Building a Comprehensive Evaluation Process

Measuring and evaluating learning has earned a place among the critical issues in the learning and development and performance improvement fields. For decades, this topic has been on conference agendas and discussed at professional meetings. Journals and newsletters regularly embrace the concept, dedicating increased print space to it. Professional organizations have been created to exchange information on measurement and evaluation, and more than twenty-five books provide significant coverage of the topic. Even top executives have an increased appetite for evaluation data.

Although interest in the topic has heightened and much progress has been made, it is still an issue that challenges even the most sophisticated and progressive learning and development departments. While some professionals argue that having a successful evaluation process is difficult, others are quietly and deliberately implementing effective evaluation systems. The latter group has gained tremendous support from the senior management team and has made much progress. Regardless of the position taken on the issue, the reasons for measurement and evaluation are intensifying. Almost all learning and performance improvement professionals share a concern that they must show the results of learning investments. Otherwise, funds may

be reduced or the department may not be able to maintain or enhance its present status and influence within the organization.

The dilemma surrounding the evaluation of learning is a source of frustration with many senior executives—even within the field itself. Most executives realize that learning is a basic necessity when organizations experience significant growth or increased competition. They intuitively feel that providing learning opportunities is valuable, logically anticipating a payoff in important bottom-line measures, such as productivity improvements, quality enhancements, cost reductions, time savings, and improved customer service. Yet the frustration comes from the lack of evidence to show that programs really work. While results are assumed to exist and learning programs appear to be necessary, more evidence is needed, or executives may feel forced to adjust funding in the future. A comprehensive measurement and evaluation process represents the most promising, logical, and rational approach to show this accountability. This book shows how to measure the contributions of learning and development and performance improvement programs.

KEY QUESTIONS

When individuals pursue a comprehensive process, they often have anxiety, issues, and concerns. They have important questions that they want resolved. Exhibit 1.1 shows a list of the typical questions that in-

- How can I move up in the evaluation chain?
- How can I collect data efficiently?
- What data should be collected at each level?
- How can I design a practical evaluation strategy that has credibility with all stakeholders?
- What support do I need for evaluation?
- How can I integrate data in a management scorecard?
- How should evaluation data be used?
- How can I get the internal support to design and implement my evaluation strategy?
- How can I proceed if the evaluation reveals an unacceptable result?
- How can I develop practical and credible tests?
- How can I use the evaluation process to implement a result-based philosophy?
- How can I make cost-effective decisions at each evaluation level?
- How can I convince clients that my program is linked to business performance measures?

Exhibit 1.1. Typical Questions.

dividuals face, regardless of the type of organization or the organization's stage of growth and development. Each of these issues, as well as many others, is detailed in this book. Each question is covered with responses that can help resolve many measurement and evaluation system challenges.

Global Evaluation Trends

Measurement and evaluation have been changing and evolving—in both the private and public sectors—across organizations and cultures, not only in the United States, but across all developed countries. The following trends have been identified:

- Organizations are increasing their investments in measurement and evaluation with best practice groups spending 3 to 5 percent of the learning and development budget on measurement and evaluation.

- Organizations are moving up the value chain, away from measuring reaction and learning to measuring application, impact, and occasionally ROI.

- The increased focus on measurement and evaluation is largely driven by the needs of the clients and sponsors of learning projects, programs, initiatives, and solutions.

- Evaluation is an integral part of the design, development, delivery, and implementation of programs.

- A shift from a reactive approach to a proactive approach is occurring, with evaluation being addressed early in the cycle.

- Measurement and evaluation processes are systematic and methodical, often designed into the delivery process.

- Technology is significantly enhancing the measurement and evaluation process, enabling large amounts of data to be collected, processed, analyzed, and integrated across programs.

- Evaluation planning is becoming a critical part of the measurement and evaluation cycle.

- The implementation of comprehensive measurement and evaluation processes usually leads to increased emphasis on initial needs analyses.

- Organizations with comprehensive measurement and evaluation systems in place have enhanced their program budgets.

- Organizations without comprehensive measurement and evaluation systems have reduced or eliminated their program budgets.
- The use of ROI is emerging as an essential part of many measurement and evaluation systems. It is a fast-growing metric—70 to 80 percent of organizations have it on their wish lists.
- Many successful examples of comprehensive measurement and evaluation applications are available in all types of organizations and cultures.

These trends are creating a never-ending appetite for more information, resources, knowledge, and skills in the measurement and evaluation process.

Measurement and Evaluation Challenges

Although measurement and evaluation are increasing, why aren't organizations doing more? Even though the need is evident, establishing the process can be formidable. Twelve basic barriers to conducting meaningful evaluations exist. They are described in detail below:

1. TOO MANY THEORIES AND MODELS. Since Kirkpatrick published articles on the four levels of evaluation in the late 1950s, dozens of evaluation books have been written, primarily for the social sciences, education, and government organizations. Then add the twenty-five-plus models and theories for evaluation offered to practitioners to help them measure the contributions of learning and development, each claiming a unique approach and a promise of addressing evaluation woes. Choosing the model to follow can seem as daunting as establishing world peace.

2. MODELS ARE TOO COMPLEX. Evaluation can be a difficult issue. Because situations and organizations are different, implementing an evaluation process across multiple programs and organizations is sometimes quite complex. The challenge is to develop models that are theoretically sound, yet simple and usable.

3. LACK OF UNDERSTANDING OF EVALUATION. Learning this process hasn't been easy for practitioners. Some books on the topic have over

six hundred pages, making absorbing the information through reading alone impossible for practitioners. Not only is understanding the evaluation processes essential for the evaluator, but the entire learning and development staff must learn parts of the process and understand how they fit into their roles. To remedy this situation, the organization must focus on how expertise is developed and disseminated within the organization.

4. THE SEARCH FOR STATISTICAL PRECISION. The use of complicated statistical models is confusing and difficult to absorb for most practitioners. Statistical precision is needed when a high-stakes decision is being made and when plenty of time and resources are available. Otherwise, simple statistics are appropriate.

5. EVALUATION IS OFTEN CONSIDERED A POST-PROGRAM ACTIVITY. When evaluation is considered an add-on activity, it loses the power to deliver the needed results. The most appropriate way to use evaluation is to consider it early—prior to program development—at the time of conception. With this approach, an evaluation is conducted efficiently and the quality and quantity of data collected are enhanced.

6. FAILURE TO SEE THE LONG-TERM PAYOFF OF EVALUATION. Developing the long-term payoff of evaluation requires examining multiple rationales for pursuing evaluation. Evaluation can be used to:

- Determine success in accomplishing learning and performance improvement program objectives.
- Prioritize resources for learning and performance improvement programs.
- Enhance the accountability of learning and performance improvement programs.
- Identify the strengths and weaknesses of the learning and development process.
- Compare the costs to the benefits of programs.
- Decide who should participate in future programs.
- Test the clarity and validity of tests, cases, and exercises.
- Identify which program participants were most successful.
- Reinforce major points made to the participants.

- Improve the quality of learning and development.
- Assist in marketing future programs.
- Determine whether the program was an appropriate solution for the specific need.
- Establish a database that can assist management in decision making.

7. LACK OF SUPPORT FROM KEY STAKEHOLDERS. Important customers, those who need and use evaluation data, sometimes don't provide the support needed to ensure the process's success. Specific steps must be taken to win support and secure buy-in from key groups, including senior executives and the management team. Executives must see that evaluation produces valuable data to improve programs and validate results. When the stakeholders understand what is involved, they may offer more support.

8. EVALUATION HAS NOT DELIVERED THE DATA SENIOR MANAGERS WANT. Today, clients and sponsors ask for data beyond reaction and learning. They need data on the application of new skills on the job and the corresponding impact of this application on the business units. Sometimes, they want ROI data for major programs. They request data about the business impact of learning—both from short-term and long-term perspectives. Ultimately, these are the executives who fund learning and development. If the desired data are not available, future funding could be in jeopardy.

9. IMPROPER USE OF EVALUATION DATA. Improper use of evaluation data can lead to four major problems, which are described here:

1. Too many organizations do not use evaluation data at all. Data are collected, tabulated, catalogued, filed, and never used by any group, other than the individual who initially collected the data.
2. Data are not provided to the appropriate audiences. Analyzing the target audiences and determining the specific data needed for each group are important steps for communicating results.
3. Data are not used to drive improvement. If not part of the feedback cycle, evaluation falls short of its intended purpose.
4. Data are used for the wrong reasons—to take action against an individual or group or to withhold funds, rather than to im-

prove processes. Sometimes, the data are used in political ways to gain power or advantage over another person.

10. LACK OF CONSISTENCY. For evaluation to add value and be accepted by stakeholders, it must be consistent in its approach and methodology. Tools and templates have to be developed to support the method of choice to prevent perpetual reinvention of the wheel. Without this consistency, evaluation consumes too many resources and raises too many concerns about the quality and credibility of the process.

11. LACK OF STANDARDS. Closely parallel with consistency is the issue of standards. Standards are rules for making evaluation consistent, stable, and equitable. Without standards, little credibility in processes and stability of outcomes exists.

12. SUSTAINABILITY. A new model or approach often has a short life and, therefore, is not sustained. Evaluation must be integrated into the organization so that it becomes routine and lasting. To accomplish this, the evaluation process must gain the respect of key stakeholders at the outset. The evaluation process must be well documented, and stakeholders must accept their responsibilities to make it work. Without sustainability, evaluation will be on a roller-coaster ride, where data are collected only when programs are in trouble and less attention is provided when they are not.

Benefits of Measurement and Evaluation

Although the benefits of measurement and evaluation may appear obvious, several distinct and important payoffs can be realized.

RESPOND TO REQUESTS AND REQUIREMENTS. Today's executives and administrators need information about application and implementation in the workplace and their corresponding impacts on key business measures. In some cases, they ask for an ROI analysis. Developing a comprehensive measurement and evaluation system is the best way to meet these requests and requirements.

JUSTIFY BUDGETS. Some learning and performance functions use evaluation data to support a requested budget, while others use the data to prevent the budget from being slashed, or in drastic cases, eliminated entirely. Additional evaluation data can show where programs

add value and where they do not. This approach can lead to protecting successful programs as well as pursuing new programs.

IMPROVE PROGRAM DESIGN. A comprehensive evaluation system should provide information to improve the overall design of a program, including the areas of learning design, content, delivery method, duration, timing, focus, and expectations. These processes may need adjustment to improve learning, especially during new program implementation.

IDENTIFY AND IMPROVE DYSFUNCTIONAL PROCESSES. Evaluation data can determine whether the up-front analysis was conducted properly, thereby aligning the program with the organizational needs. Additional evaluation data can help pinpoint inadequacies in implementation systems and identify ways to improve them.

ENHANCE LEARNING TRANSFER. Learning transfer is perhaps one of the biggest challenges that the learning and development field faces. Research shows that 60 to 90 percent of job-related skills and knowledge acquired in a program are not implemented on the job. A comprehensive evaluation system can identify specific barriers to implementing learning. Evaluation data can also highlight supportive work environments that enable learning transfer.

ELIMINATE UNNECESSARY OR INEFFECTIVE PROGRAMS. Program evaluation can provide rational, credible data to help support the decision to implement or discontinue a program. In reality, if the program cannot add value, it should be discontinued. *One caveat:* Eliminating programs should not be a principal motive or rationale for increasing evaluation efforts. Although it is a valid use of evaluation data, program elimination is often viewed negatively.

EXPAND OR IMPLEMENT SUCCESSFUL PROGRAMS. The flip side of eliminating programs is expanding their presence or application. Positive results may signal the possibility that a program's success in one division or region can be replicated in another area, if a similar need exists.

ENHANCE THE RESPECT AND CREDIBILITY OF THE STAFF. Collecting and using evaluation data—including application, impact, and ROI—builds respect for learning and respect for the learning and perfor-

mance staff. Appropriate evaluation data can enhance the credibility of the learning and development and performance improvement functions when the data reveal the value added to the organization.

SATISFY CLIENT NEEDS. Satisfying clients is a critical objective for the learning and performance improvement function. If clients are not pleased with the data, they may decline the opportunity to use the staff in the future. If they are satisfied, they may use the program again and even recommend the program to others.

INCREASE SUPPORT FROM MANAGERS. Participants' immediate managers need convincing data about the success of learning. They often do not support learning processes because they see little value in taking employees away from the job to be involved in a program with little connection to their business units. Data showing how learning helps them achieve their objectives will influence their support.

STRENGTHEN RELATIONSHIPS WITH KEY EXECUTIVES. Senior executives must perceive the learning and development staff as business partners who should be invited to the table for important decisions and meetings. A comprehensive measurement and evaluation process can show the contribution of the function and help strengthen this relationship.

SET PRIORITIES FOR LEARNING AND DEVELOPMENT. A comprehensive measurement system can help determine which programs represent the highest priority. Evaluation data can show the payoff or potential payoff of important and expensive programs—the programs that support strategic objectives.

REINVENT LEARNING AND DEVELOPMENT AND PERFORMANCE IMPROVE-MENT. Measurement and evaluation reveal the link between learning and the business and can drive increased alignment in the future. Connecting learning with business objectives requires a continuous focus on critical organizational needs and results that can and should be obtained from programs.

ALTER MANAGEMENT'S PERCEPTIONS OF LEARNING AND DEVELOPMENT. Middle-level managers often see learning as a necessary evil. A comprehensive evaluation process may influence these managers to view learning as a contributing process and an excellent investment. It can

also help shift the perception of learning from a dispensable activity to an indispensable, value-adding process.

ACHIEVE A MONETARY PAYOFF. In some situations, an actual monetary value can be calculated for investing in measurement and evaluation. This is particularly true with implementation of ROI, and many organizations have even calculated "the ROI on the ROI process." They determine the payoff of investing in a comprehensive measurement and evaluation process—the ROI methodology. The payoff is developed by detailing specific economies, efficiencies, and direct cost savings generated by the evaluation process.

These key benefits, inherent with almost any type of evaluation process, make additional measurement and evaluation an attractive challenge of the learning and performance function.

The Myths of Measurement and Evaluation

Practitioners recognize that additional measurement and evaluation is needed. However, regardless of the motivation to pursue evaluation, they struggle with how to address the issue. They often ask, "Does it really provide the benefits to make it a routine, useful tool?" "Is it feasible within our resources?" "Do we have the capability of implementing a comprehensive evaluation process?" The answers to these questions often lead to debate and controversy. Controversy stems from misunderstandings about what the additional evaluation can and cannot do and how it can or should be implemented in organizations. The following is a list of myths, including the appropriate clarifications:

MEASUREMENT AND EVALUATION, INCLUDING ROI, IS TOO EXPENSIVE. When considering additional measurement and evaluation, cost is usually the first issue to surface. Many practitioners think that evaluation adds cost to an already lean budget that is regularly scrutinized. In reality, when the cost of evaluation is compared to the budget, a comprehensive measurement and evaluation system can be implemented for less than 5 percent of the total direct learning and development or performance improvement budget.

EVALUATION TAKES TOO MUCH TIME. Parallel with the concern about cost is the time involved in evaluation—time to design instruments,

collect data, process the data, and communicate results to the groups that need them. Dozens of shortcuts are available to help reduce the total time required for evaluation.

SENIOR MANAGEMENT DOES NOT REQUIRE IT. Some learning and development staff think that if management does not ask for additional evaluation and measurement, the staff does not need to pursue it. Sometimes, senior executives fail to ask for results because they think that the data are not available. They may assume that results cannot be produced. Paradigms are shifting, not only within learning and performance improvement, but within senior management groups as well. Senior managers are beginning to request higher-level data that shows application, impact, and even ROI.

MEASUREMENT AND EVALUATION IS A PASSING FAD. While some practitioners regard the move to more evaluation, including ROI, as a passing fad, accountability is a concern now. Many organizations are asked to show the value of programs. Studies show this trend will continue.

EVALUATION ONLY GENERATES ONE OR TWO TYPES OF DATA. Although some evaluation processes generate a single type of data (reaction-level, for example), many evaluation models and processes generate a variety of data, offering a balanced approach based on both qualitative and quantitative data. The process in this book collects as many as seven different types of qualitative and quantitative data, within different timeframes, and from different sources.

EVALUATION CANNOT BE EASILY REPLICATED. With so many evaluation processes available, this issue becomes an understandable concern. In theory, any process worth implementing should be one that can be replicated from one study to another. Fortunately, many evaluation models offer a systematic process, with certain guiding principles or operating standards to increase the likelihood that two different evaluators will obtain the same results.

EVALUATION IS TOO SUBJECTIVE. Subjectivity of evaluation has become a concern, in part because of the studies conducted using estimates and perceptions that have been published and presented at conferences. The fact is that many studies are precise and are not based on estimates. Estimates usually represent the worst-case scenario or approach.

IMPACT EVALUATION IS NOT POSSIBLE FOR SOFT-SKILL PROGRAMS. This concern is often based on the assumption that only technical or hard skills can be evaluated, not soft skills. For example, practitioners might find measuring the success of leadership, team-building, and communication programs difficult. What they often misunderstand is that soft-skills learning and development programs can, and should, drive hard-data items, such as output, quality, cost, and time.

EVALUATION IS MORE APPROPRIATE FOR CERTAIN TYPES OF ORGANIZATIONS. Although evaluation is easier in certain types of programs, generally, it can be used in any setting. Comprehensive measurement systems are successfully implemented in health care, nonprofit, government, and educational areas, in addition to traditional service and manufacturing organizations. Another concern expressed by some is that only large organizations have a need for measurement and evaluation. Although this may appear to be the case (because large organizations have large budgets), evaluation can work in the smallest organizations and simply must be scaled down to fit the situation.

IT IS NOT ALWAYS POSSIBLE TO ISOLATE THE EFFECTS OF LEARNING. Several methods are available to isolate the effects of learning on impact data. The challenge is to select an appropriate isolation technique for the resources available and the accuracy needed in the particular situation.

A PROCESS FOR MEASURING ON-THE-JOB IMPROVEMENT SHOULD NOT BE USED. This myth is believed because the learning and development staff usually has no control over participants after they leave the program. Belief in it is fading, though, as organizations realize the importance of measuring the results of workplace learning solutions. Systems and processes can be implemented to influence application. Expectations can be created so that participants anticipate a follow-up and provide data.

A PARTICIPANT IS RARELY RESPONSIBLE FOR THE FAILURE OF PROGRAMS. Too often, participants are allowed to escape accountability for their learning experiences. It is too easy for participants to claim that the program was not supported by their managers, it did not fit the culture of the work group, or that the systems or processes were in conflict with the skills and processes presented in the program. Today,

participants are held more accountable for the success of learning in the workplace.

EVALUATION IS ONLY THE EVALUATOR'S RESPONSIBILITY. Some organizations assign an individual or group the primary responsibility for evaluation. When that is the case, other stakeholders assume that they have no responsibility for evaluation. In today's climate, evaluation must be a shared responsibility. All stakeholders are involved in some aspect of analyzing, designing, developing, delivering, implementing, coordinating, or organizing a program.

SUCCESSFUL EVALUATION IMPLEMENTATION REQUIRES A DEGREE IN STATISTICS OR EVALUATION. Having a degree or possessing some special skill or knowledge is not a requirement. An eagerness to learn, a willingness to analyze data, and a desire to make improvements in the organization are the primary requirements. After meeting these requirements, most individuals can learn how to properly implement evaluation.

NEGATIVE DATA ARE ALWAYS BAD NEWS. Negative data provide a rich source of information for improvement. An effective evaluation system can pinpoint what went wrong so that changes can be made. Barriers to success as well as enablers of success can be identified. Such data will generate conclusions that show what must be changed to make the process more effective.

KEY STEPS AND ISSUES

Instead of examining a particular model or process, identifying some of the key issues, steps, and processes involved in measurement may be helpful. All these must be addressed in some way to have a comprehensive process.

Stakeholders

Many stakeholders are involved in comprehensive measurement and evaluation systems. A stakeholder is defined as any individual or group interested or involved in the program. Stakeholders may include the functional manager where the program is located, the participants, the organizer, the program leader, facilitators, and key clients, among

others. Below are descriptions of these stakeholders, and they will be referred to routinely throughout the book.

SPONSOR/CLIENTS. The individual(s) who fund, initiate, request, or support a particular project or program. Sometimes referred to as the sponsor, it is the key group—usually at the senior management level—who cares about the program's success and is in a position to discontinue or expand the program.

PARTICIPANTS. These are the individuals who are directly involved in the program. The term "employee," "associate," "user," or "stakeholder" may represent these individuals. For most programs, the term "participant" appropriately reflects this group.

IMMEDIATE MANAGERS. These are individuals who are one level above the participant(s) involved in the program. For some programs, this is the team leader for other employees. Often they are middle managers, but most important, these people have supervisory authority over the participants in the program.

CEO/MANAGING DIRECTOR/AGENCY EXECUTIVE. This person is the top executive in an organization. The top executive could be a plant manager, division manager, regional executive, administrator, or agency head. The CEO is the top administrator or executive in the operating entity where the program is implemented.

THE ORGANIZATION. The organization is the entity within which the particular program or process is evaluated. Organizations may be companies (either privately held or publicly held); government organizations at the local, state, federal, and international levels; non-profits; or non-governmental organizations. They may also include educational institutions, associations, networks, and other loosely organized bodies of individuals.

PROGRAM MANAGER. The individual(s) responsible for the project, program, initiative, or process. This is the individual who manages the program and is interested in showing the value of the program before it is implemented, during its implementation, and after it is implemented.

PROGRAM TEAM. The individuals involved in the program, helping to analyze and implement it. These are individual team members who may be full- or part-time on this particular program. For larger-scale programs, these individuals are often assigned full-time, on a temporary basis, or, sometimes, on a permanent basis. For small programs, these may be part-time duties.

EVALUATOR. This individual evaluates the program. This person is responsible for measurement and evaluation, following all the processes outlined in this book. If this is a member of the program team, extreme measures must be taken to ensure this person remains objective. It may also be a person who is completely independent of the program. This individual performs these duties full- or part-time.

FINANCE AND ACCOUNTING STAFF. These individuals are concerned about the cost and impact of the program from a financial perspective. They provide valuable support. Their approval of processes, assumptions, and methodologies is important. Sometimes, they are involved in the program evaluation; at other times they review the results. During major programs, this could include the organization's finance director or chief financial officer.

ANALYSTS. These individuals collect the data to determine whether the program is needed. They are also involved in analyzing various parts of the program. Analysts are usually more important in the beginning, but may provide helpful data throughout the program.

BYSTANDERS. The bystanders are the individuals who observe, sometimes at a distance, the program. They are not actively involved as stakeholders, but are concerned about the outcomes, including the money. These bystanders are important because they can become cheerleaders or critics of the program.

Levels and Steps

Most of the existing models have been developed to enhance, modify, or improve what was initially published fifty years ago by Don Kirkpatrick. His basic premise of considering evaluation as steps of measuring reaction, learning, behavior, and results, brought a novel and useful approach

to practitioners. Although a few of the models take different approaches, the most-used, essential framework is Kirkpatrick's steps, now labeled levels. In the 1980s, Phillips added a fifth level and modified Kirkpatrick's level descriptions. The reality is that the Kirkpatrick-Phillips-based evaluation probably accounts for 80 to 90 percent of the models in use today, globally.

It may be helpful to examine measurement and evaluation of learning as a value chain, where data are collected at different times (sometimes from different sources) to provide a process. Figure 1.1 shows this value chain—fundamental to much of the current work in evaluation.

This concept shows how value is developed and also provides data from different perspectives. Some stakeholders are interested in knowing about the inputs so that they can be managed and made more efficient; others are interested in reaction; still others are interested in learning. More recently, clients and sponsors have become more interested in actual behavior change (application) and the corresponding business impact, while a few stakeholders are concerned about the actual return on investment.

Chain of Impact

The collected data are arranged as a chain of impact, shown in Figure 1.2. The chain of impact described in this figure must to be evident if the particular learning program or performance improvement project is adding business value. All stakeholders must be closely involved in the program to understand this chain of impact. The sponsor must see this chain as the data are generated throughout the process. Participants must realize that they have a critical role and that their involvement and success are shown through the chain. The designers, developers, and facilitators have to understand that the chain of impact is critical. It can be broken, essentially at any stage, and the evaluation data will indicate whether it is broken and where it is broken. Was it broken because of adverse reaction, no learning, no application? Or was there no connection to a business measure? The information described in this book will clearly indicate whether the chain of impact is intact and where it can be strengthened. Also, it will show when it breaks and there is no value.

When the chain of impact is considered throughout the process of evaluation, some interesting characteristics begin to evolve, as shown in Figure 1.3. The evaluation data are collected throughout the chain.

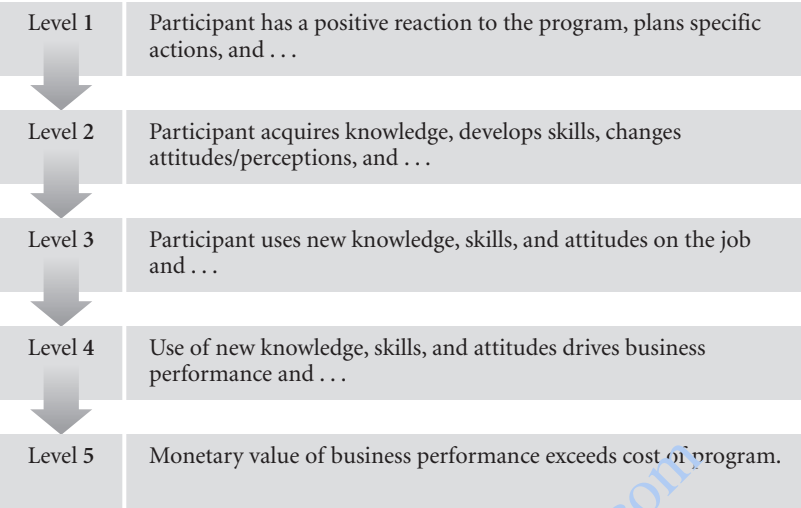| Level	Measurement Focus	Typical Measures
0. Inputs and Indicators	Inputs into the program, including indicators representing scope, volumes, costs, and efficiencies	Types of topics, content Number of programs Number of people Hours of involvement Costs
1. Reaction and Planned Action	Reaction to the program, including plans to take action	Relevance Importance Usefulness Appropriateness Intent to use Motivational
2. Learning and Confidence	Learning how to use the content and materials, including the confidence to use what was learned	Skills Knowledge Capacity Competencies Confidences Contacts
3. Application and Implementation	Use of content and materials in the work environment, including progress with implementation	Extent of use Task completion Frequency of use Actions completed Success with use Barriers to use Enablers to use
4. Impact and Consequences	The consequences of the use of the content and materials expressed as business impact measures	Productivity Revenue Quality Time Efficiency Customer satisfaction Employee engagement
5. ROI	Comparison of monetary benefits from the program to program costs	Benefit/cost ratio (BCR) ROI (%) Payback period

Figure 1.1. The Types and Levels of Data.

Level 1	Participant has a positive reaction to the program, plans specific actions, and . . .
Level 2	Participant acquires knowledge, develops skills, changes attitudes/perceptions, and . . .
Level 3	Participant uses new knowledge, skills, and attitudes on the job and . . .
Level 4	Use of new knowledge, skills, and attitudes drives business performance and . . .
Level 5	Monetary value of business performance exceeds cost of program.

Figure 1.2. Seeking the Chain of Impact.

Chain of Impact	Value of Information	Focus	Power to Show Results	Frequency of Use	Difficulty of Assessment
	Lowest	Consumer	Lowest	Frequent	Easy
Reaction					
Learning					
Application					
Impact					
ROI					
	Highest	Client	Highest	Infrequent	Difficult

Consumers: The customers who are actively involved in the process.
Clients: The customers who fund, support, and approve the project.

Figure 1.3. Characteristics of Evaluation Levels.

The data are more valuable as the process moves from reaction to ROI, at least from the client's perspective. The lower levels of data, for example, reaction and learning, are mostly consumer-oriented data, taken directly from the consumer. Reaction data are a consumer satisfaction index. Learning data are often provided to the consumer to build confidence, but impact and ROI data are more client-focused. They are the type of data that clients want to see from their learning and performance improvement programs. However, while the power to show results increases as data move through the chain, evaluating the data becomes more expensive and more difficult. However, the reverse is true for usage. As expected, a high level of data collection activity occurs at Level 1, but a low level of activity occurs around Level 4 and Level 5. And some good reasons for this exist, as will be described later in the book.
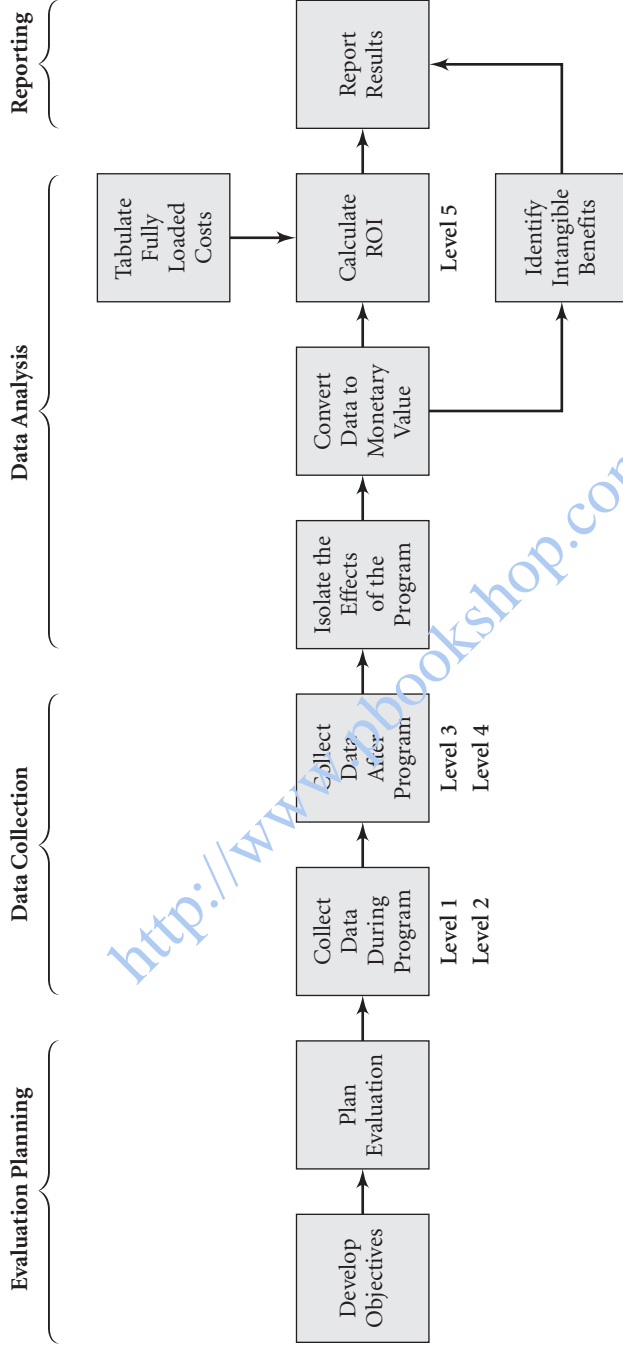
ROI Process Model

Measurement and evaluation must be systematic, methodical, and routine. Figure 1.4 shows the ROI Process Model used in this book. It illustrates the different steps in the process, beginning with objectives and proceeding through until an impact study is generated. However, the evaluation can stop at any level along the process. The first level, as described earlier, is basically inputs to the process. The data collected during programs at Level 1 and Level 2 and on a post basis for Levels 3 and 4, and Level 5 are an ROI process level.

Objectives

In the learning and development and performance improvement fields, the primary focus has been on developing learning objectives. However, for many, if not most, programs, objectives need to be enhanced to include Level 3 and Level 4 objectives. One of the most important developments in measurement and evaluation is the creation of higher levels of program objectives. Program objectives correspond with the different levels on the value chain. Ideally, the levels of objectives should be in place at the highest level desired for evaluation. Essentially, the levels of objectives are:

• Input objectives (number of programs, participants, hours, etc.)—
 Level 0

Figure 1.4. ROI Methodology Process Model.

Source: © ROI Institute. All rights reserved.

- Reaction and satisfaction objectives—Level 1
- Learning objectives—Level 2
- Application objectives—Level 3
- Impact objectives—Level 4
- ROI objectives—Level 5

Exhibit 1.2 shows an example of the multiple levels of objectives taken from a coaching program.

LEVEL 1. REACTION OBJECTIVES

After participating in this coaching program, the executive will

- Perceive coaching to be relevant to the job
- Perceive coaching to be important to job success at the present time
- Perceive coaching to be value added in terms of time and funds invested
- Rate the coach as effective
- Recommend this program to other executives

LEVEL 2. LEARNING OBJECTIVES

After completing this coaching program, the executives should improve their understanding or skills for each of the following:

- Uncovering individual strengths and weaknesses
- Translating feedback into action plans
- Involving team members in programs and goals
- Communicating effectively
- Collaborating with colleagues
- Improving personal effectiveness
- Enhancing leadership skills

LEVEL 3. APPLICATION OBJECTIVES

Six months after completing this coaching program, executives should

- Complete the action plan
- Adjust the plan accordingly, as needed for changes in the environment
- Identify barriers and enablers
- Show improvements on the following items:
  Uncovering individual strengths and weaknesses
  Translating feedback into action plans

Exhibit 1.2. Examples of Objectives.

(Continued)

Involving team members in programs and goals

Communicating effectively

Collaborating with colleagues

Improving personal effectiveness

Enhancing leadership skills

LEVEL 4. IMPACT OBJECTIVES

After completing this coaching program, executives should improve at least three specific measures in the following areas:

• Sales growth

• Productivity/operational efficiency

• Direct cost reduction

• Retention of key staff members

• Customer satisfaction

**LEVEL 5. ROI OBJECTIVE**

• The ROI value should be 25 percent.

Exhibit 1.2. Examples of Objectives, Cont'd.

Before an evaluation is conducted, these objectives must be identified and developed. Ideally, they should be developed early when the program is designed. If they are not readily available, they'll have to be included to take the evaluation to the desired level.

Evaluation Planning

The time at which evaluation is considered has changed dramatically in recent years. The traditional instructional systems design model, called ADDIE (Analysis, Design, Development, Implementation, and Evaluation), has been replaced with a new model, as illustrated in Figure 1.5. Evaluation must be considered at the conception of the program and often throughout the process. If evaluation is not considered early, serious limitations occur in the quality and quantity of data that are collected for evaluation. ADDIE places evaluation at the end, and unfortunately, professional practitioners waited until after implementation to think about evaluation. That's too late.

Evaluation must be planned—overall and individually—for each program. When evaluation is conducted only at reaction levels, not

Analysis

↓

Design

↓

Development

↓

Implementation
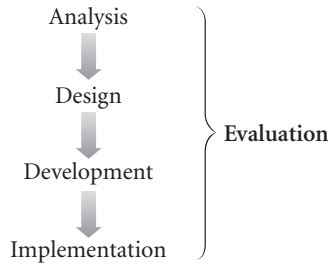
Evaluation

Figure 1.5. The New Design Model.

much planning is involved, but as evaluation moves up the value chain, increased attention and efforts need to be placed on planning. During the typical planning cycle, the purpose of evaluation must be reviewed for specific solutions and to determine where the evaluation will stop on the value chain. The feasibility of evaluating at different levels is explored, and two planning documents are developed when the evaluation migrates to application, impact, and ROI: the data collection plan and the data analysis plan. These documents are sometimes used in combination, but are often developed separately.

Data Collection

One important issue is the timing of data collection. In some cases, pre-program measurements are taken to compare with post-program measures, and in some cases, multiple measures are taken. In other situations, pre-program measures are not available and specific follow-ups are still taken after the program. The important issue is to determine the timing for the follow-up evaluation.

Another important issue is the data collection method used. Data are collected using the following methods:

- *Surveys* are administered to determine the extent to which participants are satisfied with the program, have learned the skills and knowledge, and have used different aspects of the program.

- *Questionnaires* are usually more detailed than surveys and can be used to uncover a wide variety of data. Participants provide responses to several types of open-ended and forced-response questions.

- *Tests* are conducted to measure changes in knowledge and skills. Tests come in a wide variety of formal (criterion-referenced tests, performance tests, simulations, and skill practices) and informal (facilitator assessment, self assessment, and team assessment) methods.

- *On-the-job observation* captures actual skill application and use. Observations are particularly useful in customer service training and are more effective when the observer is either invisible or transparent.

- *Interviews* are conducted with participants to determine the extent to which learning has been used on the job.

- *Focus groups* are conducted to determine the degree to which a group of participants has applied the training to job situations.

- *Action plans and program assignments* are developed by participants during the program and are implemented on the job after the program is completed. Follow-ups provide evidence of program success.

- *Performance contracts* are developed by the participant, the participant's supervisor, and the facilitator, who all agree on job performance outcomes.

- *Business performance monitoring* is useful when various performance records and operational data are examined for improvement.

Analysis

Evaluation requires analysis. Even if the evaluation stops at Level 1, analysis is required, usually involving simple averages and standard deviations. As organizations progress up the value chain, additional analyses are required. In some cases, not only are the averages and standard deviations used, but simple hypotheses testing and correlations may be required; however, these are very unusual situations. For the most part, analysis is simply tabulating, organizing, and integrating data and then presenting results in meaningful ways for the audience to understand and appreciate.

Isolation of the Effects of Learning and Development

An often-overlooked issue in some evaluations is the process of isolating the effects of learning on output data. This step is important

because many factors will usually influence performance data after a learning program is conducted. Several techniques are available to determine the amount of output performance directly related to the program. These techniques will pinpoint the amount of improvement directly linked to the program, resulting in increased accuracy and credibility of the evaluation data. The following techniques have been used by organizations to tackle this important issue:

- *A control group* arrangement is used to isolate learning's impact. With this strategy, one group participates in a program, while another, similar group does not. The difference in the performance of the two groups is attributed to the program. When properly set up and implemented, the control group arrangement is the most effective way to isolate the effects of learning and development.

- *Trend lines and forecasting* are used to project the values of specific output variables as if the learning program had not been undertaken. The projection is compared to the actual data after the program is conducted, and the difference represents the estimate of the impact of learning.

- *Participants or managers* estimate the amount of improvement related to the learning and development program. With this approach, participants or managers are provided with the total amount of improvement that is actually related to the program.

- *Other experts,* such as customers, provide estimates of the impact of learning on the performance variable. Because the estimates are based on previous experience, these experts must be familiar with the type of program and the specific situation.

Conversion of Data to Monetary Values

To calculate the return on investment, business impact data collected in the evaluation are converted to monetary values and compared to program costs. This requires that a value be placed on each unit of data connected with the program. Several techniques are available to convert data to monetary values. In many cases, standard values are available as organizations have attempted to place value on measures they want to increase and develop costs for measures they want to avoid. When these are not available, the records (or a combination of records) may show the cost or value of the measure. Also, internal

experts, external experts, or external databases can be sources of values. Sometimes, participants, supervisors, and other conveniently available staff members can provide the values.

This step is necessary for determining the monetary benefits from a learning program. The process is challenging, particularly with soft data, but can be methodically accomplished using one or more of the techniques described above.

The Cost of Programs

The cost of learning is usually developed from one of two perspectives:

1. For budgets, program approvals, and general information requests, costs are often reported systematically within the organization, and usually include only the direct costs. Executives and administrators are often interested in the direct costs. In some cases, these reports are changing to include other indirect costs.

2. When the actual ROI is calculated, the costs must be fully loaded to include all direct and indirect costs. In these situations, the cost components should include:

   - Needs assessment, design, and development, possibly prorated over the expected life of the program
   - All program materials provided to each participant
   - Instructor/facilitator, including preparation time as well as delivery time
   - Facilities for the learning program
   - Travel, lodging, and meal costs for the participants, if applicable
   - Salaries, plus employee benefits of the participants of the learning program
   - Administrative and overhead costs of the workplace learning and performance function, allocated in some convenient way
   - Evaluation, including planning, data collection, analysis, and reporting

The conservative approach is to include all these costs so that the total is fully loaded.

The Return on Investment Calculation

Dramatic changes have occurred in the need for data around learning and development programs. Many executives and managers have taken the approach, "Show me the money." Figure 1.6 shows how the "Show Me" request has evolved, leading up to an actual request for ROI. For some professionals, this is an issue that cannot be ignored because of the serious consequences. Executives and managers want this type of data, and it must be delivered. This requires at least a few major programs to be elevated to the ROI analysis level.
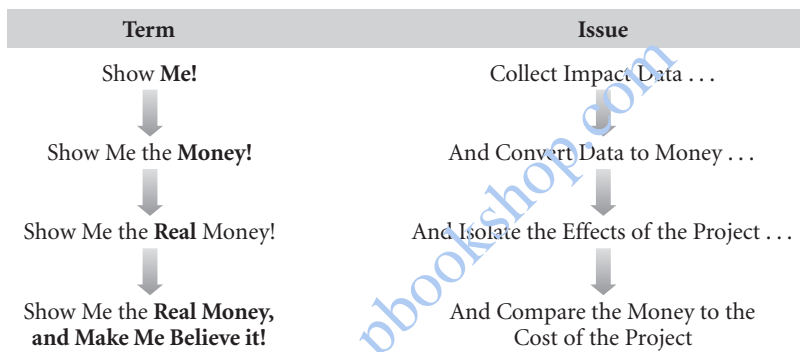
Term	Issue
Show **Me!**	Collect Impact Data . . .
Show Me the **Money!**	And Convert Data to Money . . .
Show Me the **Real** Money!	And Isolate the Effects of the Project . . .
Show Me the **Real Money, and Make Me Believe it!**	And Compare the Money to the Cost of the Project

Figure 1.6. The "Show Me" Evolution.

When the ROI is actually developed, it should be calculated systematically, using standard formulas. Two formulas are available. The benefit/cost ratio is the program benefits divided by the costs. In formula form, it is:

$$BCR = \frac{\text{PROGRAM BENEFITS}}{\text{PROGRAM COSTS}}$$

The return on investment calculation uses the net benefits divided by program costs. The net benefits are the program benefits minus the costs. In formula form, the ROI is:

$$ROI\ (\%) = \frac{\text{NET PROGRAM BENEFITS}}{\text{PROGRAM COSTS}} \times 100$$

This is the same basic formula used in evaluating other investments for which the ROI is traditionally reported as earnings divided by investment. An example of the benefit/cost ratio and ROI is illustrated below. A training program is delivered to fifty participants. Consider that, following the training program, the first-year program benefits (from Level 4 business impact data) are found to be $300,000 from the fifty participants, and the fully loaded costs to train these fifty participants is $200,000.

$$BCR = \$300,000/\$200,000 = 1.50{:}1$$

$$ROI = \$100,000/\$200,000 \times 100 = 50\%$$

The ROI calculation of net benefits ($300,000 minus $200,000) divided by total costs brings an ROI of 50 percent. This is what is earned after we get back the $200,000 spent for the program. The ROI calculation accounts for the program costs and shows the resulting net gain.

The BCR (benefit/cost ratio) calculation above uses the total benefits in the numerator. Therefore, the expressed BCR of 1.50:1 does not account for replacing the money expended. This is why, when using the same values, the BCR will always be 1 greater than the ROI. The BCR of 1.50:1 in the example means that, for every dollar spent, $1.50 is gained back. One dollar has to pay for the investment, so the net is $0.50 (as expressed in the ROI calculation).

Intangible Benefits

In addition to tangible benefits, most learning programs will influence intangible, non-monetary, benefits. Intangible benefits may include:

- Increased job satisfaction
- Increased organizational commitment
- Improved teamwork
- Improved customer service
- Reduced complaints
- Reduced conflicts

During analysis, hard data—such as output, quality, and time—are usually converted to monetary values. The conversion of soft data is at-

tempted. However, if the process used for conversion is too subjective or inaccurate and the resulting values lose credibility during the process, then the data are listed as intangible benefits with an appropriate explanation given. For some programs, intangible benefits are extremely valuable, often carrying as much influence as hard data items.

Data Reporting

This critical step is often not given the proper attention and planning needed to ensure that it is successful. This step involves developing appropriate information as impact studies, executive summaries, one-page summaries, and other brief reports. The heart of the step includes the different techniques used to communicate to a wide variety of target audiences. In most situations, several audiences are interested in and need to know the information. Careful planning to match the communication method with the audience is essential to ensuring that the message is understood and appropriate actions are taken.

Operating Standards

To ensure consistency and replication of evaluation studies, operating standards should be developed and applied in the measurement and evaluation process. The results of an evaluation must stand alone and should not vary based on the individual who is conducting the study. The operating standards detail how each step and issue of the process should be addressed. The standards presented in this book are called the guiding principles. They are listed below:

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 are available for a population or from a specific source, it is assumed that little or 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. The results from the ROI methodology must be communicated to all key stakeholders.

These specific standards not only serve a way to consistently address each step, but also provide a much-needed conservative approach to the analysis. A conservative approach will build credibility with the target audience.

Implementation Issues

A variety of organizational issues and events will influence the successful implementation of measurement and evaluation. These issues must be addressed early to ensure that evaluation is successful. Specific topics or actions may include:

- A policy statement concerning results-based learning and development
- Procedures and guidelines for different elements and techniques of the evaluation process
- Meetings and formal sessions to develop staff skills with measurement and evaluations
- Strategies to improve management commitment and support for measurement and evaluation

- Mechanisms to provide technical support for questionnaire design, data analysis, and evaluation strategy

- Specific techniques to place more attention on results

  Measurement and evaluation can fail or succeed based on these implementation issues.

FINAL THOUGHTS

More attention needs to be focused on measurement and evaluation; this is almost universally agreed on. The use of measurement and evaluation is expanding. The payoff is huge. The process is not very difficult. The approaches, strategies, and techniques are not overly complex and can be useful in a variety of settings. The combined and persistent efforts of practitioners and researchers will continue to refine the techniques and create successful applications. In the next chapter, the first step of the evaluation process, the needs analysis, will be explored.