IN THIS ARTICLE

Evaluation Design, Return on Investment

ROI: The Search

BY JACK J. PHILLIPS

FOR MANY YEARS, MEASURING THE RETURN ON INVESTMENT for training and development has been a critical issue—on meeting agendas, in the literature, and on the minds of top executives. Although interest has heightened and some progress has been made, the topic still challenges even the most sophisticated and progressive HR departments. Some HR professionals argue that measuring ROI for training isn't possible; others quietly and deliberately develop ROI measures. But overall, most practitioners acknowledge that they must show a return on the investment in training so that they can maintain training funds and enhance HR's status.

Currently, it's difficult to pinpoint the state of ROI within the field. Many HR managers are reluctant to disclose internal practices. And many say that little progress has been made, even in the most progressive organizations. It's also difficult to find case studies that show specifically what organizations have done in measuring training ROI.

Recognizing the void, the American Society for Training and Development began collecting case studies with real-life examples of measuring the return on investment in training. ASTD contacted more than 2,000 people in HR directly, including senior practitioners, authors, researchers, consultants, and conference presenters. ASTD also contacted organizations that were perceived as profitable, respected, and admired. In addition, announcements appeared in the *National Report* and other ASTD publications inviting case submissions. To meet requirements, organizations had to be willing to describe the specific steps, issues, and concerns involved in their efforts to measure training ROI.

More than 150 respondents requested specific guidelines for developing a case; 40 were willing to submit cases. ASTD stopped collecting cases once 30 were delivered. Of those, 18 were selected for publication in *Measuring Return on Investment* (ASTD, 1994). A second volume is planned for publication in late 1996.

Building onto level 4

The ROI process adds a fifth level to the Level-4 evaluation model developed by Donald Kirkpatrick. See figure 1. At level 1, participants' satisfaction with the training program is measured, and a list of their plans for implementing the training is included.

At level 2, measurements focus on what participants learned during training. At level 3, the measures assess how participants applied learning on the job. At level 4, the measures focus on the business results achieved by participants when the training

Illustration by Sandra Hendler

for Best Practices

objectives are met. The fifth, and ultimate, level of evaluation is the return on investment. It compares the training's monetary benefits with the costs.

Most organizations conduct evaluations to measure satisfaction; few conduct evaluations at the ROI level. Both are desirable. Evidence shows that if measurements aren't taken at each level, it's difficult to show that any improvement can be attributed to the training.

The model in figure 2 is a framework for developing ROI. Many of the companies in the case studies followed this model. It tracks the steps in measuring ROI—from collecting post-program data to calculating the actual return. The model assumes that training costs will be compared with monetary benefits and that all training programs will also have intangible, but reportable, benefits.

The process begins with the collection methods of post-program data. Such methods are at the heart of any evaluation. Which methods to use depends on the evaluation's purposes, instruments, measurement levels, design, and cost of data collection.

This is the first in a series of three articles about measuring the return on investment in training. Real-world case studies provide a look at how the search is going.

Two common formulas for calculating return on investment are a benefit/cost ratio (BCR) and ROI. To find the BCR, you divide the total benefits by the cost. In the ROI formula, the costs are subtracted from the total benefits to produce the net benefits, which are then divided by the costs.

For example, a literacy-skills training program at Magnavox produced benefits of \$321,600 with a cost of \$38,233. The BCR is 8.4. For every \$1 invested, \$8.4 in benefits were returned. The net benefits are \$321,600 – \$38,233 = \$283,367. ROI is \$283,367 ÷ \$38,233 x 100 = 741 percent. Using the ROI formula, for every \$1 invested in the program, there was a return of \$7.4 in net benefits.

Typically, the benefits are annual, the amount saved or gained in the year after training is completed. The benefits may continue after the first year, but the effect begins to diminish. In a conservative approach, long-term benefits are omitted from calculations. In the total cost of a program, the development cost is usually front-loaded and prorated over the first year

Figure 1: Level-5 Evaluation

Here are some questions for conducting a level-5 evaluation.

Level	Questions
1 reaction and planned action	▸ What are participants' reaction to the program? ▸ What do they plan to do with what they learned?
2 learning	▸ What skills, knowledge, or attitudes have changed? By how much?
3 applied learning on the job	▸ Did participants apply what they learned on the job?
4 business results	▸ Did the on-the-job application produce measurable results?
5 return on investment	▸ Did the monetary value of the results exceed the cost of the program?

of implementation. Or, you can prorate development costs over the projected life of a program.

Some recommendations

The case studies in the ASTD project represent a wide range of settings, strategies, and approaches in manufacturing, service, and government organizations. The training audiences varied from all employees to managers only to specialists only. Though most of the programs focused on training and development, others included such areas as total quality, performance management, and employee selection.

The cases provide a rich source of information on the strategies and thought processes of some of the best practitioners, consultants, and researchers in the field. The companies' returns on investment ranged from 150 to 2,000 percent.

Several common approaches have emerged. They could be considered best practices or just recommendations. Whichever, they seem to have worked well for the companies in the case studies.

Set targets for each evaluation level. Some organizations set a target for each level of evaluation, a target being the percentage of HR programs that will be measured at that level.

For example, many organizations require 100 percent of their training programs to be evaluated at level 1 because it's fairly easy to measure participants' reactions. Level 2 (learning) is also relatively easy to measure. Typically, the target range is 40 to 70 percent, depending on the type of program.

Level-3 evaluation (on-the-job application) involves more time and expense to conduct follow-up evaluations so targets tend to be lower at 30 to 50 percent. Level-4 (business results) and level-5 (ROI) evaluations require significant resources and budgets so their targets tend to be small: 10 percent for level 4 and five percent for level 5.

Establishing evaluation targets has several advantages. One, it provides measurable goals for assessing the progress of all training or a particular segment. It also focuses attention on accountability and communicates a strong message to the HR staff about the need for measurement and evaluation.

Use a variety of approaches to collect evaluation data

Evaluate at the micro level. Measurement and evaluation usually focus on a single program or a few tightly integrated programs. ROI measurement is more effective when applied to one program that can be linked to a direct payoff. When all of the courses in a series must be completed before their common objectives are met, it may be appropriate to wait to evaluate the series as a whole. The decision to evaluate several courses in a series should take into account the training goals, timing of the courses, and cohesiveness of the series.

It can be difficult to evaluate a series conducted over a long period of time. A cause-and-effect relationship becomes more confusing and complex. Also, it is hard to evaluate an entire function, such as management development, career development, executive education, or technical training.

Use a variety of methods. The companies in the case studies use a variety of approaches to collect evaluation data. They don't latch onto one or two practices and stay with them regardless. They recognize that every program, setting, and situation is different. They know that techniques such as interviews, focus groups, and questionnaires work well in some situations and that action plans, contracts, and performance monitoring are needed in others.

These companies use internally developed criteria to match a particular data-collection method with the training program.

Isolate the effects of training. A critical aspect of the evaluation process is trying to isolate the effect of the training from other factors occurring during the same period that could affect business results. Most of the time, training can take only partial credit for improvements in on-the-job performance. When planning to measure ROI, the case-study organizations go beyond a standard control-group analysis to use one or more techniques for isolating extraneous factors.

Use sampling wisely. It's rare for organizations to use statistical sampling in selecting a sample of training programs in which to measure ROI. For most, the result would be too many calculations. For the sake of practicality, many organizations decide to evaluate just one or two sessions of their most popular training programs. Others select one program from each major training segment. It's recommended that organizations calculating ROI for the first time select only one course to measure, as a learning experience.

If sampling is used, it's important to be statistically sound. But it's more important to consider the tradeoff between the available resources and what kind of ROI calculations management is willing to accept. Remember: The primary goals of an ROI calculation are to convince the HR staff that the process works and show senior-level managers that training can make a difference. With that in mind, it's best to get the input and approval of top management in developing your sampling approach.

The sample number depends on the following variables:

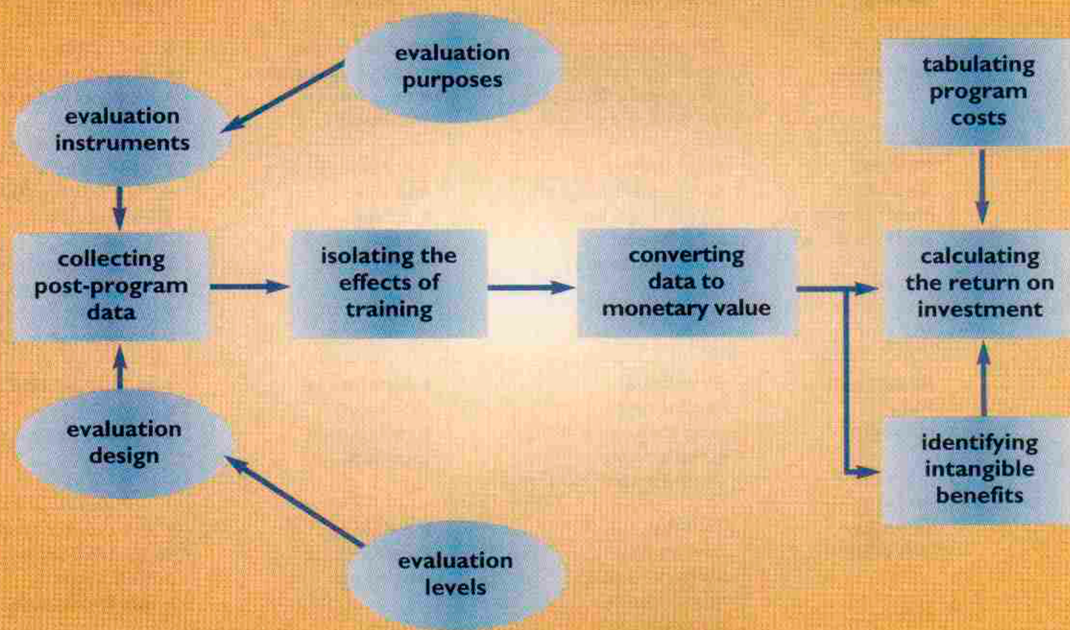
- the HR staff's expertise on evaluation
- the type of training programs being evaluated
- the resources allocated for evaluation

A Look at the Case Studies

Setting	Target Group	Program Description	Evaluation Process	Results
Bottling company (Coca Cola)	First-level supervisors	Eight half-day workshops covering supervisory roles, setting goals, developing the team, and so forth	▸ Action planning ▸ Follow-up session ▸ Performance monitoring	▸ 1,447% ROI ▸ Benefit/cost ratio 15:1 ▸ Variety of measures
Paper products company	Managers, supervisors, hourly employees study teams, skill	Organization development program (Workshops, action ▸ Performance building programs)	▸ Follow-up with interviews ▸ Survey ▸ Performance monitoring	▸ Variance from standard +$106,000 ▸ Efficiency 4% improvement ▸ Waste 36% improvement ▸ Absenteeism 35% improvement ▸ Safety 25% improvement ▸ Housekeeping 27% improvement
Health Maintenance Organization	All managers and all employees	Organization development program (team building, group meetings, customer service training	▸ Performance monitoring ▸ Management estimation	▸ 20,700 New HMO members ▸ 1,270% RO ▸ Benefit/cost ratio 13.7:1
Large commercial bank	Consumer loan officers	Two-day sales training program—focus on increase in consumer loans	▸ Follow-up ▸ Performance monitoring	▸ 30% increase in consumer loans ▸ 2,000% ROI ▸ Benefit/cost ratio 21:1
Information services company	Supervisors	Twelve two-and-one-half hour sessions on behavioral modeling	▸ Follow-up with surveys	▸ 336% ROI
Electric & gas utility	Managers & supervisors	Applied behavior management which focused on achieving employee involvement to increase quality, productivity, and profits	▸ Action planning (variety of projects) ▸ Performance monitoring	▸ 400% ROI ▸ Benefit/cost ratio 5:1
Oil company	Dispatchers	Skills training program including customer customer interaction skills, problem solving, and teamwork	▸ Follow-up observations ▸ Performance monitoring	▸ Customer complaints reduced by 85% ▸ Absenteeism reduced by 77% ▸ Reduction in pull-outs saved $283,800 ▸ 383% ROI ▸ Benefit/cost ratio 4.8:1
Bakery (Multi-Marques, Inc.)	Supervisors/ administration services	Fifteen-hour supervisory skills training including the role of training	▸ Action planning (work process analyses) ▸ Performance monitoring	▸ 215% ROI ▸ Benefit/cost ratio 3.3:1
Avionics (Litton Industries)	All employees	Self-directed work teams	▸ Action planning ▸ Performance monitoring	▸ Productivity increased 30% ▸ Scrap rate reduction 50% ▸ 700% ROI
Truck leasing (Penske Truck Leasing)	All supervisors	20-hour program on supervisory skills using behavioral modeling	▸ Performance monitoring	▸ Turnover reduction of 6% ▸ Abseeteeism reduction of 16.7%

These cases appear in Measuring Return on Investment, *published by the American Society for Training and Development, Alexandria, VA. 1994. Jack J. Phillips, editor.*

Figure 2: A Model for Calculating ROI



- the degree of support from management for training and development
- the organization's commitment to measurement and evaluation
- the amount of pressure from others to show ROI calculations.

Other variables endemic to the particular organization may apply.

Convert program results to monetary values. The organizations in the case studies seek a specific return on investment. Consequently, data on business results must be converted to monetary benefits. These companies aren't content to show just that training can result in such improvements as increased productivity and decreased employee turnover. They take the process a step further by converting such improvements to monetary units so that the improvements can be compared to costs and further developed into an ROI calculation.

For such hard-data items as productivity, quality, and time, the conversion to monetary units is relatively easy; soft-data items such as customer satisfaction, employee turnover, and job satisfaction aren't so easy to convert. Still, there are techniques for making the conversions with a reasonable amount of accuracy, and several strategies are used in these case studies. (Note: this important topic is covered in the third article of this series.)

The search continues

Because the variables that can affect performance are numerous and complex, it can be difficult to determine how much of a change is due to training. Most ROI figures aren't precise, though they tend to be as accurate as many other estimates that organizations routinely make.

After garnering the cases for the ASTD project, there are still unanswered questions about measuring ROI.

Cost standards. The methods used to monitor costs vary widely. What one organization considers a cost of training, another does not. The HR field needs standard cost data. It's becoming increasingly difficult to compare costs from one program to another. Most efforts to solve the problem have failed. In the interim, it's necessary to describe the cost components that make up the total cost of any effort to measure ROI.

Evaluation design. Many organizations don't design their evaluations to isolate the effects of training. Control groups are rarely used even though they can be used effectively without the disruption, problems, and inconvenience usually feared by practitioners. Though a control-group approach is preferable, other evaluation designs such as trend-line analysis, forecasting, and estimations can be useful.

Standard methodology. Evaluation techniques vary, though there are only so many ways that data can be collected and analyzed. Often, data-collection methods are used without regard to their advantages or disadvantages. The different labels and terminology adds to the confusion. As professionals, we need to standardize and publicize evaluation methods.

Statistics. Many HR practitioners avoid statistics. But statistical analysis can provide a sound basis for conclusions about training results. And though many top managers don't understand statistical analysis, they need to feel confident that any conclusions about training results are supported by appropriate methodology.

In several of the case studies and other evaluation projects, the power of statistics is largely ignored. Sample sizes are so small that the results can't be considered supportable, at least statistically.

Converting data to monetary values. Because of the subjective nature of this process, the results of many HR programs aren't converted to monetary units. Yet, this conversion is an essential step in ROI calculations in which monetary benefits must be compared with costs. It should be a fundamental requirement for some level-4 evaluations.

Evaluate, evaluate, evaluate

In the cases cited in the ASTD project, a variety of evaluation methods were used to determine the success of training.

Follow-up assignments. Perhaps the easiest approach to post-training data collection is to ask participants to complete a task or project, to serve as evidence of their successful application of the training content. Typically, assignment results are reported to participants' supervisors. This approach is especially helpful for level-3 evaluations and when management support isn't strong.

Surveys and questionnaires. These can capture participants' accomplishments and behavioral changes after training. You can collect responses from all participants or a sampling. Surveys and questionnaires are inexpensive, as well as easy to implement and tabulate. They're most appropriate for on-the-job application (level 3) and business results (level 4).

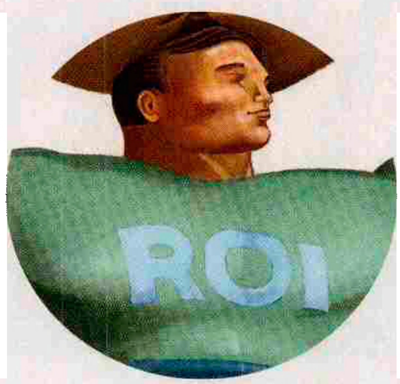
One-on-one interviews. Interviewing participants individually is an excellent way to capture changes in job-related behaviors and to garner specific details. More versatile than questionnaires, interviews can probe issues, concerns, and actions related to the training. They're suited to level-3 evaluations.

Focus groups. An extension of interviewing, focus groups involve collecting post-training information from eight to 12 participants in a structured setting. Focus-group members are asked specific questions about what they have changed as a result of training. The exchange of information often triggers creative thinking among participants, which provides high-quality data. A focus-group follow-up is appropriate for level-3 evaluations.

Observation. Direct on-the-job observation of participants after training can show whether they are applying new skills. This level-3 approach works best when participants are unaware that they're being observed. For example, to measure changes in customer service, an organization can hire shoppers to observe salespeople.

Action planning. The most powerful way to measure training's effect is through the use of action plans. Participants apply their new skills or knowledge in a task or project and then document their progress in achieving measurable objectives outlined in an action plan. Their supervisors may or may not be involved. In some organizations, progress is audited by a training coordinator. Or, participants submit their action plans to the training department to substantiate whether the desired results are attained.

The training should provide a module on how to develop an action plan. Both level-3 and level-4 evaluations can benefit from this approach.



The most powerful measurement is through the use of action plans

Performance contracts. These contain a pretraining agreement between participants, their supervisors, and sometimes the facilitator. The parties meet prior to training in order to develop measurable goals related to the training content. Later, they can determine whether the goals were met.

Special follow-ups. It can be effective to reconvene participants one to three months after the initial training segment was conducted so they can report on their success. Follow-up sessions also provide opportunities for additional training, such as refining new skills. This approach is appropriate for level-3 and level-4 evaluations.

Performance tracking. The most credible post-evaluation approach is to track department, work-unit, or individual performance after training is completed in such areas as productivity, quality, cost, and time—and in soft-data areas such as customer satisfaction.

This approach requires examining the organization's overall performance data to obtain before-and-after comparisons of each data item. Because it can provide the most convincing evidence, it's often the preferred approach of senior managers.

The search for the best practices has revealed some important concerns. It's almost universally agreed that more attention regarding ROI is needed. But only a few successful examples of ROI calculation exist. The process isn't as difficult as it may seem. The approaches and techniques can be useful in a variety of settings. Practitioners and researchers must continue to refine the techniques and show successful applications.

In the next issue of *T&D*, the second article in this series will focus on approaches for isolating the effects of training. ■

Jack J. Phillips *is principal consultant with Performance Resources Organization. He can be reached at Box 1969, Murfreesboro, TN 37111-1969. Phone 615/896-7694, fax 615/896-7181.*