Measuring for Success

What CEOs Really Think about Learning Investments

Jack J. Phillips
and
Patricia Pulliam Phillips

Alexandria, Virginia

14 13 12 11 10 1 2 3 4 5 6 7 8

Library of Congress Control Number: 2008937405

ISBN-10: 1-56286-588-9
ISBN-13: 978-1-56286-588-7

ASTD Press Editorial Staff:
Director: Dean Smith
Manager, ASTD Press: Jacqueline Edlund-Braun
Manager, Acquisitions and Author Relations: Mark Morrow
Senior Associate Editor: Tora Estep
Senior Associate Editor: Justin Brusino
Editorial Assistant: Georgina Del Priore

Copyeditor: Alfred Imhoff
Indexer: April Michelle Davis
Proofreader: Kris Patenaude
Interior Design and Production: Kathleen Schaner
Cover Design: Katherine Warminsky
Cover Illustration: Fry Design Ltd.

Printed by Versa Press Inc., East Peoria, IL, www.versapress.com

Contents

◆ Chapter 10

The Learning Scorecard

Executives Will Pay Attention If It's Relevant and Insightful

Developing a comprehensive learning and development function and process is a strategic decision, and top executives understand the need to bring a strategic focus to learning. They see learning as a driver to reach strategic goals and objectives. Though some learning and development functions are flourishing and achieving the desired results, others are perceived as being disconnected from business strategies and operating in a world unrelated to the work of the organization. They sometimes stray from the original goal and fail to demonstrate the actual contribution to the organization. They lack a process to show the management team the measurable value of learning and development and how programs and solutions are linked to strategy, influencing important business measures.

In the CEO survey, 22 percent of respondents said they have a learning scorecard. Our composite research with executives reveals six major concerns about a macro-level scorecard for learning and development. These translate into these specific types of information needed by the executives:

1. The impact of learning and development at the macro level, across all programs and other learning solutions.

2. Brief reports rather than detailed impact studies, at least for a few major programs. Executives want information they can quickly understand and digest.
3 The connection to business objectives and data to show that learning and development is making a contribution and driving certain business improvement measures.
4. The overall contribution of the process, but not necessarily the ROI analysis, for the entire learning and development function. Executives need some indication that the function is adding value, and that every program is evaluated at some level.
5. Different types of data are needed, both tangible and intangible, gathered in different time frames, and often from different sources.
6. The alignment between the major learning and development programs, strategic objectives, and operating goals.

These important needs of senior executives can be met with the learning scorecard.

For the learning and development team, additional data is often needed to ensure that processes are efficient and effective. Improvements are often identified to enhance future results. Thus, process improvement is a key focus of the evaluation data. In addition, the learning and development team needs data on application, impact, and ROI to:

1. Enhance the perception of the learning and development so that the key stakeholders (including executives) have positive impressions about the contribution and usefulness of major initiatives.
2. Build credibility for learning and development with all stakeholders. Data must provide convincing evidence from a realistic, valid, and reliable approach to show that the comprehensive measurement and evaluation process is credible.
3. Justify future expenditures. Executives who need to see that the contribution is significant and the payoff is appropriate approve most learning and development growth. An absence of this type of data may result in budget decreases instead of increases.

4. Enhance management support. Evaluation data should convince management that their support is important, necessary, and a critical part of learning and development success.
5. Provide information for benchmarking with other learning and development functions so that best practice comparisons are possible.

Collectively, these drivers and needs from different target audiences create an unprecedented demand for a measurement system that will collect, distribute, and interpret the necessary data on a routine basis.

Why a Scorecard?

In the past several years, a shift to results-based processes has been evolving in the learning and development function. Initially, learning and development reporting was activity based and not necessarily connected directly to many of the business strategies. The reports from functional specialists lacked accountability and the processes needed to show the value of their contribution. Overall, efforts to measure the impact of learning and development were nonexistent.

Today, things have changed. Many learning and development processes are linked to business strategies, and there is a more comprehensive approach to measurement and evaluation. Table 10-1 depicts this paradigm shift from activity-based to results-based programs. Though the last item on the list is the focus of this chapter—the method of communicating progress and reporting results—the issue involves many of the topics in the results-based approach discussed throughout this book. In the traditional, activity-based mode, the learning and development team would report results from an input perspective, where investment and commitment are emphasized, such as the number of hours consumed in learning activities, the number of participants involved in programs, and the amount of money invested in the process. Today, progressive learning and development functions present a complete scorecard that shows the impact of learning, revealing eight types of data, from input to ROI.

Table 10-1. The Paradigm Shift in Learning and Development

Activity-Based	Results-Based
No business need for the program	Program linked to specific business
No assessment of performance issues	Assessment of performance effectiveness
No specific measurable objectives	Specific objectives for behavior and business impact
No effort to prepare participants to achieve results	Results expectations communicated to participants
No effort to prepare the work environment to support transfer	Environment prepared to support application
No effort to build partnerships with key managers	Partnerships established with key managers and clients
No measurement of results or ROI analysis	Measurement of results and ROI analysis
Reporting on programs is input focused	Reporting on programs is output focused

This learning scorecard primarily provides information to the client group, including top executives. However, it also provides useful measures for the learning and development team. Multiple customer demands are at the core of the scorecard's development. Although there are many stakeholders, two broad groups of customers are served by the learning and development function: the *consumers* and the *clients*. The consumers are the participants involved in the process. The clients are the executives who fund, support, request, or approve programs.

The learning and development team needs immediate information from the consumers in the form of feedback about its processes and programs. The staff must know the extent to which participants are learning new skills and knowledge from the processes in which they participate. The client group, on the other hand, is more interested in application, impact, and ROI. This group is interested in obtaining significant behavioral change from employees as they interact with customers, suppliers, and team members. They are also interested in the actual linkage to business impact—to have some assurance that the programs are helping the operating units achieve their goals. Some executives want the actual ROI,

to ensure that major learning and development programs represent an important payoff for the organization. The demands of these two groups create the need for a balanced approach to measurement, reporting a variety of data, both qualitative and quantitative, collected in different time frames from different individuals.

Shifting to Output-Focused Reporting

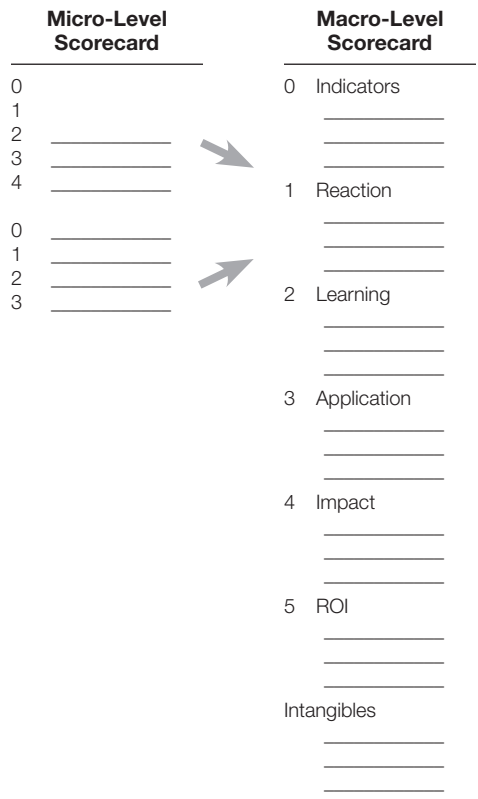
Most traditional reporting is focused on the actual investment in, and activities about, learning—the input side of the statistics. CLOs use a variety of statistics, such as number of hours of learning programs, number of people involved, number of programs, number of multiple enrollments, and investment per associate. Though these statistics are useful in providing an indication of the level of support, scope, and investment, they do not reflect outcomes, such as application or business impact.

Most evaluation processes typically concentrate on micro-level activities, measuring one program at a time and reporting the results. But for a learning and development function with several hundred learning programs, it is practically impossible to measure every program at every level of evaluation; nor is it desirable. For some programs, however, it is extremely important to show business impact and ROI.

The challenge is to measure as many programs as possible at the higher levels (application, impact, and ROI) and to integrate the data in a meaningful way to show the overall contribution of the learning and development function. In essence, this process takes a micro-level activity (evaluation of a specific learning program) and presents a macro-level view (evaluation of all learning programs) of this function.

Figure 10-1 illustrates the concept. Though each program is evaluated on the micro level, only a few selected measures in each of the micro evaluations are captured for the macro evaluation. In essence, it takes the most critical, important, and executive-friendly measures to go on the learning scorecard. Building a macro report requires several important steps that address a variety of issues.

Figure 10-1. Micro-Level versus Macro-Level Scorecards



Beginning with the end in mind and picturing the fully developed scorecard may be helpful in progressing with the steps to develop this type of reporting process. Figure 10-2 shows an outline of a comprehensive learning scorecard. This figure shows categories only and does not reveal the actual data. About 40 measures are included in the complete report. Eight major categories of data are presented, including:

0. Inputs and indicators showing the commitment of resources, including volume and efficiencies.
1. Reaction and planned action as participants are involved in learning solutions.
2. Skill and knowledge acquisition as participants are involved in learning solutions.

Figure 10-2. Sample Outline of Reporting for a Learning and Development Function

0. Inputs/indicators
    1. Number of employees involved
    2. Total hours of involvement
    3. Hours per employee
    4. Investment as a percent of payroll
    5. Cost per participant
    6. Delivery mechanisms

I. Reaction and planned action
    1. Percent of programs evaluated at this level
    2. Ratings on 3 items vs. target
    3. Percent with action plans
    4. Percent with ROI forecast

II. Learning
    1. Percent of programs evaluated at this level
    2. Types of measurements
    3. Self-assessment ratings on 3 items vs. targets
    4. Pre/post-average differences
    5. Percent receiving certification

III. Application
    1. Percent of programs evaluated at this level
    2. Ratings on 3 items vs. targets
    3. Percent of action plans complete
    4. Barriers (list of top 10)
    5. Enablers (list of top 10)
    6. Management support profile

IV. Business impact
    1. Percent of programs evaluated at this level
    2. Linkage with business measures (list of top 10)
    3. Types of measurement techniques
    4. Types of methods to isolate the effects of programs

V. ROI
    1. Percent of programs evaluated at this level
    2. ROI summary for each study
    3. Methods of converting data to monetary values
    4. Fully loaded cost per participant and comparison

Intangibles
    1. List of intangibles (top 10)
    2. How intangibles were captured

Awards
    1. Learning and development professional awards
    2. Industry awards

3. Behavior change and actions taken as participants apply on the job what they have learned.
4. Business impact driven from the application of learning—which may use individual, team, or organizational measures.
5. ROI for selected programs showing the actual return from specific learning solutions.
6. Intangible measures that are not converted to monetary values but represent important measures.
7. Awards received by the learning and development team.

Collectively, these categories satisfy executive needs and learning and development team needs, using the classic five levels of evaluation:

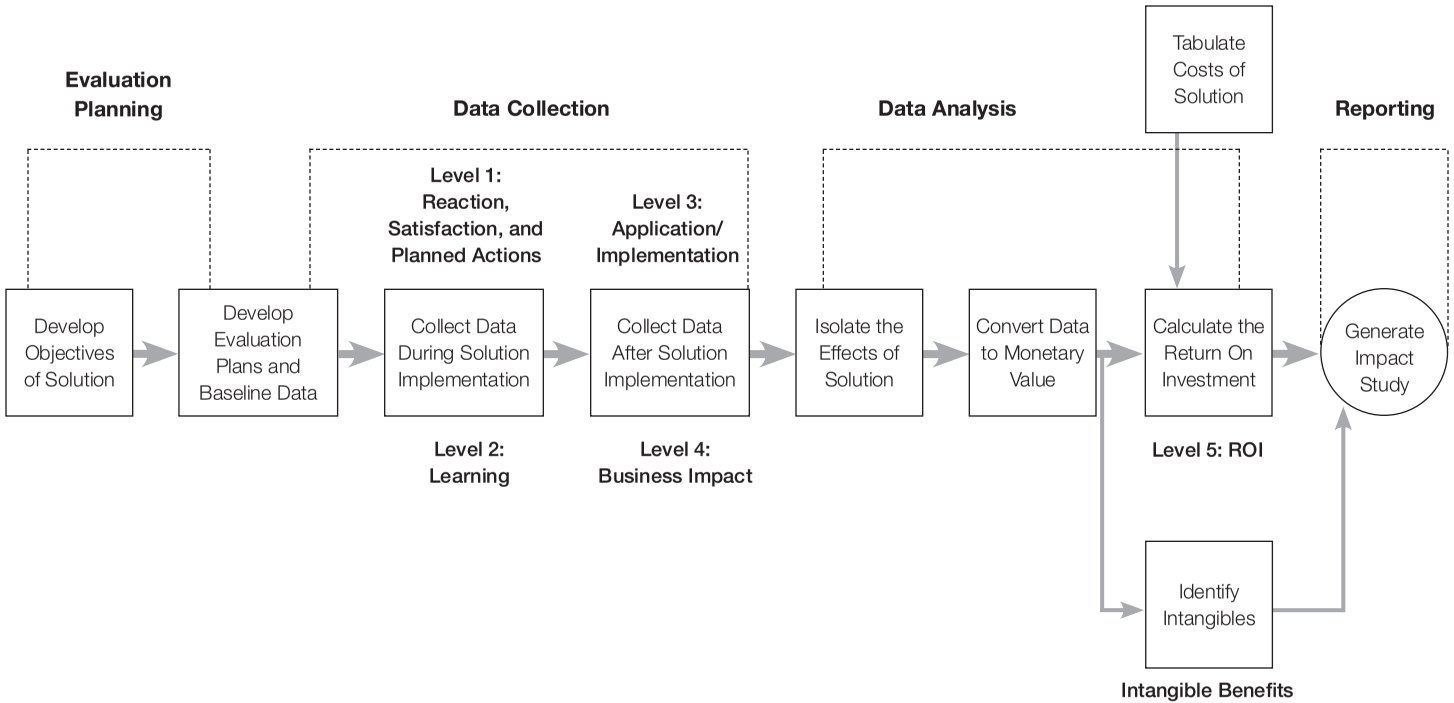
1. Reaction
2. Learning
3. Application
4. Business impact
5. ROI.

The Foundation for the Scorecard

The learning scorecard is built on a foundation that includes several important principles necessary for the development of a comprehensive measurement and evaluation process, which is referred to as the ROI methodology. The part of this process is the framework for evaluation, which represents different types of data collected at different time frames, as described in chapter 2.

A step-by-step process model is needed to plan and implement the evaluation, especially when programs are evaluated in a postprogram time frame. Whenever there is an attempt to measure application or impact, a process model is needed to ensure that the data is collected, analyzed, and reported in logical and sequential steps. Figure 10-3 shows the ROI process model used by thousands of organizations to develop impact studies on different types of learning solutions. It represents the most frequently utilized process model in the world to measure the impact of learning. Hundreds of case studies have been published on the process,

Figure 10-3. Model of the ROI Process: Calculating the ROI of a Business Performance Solution



and almost four thousand individuals have been certified to implement this process within their organization.

An important feature of this model is the number of techniques available to isolate the effects of the learning solution from other influences. In almost any setting, there will be multiple influences on output data, and the process model must utilize several techniques to measure or isolate the effects of learning from the other influences. An important part of the process model is the multiple ways in which to convert data to monetary values. As more situations require ROI analysis, this step becomes essential. Data, driven by the program, is converted to monetary values and compared with the actual cost of the program to yield the actual ROI. The results, representing six types of measures, are reported to specified target audiences.

Operating standards are necessary, ensuring consistency in utilizing the approach. These standards ensure that the results of an impact study conducted by one individual can be replicated if compared with the same study conducted by another individual. Table 10-2 shows the operating standards for the ROI process model described here. These "guiding principles" provide basic philosophies and operating rules that not only ensure consistency but also offer a conservative approach to analysis. These principles also provide the rationale for several cost-saving approaches in the analysis, while maintaining process credibility.

Selecting the appropriate programs for impact and ROI evaluation can often be a tedious process in and of itself. But it is necessary to ensure the successful implementation of a comprehensive measurement and evaluation process. When the wrong programs are selected, the evaluation is not only costly but can also be frustrating if the process and results are ineffective—or worse, completely ignored by key stakeholders.

The selection process begins with a review of the learning and development budget. To ensure that evaluation up to such a comprehensive level is feasible within the constraints of the learning and development measurement and evaluation budget (usually 3 to 5 percent of the

Table 10-2. ROI Methodology Guiding Principles: The Standards

1. When a higher-level evaluation is conducted, data must be collected at lower levels.
2. When an evaluation is planned for a higher level, the previous level of evaluation does not have to be comprehensive.
3. When collecting and analyzing data, use only the most credible sources.
4. When analyzing data, select the most conservative alternative for calculations.
5. At least one method must be used to isolate the effects of the solution/program.
6. If no improvement data is available for a population or from a specific source, the assumption is that no improvement has occurred.
7. Estimates of improvements should be adjusted for the potential error of the estimate.
8. Extreme data items and unsupported claims should not be used in ROI calculations.
9. Only the first year of benefits (annual) should be used in the ROI analysis of short-term solutions.
10. Costs of a solution, project, or program should be fully loaded for ROI analysis.
11. Intangible measures are defined as measures that are purposely not converted to monetary values.
12. Communicate the results of the ROI methodology to all key stakeholders.

total learning and development expenditures), a target-setting process is developed to reflect the actual percentage of programs evaluated at each level. This procedure often begins with 100 percent of all programs evaluated at Level 1. (For the details on Level 0 through Level 5, see appendix A.) These evaluations are usually automated, easy to accomplish, and contain very important information needed to make adjustments quickly. As measurement moves up the chain of impact, the evaluation process becomes more difficult and expensive; thus, a declining percentage of programs is evaluated at higher levels to make evaluation economically feasible. Figure 10-4 shows sample evaluation targets for a large telecommunication company. This distribution is typical of what many learning and development functions are pursuing.

Although these percentages can vary with a particular organization, they almost always reflect declining percentages, often leaving a question about the appropriateness of evaluating so few programs at Levels 4 and 5. However, if those programs are carefully selected following predetermined criteria developed with input from the management team, the necessity for

Figure 10-4. Sample Evaluation Targets for a Large Telecommunications Company

Level	Target (percent of programs evaluated at this level)
1 Reaction	100
2 Learning	70
3 Application	30
4 Business impact	10
5 ROI	5

having a greater number of evaluations is often precluded. The typical criteria utilized for selecting programs at these levels are

- the life cycle of the program
- the linkage of the program to operational goals and issues
- the importance of the program to strategic objectives
- the cost of the program
- the visibility of the program
- the size of the target audience
- the investment of time
- top executives are interested in the evaluation.

Though these criteria are common, the list can be expanded or modified for a particular learning and development function. Because impact studies are expensive and difficult, it is necessary to focus evaluations at this level only on the programs designed to make a significant difference, those that drive business value, or those that are important to the management group. This keeps the measurement and evaluation process reasonable and feasible within economic parameters.

Collectively, these building blocks provide the foundation and rationale needed to develop a comprehensive measurement system.

The Scorecard

The data reported in the scorecard (figure 10-2) is derived from the measurements captured from the processes described in the above sections. This section provides more detail on building the blocks of the scorecard itself.

Inputs and Indicators

The traditional approach to measuring the learning and development function is to report on inputs and indicators. Though these measures are important, they do not reflect the results, only the level of commitment, volume, efficiencies, and trends in processes. And though the number of indicators is vast, it is important to also include measures of interest to top managers. Ideally, the management group should approve the indicators, and the inputs reported should stimulate interest with executives. Here are a few possibilities:

1. The number of employees participating in learning programs.
2. The number of hours of learning activity per employee.
3. Various enrollment statistics, including the demographics of participants, participation rates, completion rates, and so on.
4. Investment in learning reported in a variety of ways (total costs, cost per employee, direct cost per participant, and cost as a percentage of payroll are common ways).
5. Cost recovery, if there is a charge-back system. (Some learning and development functions operate in a charge-back or profit-center mode.)
6. Status of alternative delivery of learning.

Several other statistics can be reported on issues such as technology, on-the-job learning, trends, volume, and efficiencies. Any mix is appropriate to highlight and monitor an important trend. The indicators show the degree of management's commitment to the learning process and provide a brief view of the mix of programs offered.

Reaction (Level 1)

The first outcome level, reaction, represents an important area of measurement. This is the most popular level of measurement, often used in measuring almost 100 percent of programs. The 100 percent coverage does not necessarily reflect importance as much as it does the ease of measuring and the cost-efficient way in which the data can be tabulated. Some argue that 100 percent is not needed; a sampling is sufficient. However, other factors may drive the 100 percent coverage. If planned action is a part of the data collection at this level, it is important to capture planned action from all the participants. Also, some participants want to provide feedback data, particularly if they had an exceptionally satisfying or extremely disappointing learning event.

To enable a comparison of one program with another, it is necessary to capture input for several specific reaction measures. Though specific items can vary with the learning and development function, seven items are suggested:

- relevance to the job*
- amount of new information
- recommendation to others*
- importance of the information*
- intention to use skills and knowledge*
- effectiveness of the facilitator
- effectiveness of the delivery system.

These items represent some of the most useful reactions. Also, four of the seven (followed with an asterisk) have been found to have a significant correlation with application and serve as predictors of actual applications. These items can be easily compared from one program to another, thus enabling the interaction of data across programs.

Another potential Level 1 measure is the percentage of the participants with action plans. Because of the interest in participants applying what they learn, this has become an important measure. Also, because there is some concern that transfer will not take place, the presence of an action plan enhances the possibility of the actual application. This measure will usually correlate with the extent of the actual application.

The next set of potential measures focuses on the forecasting capability of several learning and development functions. Data can be collected with the reaction questionnaire to forecast application, impact, and ROI. Additional questions can be added to the standard reaction questionnaire to generate the information. Figure 10-5 shows the types of questions necessary to develop projections for Levels 3, 4, and 5. The form can vary and be expanded and is often included as a supplemental form. When a significant number of forecasted ROI calculations are developed, the average forecasted ROI with an additional adjustment is sometimes reported. This adjustment represents an additional reduction factor other than the confidence factor from the individuals providing input. This adjustment represents the unknown barriers that will inhibit the transfer of the skill to the work setting. Sometimes this factor can be as high as 60 percent and is usually based on experience within the organization.

Learning (Level 2)

The measurement of learning presents a challenge for many comprehensive learning and development functions, because they attempt to determine the extent to which participants acquire skills and knowledge in learning solutions. A first step in this category is to report the percentage of programs evaluated at this level. This percentage often ranges from 40 to 80, depending on the definition of a learning measure. Some learning

Figure 10-5. Sample of the Types of Questions Necessary to Develop Projections for Levels 3, 4, and 5

Planned Improvements

- As a result of this program, what specific actions will you attempt as you apply what you have learned?
- Please indicate what specific measures, outcomes, or projects will change as a result of your actions.
- As a result of the anticipated changes in the above, please estimate (in monetary values) the benefits to your organization over a period of one year.
- What is the basis of this estimate?
- What confidence, expressed as a percentage, can you put in your estimate? (0% = No confidence; 100% = Certainty) _____%

measurements are very formal, such as objective tests, simulations, and structured skill demonstrations. Others are less structured and informal, such as self-assessments, facilitator assessments, and team assessments. Learning measurements should be defined in this context and reported appropriately. Some learning and development functions define a learning measure as any type of measurement, whether formal or informal. Others make a distinction between the types of measurement, and this is often reported on the scorecard. Table 10-3 provides a simple breakdown of the different types of measurement processes that can provide insight into the extent of formal versus informal assessments.

To compare learning changes from one program to another and integrate the data into a macro view on the scorecard, it is essential to have only a few identical measures. Typical issues addressed in a self-assessment of learning change are

- acquisition of the skills and knowledge
- ability to use the skills and knowledge
- confidence in the use of skills and knowledge.

The scale is adjusted for each program so that exceptional ratings (that is, 5) are the same across programs.

Two other possibilities sometimes surface that can be included in this section of the scorecard. When available, pretest and posttest differences can be reported to show the percentage of improvement. Because this is a percentage improvement, an average measure for similar programs

Table 10-3. Types of Measurement Processes for Formal versus Informal Assessments

Formal Measures	Informal Measures
Objective tests	Self-assessments
Performance testing	Team assessments
Simulations	Facilitator assessments

provides insight into the extent of learning change for all those programs. Also, when some programs are designed to lead to a certification or qualification, the percentage of participants actually meeting requirements is often reported.

Application (Level 3)

To measure the actual change in behavior on the job and progress with application, measurements must be taken after the program is completed. The first in this category of measures is the percentage of the programs evaluated at this level (3). This percentage often ranges from 20 to 50, depending on the resources available and the types of follow-up evaluations planned. Three issues are recommended as standard follow-up items when assessing changes in skills and knowledge:

- the extent of use of the skill or knowledge on the job
- the frequency of use of new skill or knowledge (on the job)
- the effectiveness of the skill or knowledge (as applied on the job).

Essentially, these items can be collected on every program to show the extent to which the skills and knowledge are applied on the job. With this approach, exceptional performance (again, a rating of 5) is the same across programs, although what constitutes exceptional performance would vary with the program. For example, for some programs, exceptional performance may be defined as using the skill several times every day. For other programs, exceptional performance may require the use of the skills once a month. The important point is that the scale is adjusted with the program so that ratings can be compared across programs.

The next recommended measures can be very critical to the success of the learning and development function. The first of the four measures is to capture the percentage of the action plans completed by the desired follow-up time. This provides data on the extent to which participants apply and complete their assignments from the program. When compared with baseline data, this measure shows changes in transfer of learning to the job. The next measure captures the barriers to skill and knowledge application on the job. Forced-choice options for anticipated barriers are

usually offered, along with space for additional barriers. Forced choices for typical barriers allow for integration across programs and can be arranged as a "top 10" list.

Table 10-4 shows the typical barriers. The enablers are those processes that enhance the transfer of learned skills and knowledge. The enablers provide insight into the reasons for success and sometimes mirror the barriers. A similar forced-choice option can be utilized with the enablers. The same process can be used to capture, tabulate, and integrate data on enablers as used with data on barriers.

A management support profile reveals the extent to which immediate managers support the programs. A management support profile taken from both the manager and the participant provides interesting information about the quality of support. Table 10-5 shows a range of follow-up questions on management support from the participants' perspective. Participants and managers are asked to check the appropriate statement that best describes the level of management support provided. Managers have their statements worded slightly different. When data is collected across programs and ranked, much can be revealed about the level of management support. Disconnects can be quickly revealed and converted into actions for improvement.

By listing the top 10 barriers, enablers, and management support issues, useful indices and trends are provided for management to see what helps

Table 10-4. Typical Barriers Listed on Follow-Up

The immediate manager does not support the skills/knowledge.
The culture in the work group does not support the skills/knowledge.
There is no opportunity to use the skills/knowledge.
There is no time to use the skills/knowledge.
The skills/knowledge could not be applied to the job.
The systems and processes did not support the use of the skills/knowledge.
Changed job and the skills/knowledge no longer apply.
The skills/knowledge taught are not appropriate in our work unit.
There is no need to apply what was learned.
Could not change old habits.

Table 10-5. Management Support Response: Options for Follow-Up Questions

My manager asked me to forget these skills—we operate differently here.
My manager questioned the appropriateness of the skills/knowledge.
My manager complained about my absence and expressed concern about the program.
My manager made no comment about the program.
My manager asked about my reaction about the program.
My manager encouraged my use of the skills/knowledge.
My manager coached my use of the skills/knowledge.

or hinders the implementation of training. Trends are monitored and the data is turned into action plans to reduce or remove impediments to success and to enhance the enablers of application.

The next measure in this category is the data collection technique. This item provides input into the different methods utilized to capture both application and business impact data. Objective versus subjective processes can be emphasized. Some organizations are moving from the subjective to the objective, while others are moving in the other direction. Either way, the tracking of techniques shows the data collection trends in learning and development. Table 10-6 shows the typical options available for postprogram data collection.

Business Impact (Level 4)

The connection to business measures can be captured when studies are conducted to show business impact or when follow-up questionnaires are utilized to capture application data. In this category of measures, the first measure is the percentage of programs evaluated at this level. The percentage is usually around 10 to 20 when all the processes are in place to produce an impact study. However, this percentage can increase significantly if questionnaires, interviews, and focus groups are modified to capture business impact data, in addition to application data. When summarized and reported, this becomes important data.

Table 10-6. Methods of Collecting Postprogram Data

Method	Level 3	Level 4
Surveys	✔	
Questionnaires	✔	✔
Observation on the job	✔	
Interviews	✔	
Focus groups	✔	
Action planning	✔	✔
Performance contracting	✔	✔
Performance monitoring		✔

As shown in figure 10-6, the linkage of the program to business measures can be captured on a follow-up questionnaire. Only responses where there is a significant and very significant influence are tabulated. Impact study data, showing the movement of business impact measures, is reported on a different line of the report. Either way, the connection between the program and the business measure is reported. When presented as the top 10 list, it is possible to quickly examine disconnects or alignment issues. For example, if the most important business strategy is to improve customer satisfaction, then the program's connection to customer satisfaction measures should surface at the top of the list. If it does not, disconnects may be occurring or there could be misalignment.

Another measure reported on the learning scorecard is the method used to isolate the effects of the program. A variety of ways is available, and some organizations are moving to more research and analytical methods while others are using subjective estimations to save time and costs. Changes in the use of methods are identified as the different techniques are highlighted. Table 10-7 shows the possible techniques for isolating the effects of a program.

Return on Investment (Level 5)

From the viewpoint of many executives, the ultimate level of evaluation is the actual return on investment, where the monetary benefits are

Figure 10-6. Sample Follow-Up Questionnaire on Linkages to Business Impact Measures

Indicate the extent to which you think this program has influenced each of these measures in your work unit, department, or business unit:

	No Influence	Some Influence	Moderate Influence	Significant Influence	Very Significant Influence
Productivity					
Quality					
Response time					
Cost control					
Employee satisfaction					
Customer satisfaction					
Other (please specify)					

compared with the cost of the program. The first measure in this ROI category is the percentage of programs evaluated at this level. This number is usually quite low, usually in the 5 to 10 percent range. Not much additional information is needed on the learning scorecard for the ROI studies, because the number of studies is quite small.

The results of ROI studies are typically reported to target audiences, and only summary data is needed on the scorecard. A brief paragraph showing the nature of the study and the actual results, including the

Table 10-7. Techniques to Isolate the Effects of the Programs

Use of a control group arrangement
Trend line analysis of performance data
Use of forecasting methods of performance data
Participants' estimate of programs impact (percent)
Supervisors' estimate of programs impact (percent)
Management's estimate of programs impact (percent)
Use of previous studies and experts
Calculating or estimating the impact of other factors
Use of customer input

ROI percentage, is appropriate, along with the information on how to obtain additional details. This category includes two other measures: the method used to convert data to monetary values, and the cost per participant. Because the data conversion methods can vary, it is important to show the methods utilized so that trends can be tracked. Many organizations are moving toward the use of standard values and the use of estimates so that the cost to develop values can be minimized. The different approaches to convert data to monetary values are

- Convert output to profit or cost savings.
- Convert the cost of quality to money saved.
- Convert employees' time to compensation.
- Use historical costs.
- Use internal and external experts.
- Use data from external databases.
- Link with other measures.
- Use participants' estimates.
- Use supervisors' and managers' estimates.
- Use learning and development teams' estimates.

The cost per participant can be as important a measure for the learning scorecard as it is for the ROI studies. This value represents a fully loaded cost and is different from the cost per employee for training reported in the indicator category, which is usually a direct cost. The total program costs categories are

- needs assessment (prorated)
- development costs (prorated)
- program materials
- instructor/facilitator costs
- facilities cost
- travel/lodging/meals
- participant salaries (and benefits) for time away from work
- administrative/overhead costs
- evaluation costs.

This total cost is then divided by the number of participants in the program. The program cost per participant is reported on the scorecard and can be compared with the direct cost of the program.

Intangible Measures

An additional category of important data is intangible measures. These measures may be included in the business impact category of a report or mentioned in a separate section, particularly if there are significant numbers of impact studies. Intangibles are defined as the measures that are purposefully not converted to monetary values, although they are very important to the organization and there is evidence of linkage to the learning and development programs.

The top 10 intangibles are sometimes reported, providing an opportunity to check for alignment and disconnects with organizational strategy. The most important intangibles in the organization—ranging from teamwork to networking to brand awareness to customer satisfaction—are listed as questions for participants to complete. On the follow-up evaluation on Level 3, this question may appear listed this way: "To what extent did this program influence this measure?" The reaction could be on a scale from 1 to 5, where 1 is no influence and 5 is a very significant influence. When these questions are combined on all of the follow-ups, a profile is revealed of the intangibles driven by the programs. The intangible driven most becomes the most important and influential for the programs, and so forth. This approach is easy to do and provides some insightful information about how the programs are linked to these important measures.

Awards

Today, many awards are made available to learning and development functions and provide recognition for some of the best functions. Some are professional awards provided by professional organizations, such as the American Society for Training and Development. Magazines, such as *Training* and *Chief Learning Officer*, provide others. Still others are through

conference organizers, such as the Corporate University Best in Class Awards. In addition, the industry in which an organization operates sometimes provides awards.

Achieving awards and listing them on the learning scorecard can be helpful, but particularly for executives who want external recognition. As you recall, our survey shows that "awards" was ranked 3 out of 8 by those responding. After all, it's usually hard work and excellent progress that brings this type of recognition. And when the award is obtained consistently, it shows a sustained best practice approach. The downside of awards is that the basis for achieving them is sometimes not very objective. Unfortunately, many awards are given based on the quality of the entry or nomination. Some heads of learning and development functions have boasted about being able to win any award that's available (as he or she has) because of the appropriate resources behind the awards. Still, they are worth pursuing and should be considered for the overall scorecard.

The Learning Scorecard—Short Version

Some senior executives are frightened to see the comprehensive nature of the learning scorecard presented in this chapter. This is particularly true when previous scorecards or attempts to develop a scorecard have been dominated by input measures, as many of them are. When this is the case, perhaps moving to the shorter version would be a first step. Figure 10-7 shows a short version of the comprehensive scorecard described in this chapter and contains only 15 measures, but all are very powerful and important to the executive group. Most of the measures are self-explanatory and reflect the measures presented in this chapter; however, a few deserve some explanation.

In the area of reaction, the perceived value of the program can be captured, even as an index, which is an average of several questions that might always be asked, and presented to your team as one measure. Also, the business impact linkage can be obtained during the program, instead of waiting until the follow-up. In this case, a question is added in the

Figure 10-7. Short Version of the Learning Scorecard

0 Inputs/indicators
1. Number of programs
2. Hours per employee
3. Cost per employee

1 Reaction
1. Perceived value (index)
2. Business impact linkage

2 Learning
1. Percent of programs evaluated at this level
2. Learning (index)

3 Application
1. Percent of programs evaluated at this level
2. Application index
3. Top barriers to learning transfer

4 Business impact
1. Percent of programs evaluated at this level
2. Business impact linkage

5 ROI
1. Percent of programs evaluated at this level

Intangibles
1. Top intangibles

Awards
1. Industry and professional awards

reaction questionnaire that lists all the key business impact measures in the organization with a statement that says, "Please provide the extent to which this program will influence this measure." A scale of 1 to 5 is used, with 1 being no influence and 5 being very significant influence. In essence, this is the very same question asked on the follow-up, but here it's collected almost 100 percent of the time and shows perceived linkage immediately.

Under learning, the index can be constructed using one or two measures but presented as an average. It could be an average of the three presented in this chapter, or it could be a single measure. On number three, the application index is an average of the three suggested in this chapter. The rest of the measures are self-explanatory.

Reporting Data

Although simple and straightforward, the data outlined in the learning scorecard can be quite cumbersome when several hundred programs involving thousands of participants are conducted each quarter. An important issue is to use consistent rating scales and consistent questions and statements. Also, the only way that complete data can be accumulated and reported economically is through the use of technology. If reaction and learning data is collected with a questionnaire during or at the end of a learning program, the data can be quickly integrated and available instantaneously.

In addition, when a standard follow-up questionnaire is utilized to capture application and impact data, most of the data items at these levels can be integrated. For higher levels of evaluation (3, 4, and 5), data volume may become smaller because convenience sampling is utilized and the task should be feasible and reasonable. Ideally, if data is collected online, it can be rolled up into an online reporting format. In the short version of the learning scorecard, the report can be developed showing only one or two measures for each of the seven categories. The fully loaded scorecard, presented in figure 10-2, can be developed over time as resources are available and data collection instruments are revised. The key is to begin with the end in mind with a mockup of the report, presented to the executive group. Both reports provide a tremendous amount of data that shift reporting from being input focused to output focused. The frequency of reporting can be monthly, quarterly, or annually—the more frequent, the better. Online reporting should be an ultimate goal.

Challenges

Three main challenges must be addressed as CLOs move from the traditional activity-based reporting focused on input to a comprehensive scorecard of successes based on output. Because a macro view of learning and development is needed, this shift is long overdue and can be accomplished with a reasonable amount of resources. The first challenge is to allocate additional resources for measurement and evaluation. The

learning scorecard can be developed, including all the measurement and evaluation processes for about 3 to 5 percent of the total learning and development budget. This is not an unreasonable amount when the value of the process is considered. Also, this additional investment can pay off in significant cost savings.

The second challenge is to approach the task in a disciplined, methodical manner. This is particularly difficult when it is not required (directly) by executives. Waiting for the executive request for more of a contribution on the scorecard (and less input data) may be too late. The best time to pursue this is when it is not mandated by executives.

The third important challenge is the actual use of the data. Not only does this reporting reveal success in terms of the contribution, but it also provides a tremendous amount of process improvement data that enables the learning and development team to make adjustments throughout the learning cycle. All the data should be collected with its ultimate use in mind. If it is generated only to fill up boxes in a chart or lines on a report, it will become worthless. Because of the possibilities, the executive group and the learning and development team need to decide how the data can best be used and interpreted.

Final Thoughts

The good news is that the challenge of reporting useful data that measures learning and development success and contribution is feasible. Success is possible in almost any setting, but it is most likely with careful planning, a framework that focuses on results, a determination to make it work, and perhaps a few additional resources. Most CEOs report that they don't see a learning and development scorecard. This is a great opportunity to develop a learning scorecard for the executives—but only if it contains data the executives want to see.