *Wall Street Journal*, p. B1.

Kaplan, R. S., & Norton, D. P. (2004). *Strategy maps: Converting intangible assets into tangible outcomes*. Boston: Harvard Business School Press.

Kirkpatrick, D. L. (1959). Techniques for evaluating training programs. *Journal of the American Society for Training & Development*, *13*, 3-9.

Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler.

Phillips, J. J. (2003). *Return on investment in training and performance improvement programs*. San Francisco: Butterworth-Heinemann.

Spitzer, D. R. (1999). Embracing evaluation. *Training Magazine*, *16*, 42-47.

Spitzer, D. R., & Conway, M. (2002). *Link training to your bottom line*. Alexandria, VA: American Society for Training and Development.

van Adelsberg, D., & Trolley, E. A. (1999). *Running training like a business*. San Francisco: Berrett-Koehler.

Dean R. Spitzer, a researcher and consultant with IBM Corporation, is an internationally recognized consultant and innovator in the area of performance improvement, high impact training, learning measurement, and organizational change. He developed the Learning Effectiveness Measurement methodology, which is being used throughout IBM and with IBM's clients. He is the author of six books and more than 130 articles and book chapters and has received numerous awards from the International Society for Performance Improvement "for outstanding contributions to advancing the field of Human Performance Technology."

Spitzer, D. R. (2005). Learning effectiveness: A new approach for measuring and managing learning to achieve business results. *Advances in Developing Human Resources, 7*(1), 55-70.