

Liberating the Lamppost: Shedding Light on Training Evaluation

by Peter Bregman and Howie Jacobson

An old joke: A training manager is searching under a street lamp for a set of keys. A passer-by offers to help. Together they search and search, in vain. Finally the passer-by asks where the keys fell. The training manager points to the other side of the street. “Why are you looking on this side when you dropped them over there?” the incredulous passer-by asks. The reply: “The light’s much better here.”

Training and development professionals have a problem. In theory, we believe in evaluating our work. In practice, however, we don’t know how to do it. We measure what is easiest to measure, instead of what is most important. We search for the keys to effective training under the lamppost, instead of where they are. This article will examine the reasons for this misdirected effort, and suggest a simple and inexpensive solution.

In Theory, It’s a Good Theory

According to prevailing theory, we are supposed to evaluate training from four related points of view: what the trainees think, what they learn, how their behavior changes following training, and how those changes affect the business (the “Kirkpatrick Model”). That seems straightforward enough: trainees value training in useful new skills and behaviors. Because they are useful, those skills and behaviors will be used in the workplace. Therefore, the organization will benefit in tangible ways.

The purpose of the four-level evaluation is to test the training program against this theory. The rosy picture described above can break down anywhere. For example, trainees may react favorably at the end of an exciting program. A few

days later, they realize that they haven’t actually learned anything. When the jazzed feeling wears off, so does any perception of value. Or, most typically, trainees learn how to behave and why, but slip back into old behaviors on the job. “Raving fans,” you hear the help desk technician growl through clenched teeth after three hours back on the job, “I’ll give them something to rave about.” Occasionally, the behaviors change, but the impact on the business is negligible. Customer service reps now practice active listening, but your networking software still crashes every time a user joins a chat room. As a training manager, you want to evaluate all four potential breakdown points. Only then will you understand how the training program relates to the desired results.

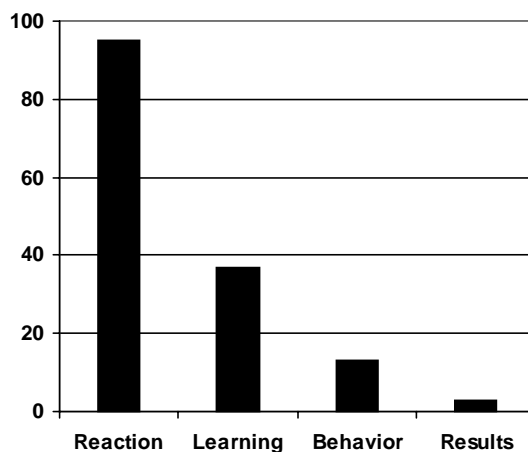
The three lower levels of evaluation (reactions, learning, and behaviors) are only of interest insofar as they serve business results. They don’t matter by themselves. Here, in a nutshell, is the problem:

- Employee reactions to training are easiest and least expensive to measure. They are what you care about least.
- Business results due to training are most difficult and costliest to measure. They are what you care about most.

Training managers typically survey trainees immediately following the last training session and consider the program fully evaluated. It is easy to see why. We prefer looking for the keys directly under the lamppost, where the light is brightest. But we know the keys are not there. Current theory does not explain how to shine a light on the side of the street where there is something important to find. At best, we try to tilt the beam to shine dimly where we need it. How can we liberate the lamppost to move where we need it?

Leaning on the Lamppost

The 2000 American Society for Training and Development (ASTD) State of the Industry report found that 95% of surveyed organizations gauged trainees' reactions to training, 37% measured their learning, 13% tracked behavior changes following training, and only 3% accounted for business results of the training.



Percent of Surveyed Organizations Measuring Each Evaluation Level (ASTD, 2000).

Since these organizations accept the need for a four-level evaluation process, what is stopping them from evaluating learning, behavior, and business results?

In a 1997 ASTD survey of human resources development professionals, respondents gave four answers most frequently:

- difficulty in determining the impact on financial performance
- time required for a proper evaluation
- inability to determine appropriate outcome measures
- cost of evaluation

The more important the outcome, the harder to measure. Let's examine why the four-level model is so difficult to use.

Reactions

Employee reaction to training is the easiest: Ten minutes before the end of the last session, the instructor hands out an evaluation questionnaire to each participant. Questions look like this: "How do you rate the subject? How do you rate the presenter?" After each question, participants are given a choice of "poor" to "excellent," and an invitation to give a comment or suggestion. That evening, the training manager takes all 49 questionnaires home (five new managers were sick on the last day, and six had to leave early) and tallies the responses over "Ally McBeal." The data, along with recommendations, are presented to headquarters the next day. The 11 absentees never give their opinions.

Reaction surveys require very little time to prepare, fill out, and tabulate. As measures of customer satisfaction, they provide immediate and often useful feedback to trainers. They are important

to the training function, but not so much to the rest of the organization.

The major limitation of trainee surveys is that they cannot tell you if the training actually worked. If the responses are highly negative, the program probably didn't do the job (though it might have). But favorable responses don't mean that trainees will retain or act on the material. I always enjoy outdoor adventure-based workshops, even when I know that the lessons won't transfer to the office. Cynical trainers have dubbed these surveys "Smile Sheets." They are necessary, but not sufficient. You might still be wasting money on training that is more fun than a day at the office, but ineffective nevertheless.

Learning

Measuring learning is harder and more time-consuming than asking trainees if they liked the training. Employees often view tests as demeaning, as a lack of trust in their abilities. Often tests do not measure learning, but how well people take tests. Good tests are hard to construct, and can be time-consuming to take if they measure all aspects of the course content. Online assessment is beginning to simplify and streamline the process, but so far has done little to solve the above problems. Just because trainees learn new skills and behaviors does not mean that they will use them in the workplace.

Behavior

Observing employee behaviors is more complicated than testing knowledge, and can also be more time-consuming. A test is an event, but behavioral observation is a process. Who should observe? How often? In what form should they record data? How is it fed

back to employees? These questions are not simply matters of convenience and efficiency; companies run the risk of alienating employees who feel that Big Brother is watching them.

Training programs exist to change behavior. If behavior at work changes in desired ways, the training has been successful. It is at this level that most training programs fail: According to most studies of non-technical training (cited in *The Learning Alliance*, by Brinkerhoff and Gill), about 90% of trainees will not be using the new skills six months later.

Business Results

Some training programs *do* change behaviors, of course. The question then becomes, How has our organization benefited from those changes? A group of senior managers learn a more efficient way to manage projects. They begin to manage projects in this new way. What results will interest the executive who paid for the training?

One obvious result is "efficiency of project management." But what does that look like? How can "efficiency" be observed? People meet deadlines. They complete tasks with less supervision. They spend less time on those tasks. But what if these "improvements" reduce customer satisfaction? Donald Kirkpatrick, in the second edition of *Evaluating Training Programs*, gives the example of a company that improved efficiency by waiting until trucks were full before sending them on runs, thus reducing transportation costs. In the short term, profits rose. But soon customers were complaining about less frequent deliveries.

Organizations are complex systems. Often something looks like an isolated improvement and actually proves disastrous for the system. Everyone knows that it is easier to do something oneself than teach someone else to do it. Suppose that increased efficiency was achieved through reduced delegation. In his excellent book, *Managing the Professional Service Firm*, David Maister points out that under-delegation of tasks can lead to morale and retention problems among junior associates. Focusing on discrete results can fool you into thinking that your training program is helping the whole business. The 1997 ASTD survey shows that many training managers do not know which results to measure.

Why not look just at financial performance? When businesspeople invest in something, they naturally want to know the return on their investment (ROI). Training managers feel pressure to demonstrate that training improves financial performance. But ROI is notoriously difficult to measure. Often, there are too many variables. When sales increase following training in communication skills, training may or may not have caused the increase. Perhaps a competitor went out of business. Maybe the increase is seasonal. Or the economy just got stronger. The 1997 ASTD survey found that human resource development professionals struggle to link training results to financial performance.

One solution is to craft a study using a similar control group: Group A receives training, while Group B does not. If the same conditions affect both groups, any differences in financial performance following the training must be caused by it. Most organizations, however, do not

have the luxury of two identical groups operating in exactly the same environments. Studies that aggressively control the differences between two groups tend not to look like the real world. For example, if someone from Group A is transferred to Group B, the study loses validity. Most controlled studies take a lot of time and a lot of money to do well. As the 1997 ASTD survey shows, human resource development professionals view excessive duration and cost as challenges to adequate evaluation.

Moving the Lamppost

There is a cheap, quick, and easy way to measure meaningful results. It involves one mind-bending assumption: the most important business result is customer satisfaction.

Every employee in an organization has customers. Many employees serve internal customers: their managers, their direct reports, their peers on project teams, members of line units that rely on their work, and others. Some employees serve external customers as well: buyers, suppliers, business partners, clients, and others. Doing a good job means providing quality service to all of one's customers.

Every organization has internal and external customers. Employees and clients are both customers, evaluating how they are treated and whether their needs are being met or exceeded by the organization. As Maister points out, professional service firms must create super-satisfied employees and super-satisfied customers if they are to succeed financially. The ultimate bottom line (profits) depends on customer and employee satisfaction. Employee satisfaction is crucial, ultimately, for

external customer satisfaction. Phil Terry, CEO of Creative Good, a New York-based internet strategy firm, puts it succinctly, “Happy employees create happy customers.”

Every business result that matters leads to improved customer satisfaction, either directly or indirectly. An event or process or attitude that does not make customers happy is bad for business, even if it improves efficiency or pads short-term profits. Recent books like *Net Ready*, by Amir Hartman, John Kador, and John Sifonis and *The Cluetrain Manifesto*, by Rick Levine, Doc Searls, David Weinberger, and Christopher Locke, identify a key feature of the New Economy: a fundamental shift in power from companies to their customers. The days of Henry Ford’s “Give ‘em any color they want, as long as it’s black” are over. To succeed, companies will need to partner with their customers to provide customized products and services.

Why Customer Satisfaction is Easy to Measure: Four Reasons

Changes in customer satisfaction due to training are easy to measure for four reasons. First, we can use reaction surveys. As we have seen, they are easy and inexpensive to create, administer, tabulate, and interpret. They are also a reliable measure of customer satisfaction. We just administer them to a different population: the customers of the training participant.

Second, reaction sheets for customers get high response rates. Customers want to give feedback, especially if they believe it will be taken seriously. Most are delighted to offer an opinion that may help them get better service in the

future, or that allows them to express frustration or appreciation. People care deeply about how they are affected by other people.

Third, it is easy to make a valid connection between the effects of training on an employee and how customers feel about that employee. By definition, a training program succeeds when it changes employees’ behaviors *in ways that matter to their customers*.

- Customer service and sales reps receive training in active listening. Ask their customers and prospects whether their concerns are being addressed better than before.
- A manager learns to communicate performance feedback more effectively; ask his direct reports whether they understand better how they are doing and what they need to do differently, and whether the feedback motivates or discourages them.
- An associate learns to be more organized in her work; ask her manager and her clients if she meets her deadlines more consistently and produces better quality work.

Fourth, organizations that use 360° reviews already ask their customers to evaluate their employees’ performance. To evaluate the effects of training, simply ask targeted questions about behaviors relevant to the training, and ask at times that relate to the training schedule (i.e., pre/post/three month follow-up).

Why It's Smart to Measure Customer Satisfaction: Four Additional Benefits

Evaluating training in this way confers four additional benefits. First, asking external customers how happy they are with your services makes them happy. It shows that you are committed to improving your service, and that you value their patronage and don't take it for granted.

Second, it communicates to employees that their organization is serious about their development. That message can enhance the effectiveness of any training program.

Third, it communicates to employees that nothing they do is more important than increasing customer satisfaction. Competencies such as "Leadership" and "Analytical Thinking" are important, but only in the service of co-workers and external customers. In effect, organizations tell their people, "The most important job you have is making customers happy. To show you we mean it, we'll measure your success through your customer's eyes."

Fourth, it tells trainers that their work will be valued to the extent that employees do things differently after the training. For trainers accustomed to evaluation by smile sheet, this represents a revolution in accountability. Assessment drives instruction. Trainers will inevitably "teach to the test." The challenge is creating a test worth teaching to. What your customers think of you may be your ultimate business test.

We know where the keys are. Enough stumbling in shadows trying to locate them. Enough searching on the wrong side of the street. If our organizations

embrace customer satisfaction as the principal business result of training, we can liberate the lampposts. Who knows? They might dance across the street to illuminate our search for the keys to effective development.