

E Journals    electronic theses and dissertations <b>dla</b> Special Collections    University Archives © VT IMAGEBASE    Rare Books						
Digital Library & Archives	ETDs	ImageBase	Ejournals	News	EReserve	Special Collections



Current Editor:

Janet Z Burns [jburns@gsu.edu](mailto:jburns@gsu.edu)

Fall 2004

Volume 41, Number 3

[DLA Ejournal Home](#) | [JITE Home](#) | [Table of Contents for this issue](#) | [Search JITE and other ejournals](#)

## Contemporary Approaches for Assessing Outcomes on Training, Education, and HRD Programs

**Paul E. Brauchle**  
**Klaus Schmidt**  
 Illinois State University

When educators, trainers, and human resource development (HRD) practitioners are asked to produce evidence that supports the productivity of the training enterprise, they are often frustrated by the lack of simple and effective methods for assessment. Often, they resort to simple questionnaires to obtain feedback on the results of their efforts. Sometimes they just assume that if the training was based on a needs analysis or if it focused on what the company wanted, it was probably effective. However, these methods cannot easily tie training activities to the dollar values that are considered important by most organizations. This puts trainers at a disadvantage when dealing with their more financially literate colleagues. Administrators in the organization are likely to know exactly how much training costs, but they may have little idea of its real value. The training enterprise must be able to supply that information.

The costs of training are usually measured in dollars or translated to dollars, a powerful measuring scale that has enormous emotional appeal to managers. Next to a dollar measure of cost, questionnaires or assumptions based on a needs analysis often seem like weak arguments. What is needed are methods that can show the value of training in terms that managers can understand.

A number of methods have been used to assess the monetary benefits of training. Some looked at the consequences of not training employees, while others involved analyzing performance records or costing the training curve under training conditions. A well-known method is cost/benefit analysis. During a cost/benefit analysis, training costs are first calculated, then some performance values that occurred as a result of training are assessed, and an index of benefits is computed.

Swanson and Gradous (1988) derived performance values from productivity measures such as the number of items produced per shift. If more items are produced after training, the gain is calculated by subtracting post-training production figures from pre-training production figures and converted to dollars to show performance value. The benefit is then found by subtracting the cost of the training from the post-training performance value. Unfortunately, there does not seem to be a universally accepted method for performing the analysis.

Whichever methods are used to calculate the benefit of training, they are probably tacit admissions that performance improvement efforts are very important to business, and that the improvements they show, or do not show, should be considered seriously.

The training enterprise is indeed big business in America. The American Society of Training and Development (1996) reported 1995 data which indicated that employers' total expenditures on training were \$55.3 billion. By 2000, the total dollars budgeted for formal training by U.S. businesses had dropped slightly, down to \$54 billion (*Training Magazine Staff, 2000*). However, within the larger context of corporate training, some areas clearly stood out as growth centers. The corporate e-learning market was estimated to be \$1.1 billion in 2000; that figure was expected to reach \$11.4 billion by 2003 (Moe & Blodgett, 2000). Hartley (2001) noted that cooperate training was expanding rapidly throughout industry. Moe and Blodgett estimated that the global e-learning market was valued at \$300 billion and projected its expansion to \$365 billion by 2003. The 2003 ASTD *State of the Industry* report indicated that in 2002 (the most recent year in which data was available), substantial increases in training expenditures occurred (see Table 1). It seemed clear that training and development efforts represented significant investments, amounting to 2.2% of payroll in 2002 (Sugrue, 2003), up from 1.9% in 2001 (Thompson, Koon, Woodwell, & Beauvais, 2002).

**Table 1**  
*Training and Development Increases from 2001 to 2002*

Category of expenditure	2001	2002
Expenditure as a % of payroll	1.9%	2.2%
Expenditure per employee	\$734	\$826
Training hour per employee	24	28
E-learning technologies training (as a percentage of all training)	10.5%	15.4%

Source: Sugrue (2003), p. 2.

In light of these large expenditures, managers of both private and public organizations are beginning to more seriously question their ability to evaluate the impact of training. In fact, companies want to know what benefits they can reasonably expect from training (Dionne, 1996). Other authors voiced different perspectives for the increased interest in return on investment. Parsons (1995), for example, stated that it has become fashionable to analyze the financial costs and benefits of human resource development programs, whereas Holton (1995) found that the increasing global competition has led to intense pressure to demonstrate that programs are directly contributing to the bottom line of an organization. Phillips (1995) added

that the pressure to measure the return on investment (ROI) is increasing.

With so much invested, or expected to be invested, in corporate training, Blickstein (1996) raised the following questions: (a) What does management want from its training investment?, and (b) Can management's expectations be met in other than traditional bottom-line terms?

To answer these questions, a brief review of the four-level evaluation system developed by Donald Kirkpatrick is useful. In 1959, while teaching at the University of Wisconsin, Kirkpatrick proposed a model that classified training outcomes into four levels. The model was developed to provide a framework that explains evaluation. Kirkpatrick noted that some training professionals believe that evaluation means measuring changes in behavior, and others believed that the only real evaluation lies in determining final results. Still others think in terms of the comment sheets that participants complete at the end of a program. "They are all right - and yet wrong, in that they fail to recognize that all four approaches are parts of what we mean by evaluating" (Brown & Seidner, 1998, p. 95). These often-cited levels describe a rubric that can be used to evaluate programs.

### **Kirkpatrick's Model**

#### *Level 1 Evaluation (Reaction)*

Reaction may be described as how well the trainees like a particular program (Kirkpatrick, 1996). This level can be called a measure of customer satisfaction. Evaluations at this level measure how those who participate in the program react to it. Brown and Seidner (1998) asserted that effective training results in favorable trainee reactions and motivates them to learn. Abernathy (1999) summed up Level 1 evaluation by saying that it asks, "Did you like the training?" (p. 20). Many business organizations conduct effective evaluation at this level. Education tends to use it in postsecondary settings much more than at the primary or secondary levels.

#### *Level 2 Evaluation (Learning)*

The second level of the model can be described as the extent to which participants change attitudes, improve knowledge, and/or increase skills as a result of attending the particular program (Brown & Seidner, 1998). Pine and Tingley (1993) defined Level 2 in terms of measuring the content of the training. To Abernathy (1999), this level answers the question, "Did you understand the information and score well on the test?" (p. 20).

Regardless of how this level of evaluation is defined, Brown and Seidner (1998) suggested that the following questions should be asked of the trainees: (a) What knowledge was learned?, (b) What skills were developed or improved?, and (c) What attitudes were changed? If these objectives are accomplished, a change in behavior can be expected. Since it is more difficult and time consuming to measure learning, as compared to measuring trainee reactions, they proposed the following guidelines to help measure learning.

1. Use a control group if practical.
2. Evaluate knowledge, skills, and/or attitudes both before and after the program.
3. Use a paper and pencil test to measure skills.
4. Get 100% response.
5. Use the results of the evaluation to take appropriate action.

Educators have refined the second level of evaluation to a fine art, and they use it most effectively in all educational settings. Industry uses it less frequently, usually at a less sophisticated level.

### *Level 3 Evaluation (Behavior)*

Brown and Seidner (1998) defined behavior as the extent to which the change in behavior has occurred as a result of the participants' attending the training session. According to Kirkpatrick, several conditions are necessary for a behavior change to occur. The person must have a desire to change, work in the right climate, and be rewarded for changing. This level answers the question, "Did the training help you do your job better and increase performance?" (Abernathy, 1999, p. 20).

According to Blickstein (1996), the evaluations of Level 1 and Level 2 can be determined while the training program is still in session. However, Level 3 (Behavior) evaluation is much more difficult and time consuming. Brown and Seider (1998) provided the following guidelines for evaluating behavior.

1. Use a control group if practical.
2. Allow time for behavior change to take place.
3. Evaluate both before and after the program if practical.
4. Survey and/or interview one or more of the following: trainees, their immediate supervisors, their subordinates, and others who observe their behavior.
5. Get 100% response on sampling.
6. Repeat the evaluation at appropriate times.
7. Consider cost versus benefits.

Business and industry focus on this type of evaluation much more frequently than does education, with the exception of technical, vocational, and occupational programs, which must show performance gains from training.

### *Level 4 Evaluation (Results)*

This level indicates the end result from participants attending the program (Brown & Seidner, 1998). Parry (1996) believed that this level usually shows up as return on investment and thus the dollar value of the benefits of training over and above the cost of the training itself. Results-level evaluation answers the question, "Did the company or department increase profits, customer satisfaction, and so forth as a result of the training?" (Abernathy, 1999, p.20). According to Kirkpatrick's model:

Final results can include increased production, improved quality, decreased costs, reduced frequency and/or severity of accidents, increased sales, reduced turnover, and higher profits and return on investment. It has long been thought important to recognize that the most desirable approaches to delivering instruction (training) are those that are the most effective in terms of results and the most efficient in terms of cost. (Parry, 1976, p.74)

Parry appeared to agree with Blank (1982), who had stated as early as 1982 that "we need to strike a balance between effectiveness (does it work?) and efficiency (how much does it cost?)" (p. 192). Kirkpatrick (1994) then proposed the following guidelines to help evaluate at the results level of the model (these are similar, but not identical to, his proposals for Levels 2 and 3).

1. Use a control group if practical.
2. Allow time for results to be achieved.
3. Measure both before and after the program if practical.
4. Repeat the measurement at appropriate times.
5. Consider cost versus benefits. Be satisfied with evidence if proof is not possible.

At this level, business and industry are clearly the most sophisticated users of evaluation. Educators have traditionally used surveys, interviews, or anecdotal evidence to assess long-term satisfaction with programs; and they do not often try to express their program outputs in terms of monetary value. However, technology is available (e.g., Unemployment Insurance [UI] Wage studies) to enable educators to show the long-term economic value of their programs (more about this later).

Even though it has come under criticism (Holton, 1996), the four-level evaluation model has been acknowledged as the standard in the training field because of its simplicity and its ability to help people think about evaluative criteria (Alliger & Janak, 1989). Kirkpatrick's model has provided a vocabulary and a rough taxonomy that has clearly met an organizational need, and it has become well known in HRD departments around the country. It has been a highly successful framework that has been used in the training evaluation field for nearly 50 years. Recent research has focused on the charge that the four-level model often stops short of reaching meaningful long-term results (Bernthal, 1995). In a follow-up, Phillips (1997a, 1997b) suggested a number of modifications to the model, including the addition of a fifth level that specifically addresses ROI.

It is probably safe to say that both education and industry have much to contribute to the evaluation process. There are lessons to be learned from both sectors across the evaluation spectrum. Both are good at conducting Level 1 (Reaction) evaluation. Educational practice can teach us how to use Level 2 (Learning) evaluation with greater precision. Both education and industry conduct Level 3 (Behavior) evaluation; and even though only 9% to 11% actually conduct evaluations at Level 4 (Results), industry does the best job.

### **Return on Investment (ROI)**

Return on investment has been a critical issue for trainers and top executives in recent years and is a topic frequently listed on meeting agendas. This technique probably should receive more emphasis from educators than it has in the past. It has been continuously discussed in the literature (Phillips, 1996a), though not without controversy. For example, some professionals argue that it is not possible to calculate the ROI in education and training contexts, while others develop measures and ROI calculations anyway (Phillips, 1997b). Most authors appear to think that ROI calculations and related efforts belong in Level 4 evaluation, but Phillips (1997a) described examples of ROI calculations at each of the four levels. So, what is the current status of ROI in the field of performance improvement?

#### *Current Status of ROI*

Phillips (1996a) suggested that it is difficult to pinpoint the state of ROI in the profession. Some HR managers are reluctant to discuss internal practices, and it is difficult to find case studies that specifically list the strategies used by training departments in determining ROI. However, a study of more than 40 organizations found that the best companies are measuring customer requirements, testing participants, measuring what the client can and will pay for, and moving away from justification (Dixon, 1996). Dixon went on to state that one of the most striking findings was that none of the best-practice organizations was evaluating primarily to justify training or to maintain the training budget. They were selectively conducting evaluations at Levels 3 and 4, but not on a regular basis. Geber (1995) noted, with Dixon, that it would be impractical for most companies (or educational institutions) to do Level 3 and Level 4 evaluations on every single course.

Recognizing this problem, the American Society for Training and Development (ASTD) collected information on this subject from more than 2,000 HRD professionals. The results came in the form of two publications that purported to describe how return-on-investment measurements were conducted in real-life situations (Phillips, 1994, 1997c). The case studies in this project represented a variety of settings, strategies, and approaches in manufacturing,

service, and government organizations. Respondents varied from employees to managers and specialists in the training field. ROI in the companies studied ranged from 150% to 2,000%. However, educators were not part of this study, though they probably should have been.

As previously mentioned, it is possible to consider ROI a fifth level to Kirkpatrick's four-level model. In addition to the reaction level, learning level, behavior level, and the results level, a return-on-investment level would be added that compares the training's monetary benefits with the costs. A model had been developed that tracked the steps in measuring ROI from collecting post-program data to calculating the actual return (Phillips, 1996a). The model can compare training costs to monetary benefits, assuming that all training programs have reportable results as well as intangible benefits; and most do have some sort of identifiable outcomes.

### *Basic Process for Calculating ROI*

Phillips (1996c) provided the following basic process for calculating ROI.

1. Collect Level 4 evaluation data.
2. Ask, "Did on-the-job application produce measurable results?"
3. Isolate the effects of training from other factors that may have contributed to the results.
4. Convert the results to monetary benefits.
5. Total the costs of training.
6. Compare the monetary benefits with the costs.

When discussing the general concept of ROI, it is important to note the need for identifying the desired program outcomes and using the design of the evaluation to inform program planning. This is equally true in training and in education. Putting the necessary time and resources in return-on-investment analysis makes sense only if one is convinced that the training or education program was applied correctly in the right place.

### *ROI Evaluation Recommendations in the Literature*

Several recommendations emerged as a result of these and other case studies. The first is that "targets should be set for each evaluation" (Phillips, 1996a, p. 44). Some organizations set targets for each level of their training programs. For example, some organizations require 100% of their training programs to be evaluated at Level 1, 40% to 70% of their training programs at Level 2, and so forth. When an organization sets targets for accountability, it is sending a powerful message to the HRD department about the need for measurement and evaluation (Phillips). The need to set targets was part of a common belief that training design should occur at the same time as a measurement discussion (Williams, 1996). Simultaneously working on training design and measurement planning is more effective because the information required for training design is exactly the same information needed for solid measurement.

A third recommendation was that evaluation should be focused on the micro level. Respondents focused on a single program or a few tightly integrated programs in order to have an effective ROI measurement. However, the results of the study indicated that ROI measurement was more effective when applied to a single program that can be linked to a direct payoff (Phillips, 1996a).

A fourth recommendation dealt with the methods of collecting information. Phillips (1996) advocated that a variety of methods should be used to collect data. These methods could include interviews, focus groups, and questionnaires, but also action plans, contracts, and performance monitoring. Respondents used more than one or two practices to collect data

because they recognized that programs, settings, and situations were different. In order to increase objectivity, Bernthal (1995) had suggested using several different measurement methods in the same evaluation or conducting several evaluations using different methods or approaches. The key message here is that one size does not fit all.

A fifth recommendation focused on activities that are necessary precursors to ROI evaluation. Before attempting to conduct return-on-investment studies, it is necessary to isolate the effects of training from other factors that can affect business results (Phillips, 1996, 1997b). Most of the time, improvements in job performance are only partially due to training programs. Variables other than training, such as trainees' age and work experience, seasonal sales patterns, economic changes, shifts in managerial styles, equipment breakdowns, customer attitudes, etc., may influence the data and make it difficult to determine the actual effect of the training upon the ROI results (Shelton & Alliger, 1993). They further indicated that a way to measure the effects of extraneous factors is to compare the results of a control group with the results of the trainee group. Trend-line analysis, forecasting, participant estimation, supervisor estimation, management estimation, customer input, expert estimation, subordinate input, and other factors are additional ways to isolate training's effect on performance (Phillips, 1996b).

A final recommendation advised organizations that had little or no experience in calculating ROI to measure only one course at a time (Phillips, 1996a, 1996c). It is not practical for most organizations to calculate the ROI on all training programs because the result would have a large number of calculations and would be very time consuming. For this reason, most organizations take a practical approach to the problem and focus on one or two of the most important or popular training programs. If presented well, ROI calculations on just a few programs can be powerful. The respondents participating in Phillips's studies were not content to show that training can result in such improvements as increased productivity and decreased turnover. An additional step was taken to convert these improvements to monetary values that could be compared to costs for ROI calculation. For hard-data items, such as productivity, quality, and time, the ROI calculation was much easier than those calculations for soft-data items, such as customer satisfaction, employee turnover, and job satisfaction (Phillips, 1996c).

The following sections will identify and discuss some approaches that individuals and organizations have used to assess individual performance improvement and program outcomes.

### **Calculating the Return on Investment**

There are two common formulas used to calculate the return on investment (Phillips, 1996b). The first is benefit/cost ratio (BCR), and the second is ROI. The benefit/cost ratio (BCR) can be calculated using the following formula:  $BCR = \text{program benefits} / \text{program costs}$ . It uses the total benefits and the total costs to obtain an index of the worthwhile results of the training program to its overall cost. ROI is different because it gives a percentage of return on the dollars that are invested in training and development. This figure is interpreted in the same way the returns of an investment in stocks or mutual funds would be viewed. To get ROI, the training costs are subtracted from the total benefits to get the net benefits, and then the net benefits are divided by the costs. The formula for this is  $ROI (\%) = \text{net program benefits} / \text{program costs} \times 100$ . Phillips (1996a) gives the following example: Suppose a training program produces benefits of \$321,600 with a cost of \$38,233. The BCR is 8.4. For every \$1 invested, \$8.40 in benefits was returned. The net benefits were  $\$321,600 - \$38,233 = \$283,367$ . ROI is  $\$283,367 / \$38,233 = 741$ . Using the ROI formula, for every \$1 invested in the program, there was a return of \$7.41 in net benefits.

Another useful method is called payback period (Phillips, 1997a). This technique usually makes the assumption that the cash proceeds generated by a training intervention are constant

over time, and it calculates the time period needed to pay back the original investment. The formula is: Payback period = Total investment/net annual savings. In the example above, the total investment is \$38,233, and the net benefits are \$321,600. If there is no time period specified, it can be safely assumed that the net benefits are for a period of one year, because budgeting is usually done on an annual basis. Using these figures with the formula produces an answer of .1188837 years, or 43 days. In this instance, the original training investment was paid back within 43 days. These methods are widely used by industry. They have potential for education if the economic value of educational programs can be established.

### The Time Value of Money

A thorough consideration of a training or performance improvement initiative should also include the time-value of the money that is saved (or made). This concept has been used for years by production/operations managers (Riggs, 1987), but it is not often found in the training evaluation literature. Suppose that ROI calculations indicate that after three years, a training intervention will make the company \$100,000. Someone now wants to know the present value of those future dollars, knowing intuitively that money to be received in three years is worth less than money received now. The general formula for the present value of future money is  $P=F/(1+i)^n$  (Riggs, 1987, p. 132). Here,  $P$  is present dollars;  $F$  is future dollars;  $I$  is a constant;  $i$  represents the opportunity cost (the amount of interest that could have accrued on that money now instead of three years from now); and  $n$  is the number of years being considered. Using the same \$100,000 value, and assuming 10% per year on the money, the result is  $P=100,000/(1+.10)^3$ , or \$75,131.48. It can be seen that the \$100,000 in three years is worth only about 75% of what it would be worth if received today.

On the other hand, the same \$100,000 must be invested now for a program that will break even in three years. You want to know the future value of those present dollars. Riggs (1987) provided the formula for calculating the future value of present dollars as  $F=(1+i)^n$ . So  $F=(1+.10)^3$  works out to \$133,100. If those dollars had been invested at 10% per year, in three years they would have been worth \$133,100. Therefore, the value of the training program after three years must exceed \$133,100 in order to show any gain. This concept is sometimes called a "hurdle rate" (Hendrick & Moore, 1985, p. 52). A training intervention, or educational or HRD program, must equal or exceed that figure in value; or it cannot be asserted that it has had any real benefit.

### Utility Analysis

This method measures the value of job performance after training. One version is based upon a method used by Godkewitsch (1987) that was later published by ASTD in INFO-LINE publication #007 (1990), entitled "How to Conduct a Cost-Benefit Analysis". The formula for training benefits is  $B = N[(E \times V) - C]$ , and the result is a dollar amount that describes what the training is worth to the organization in terms of the performance of the workers who have participated in the training (Brauchle, 1995). In this formula,  $N$  is the number of people trained;  $E$  is the effect of the training in standard deviation units;  $V$  is the dollar value of the effect; and  $C$  is the cost of the program. Suppose a training intervention has raised the average performance of workers +.42 standard deviations, and the dollar value of that effect was \$9,600. There were 25 workers trained at a per-person cost of \$2,502. Using this information with the formula  $B = N[(E \times V) - C]$ ,  $B = 25 [(.42 \times \$9600) - \$2,502]$ , which works out to a benefit of \$38,250. This means that the training program had a benefit of \$38,250.

The advantages of using this method are as follows.

1. It uses dollars as the unit of measure for the analysis.
2. This unit of measurement is usually acceptable to managers because they understand it.
3. The method is relatively simple to use.

4. It does not require extensive computer capability or number crunching power.
5. It provides results fairly quickly, without waiting for a one-year budget report.

On the other hand, the disadvantages of using this method are as follows.

1. This a norm-referenced method, with several attendant problems.
2. It relies on supervisor evaluations, which may be of doubtful validity and even more questionable reliability.
3. The value of one standard deviation improvement (40 percent of annual salary, based on other studies) is questionable, and should not be used if there are local benchmarks against which to compare it.
4. The dollars calculated as benefits are not real dollars unless there are validated benchmarks.
5. It may be difficult to tie the benefits to long-term business results if the measures of productivity gain are taken shortly after training.
6. The results do not answer the question, "Did you train for the right skills?" (Brauchle, 1995)

In summary, this method is thought to be ". . . probably worth using if you can establish your own benchmarks for the value of one standard deviation of productivity gain. (It) is less complicated than many others and may provide useful justifications for the training enterprise" (Brauchle, 1995, p. 17-22).

### 360-Degree Appraisal Feedback

A considerable amount of money is spent annually on supervisor and manager training, and some related literature has focused on the concept of 360-degree feedback as a method of evaluating the degree to which executive performance has improved. Antonioni (1996) defined 360-degree feedback as a process whereby "...individuals appraise themselves and also receive appraisal feedback from their appraisers: immediate supervisor, peers, and from direct contributors if they are managers" (p. 72). This methodology, according to Charney and Conway (1998), "...is a performance evaluation system that uses input from all employee levels to assess performance. Training is an important factor in helping employees use the process effectively" (p. 196). In recent years this process has been popular enough to support the generation of considerable software in support of it (Bracken, Summers, & Fleenor, 1998; Fried, 1998; Meade, 1999; Ellis, 2001). It is believed to have at least five desirable outcomes:

1. Increased awareness of the appraisers' expectations;
2. Improvements in work behaviors and performance;
3. Reduction of undiscussables, specifically the appraisers' feelings and perceptions about undesired behaviors of those being appraised;
4. Increase in effective periodic informal 360-degree performance reviews; and
5. Increase in organizational learning (Antonioni, 1996).

An effective 360-degree feedback process offers a major opportunity for organizational members to improve the quality of their work relationships. It can help appraisers and those being appraised effectively define the quality requirements in their work relationships. The model teaches individuals how to give and receive constructive feedback, and provides a structure for discussing the undiscussables (Antonioni, 1996). It is thought to not only improve work behaviors, but also to increase worker performance, which in turn should improve return on investment for the organization. However, Antonioni did not explain exactly how this process could actually be translated into monetary outcomes. Church and Bracken (1997) attempted to relate the results of 360-degree feedback to organizational performance or results. Although 360-degree feedback is generally accepted as a good evaluation system that has the potential to contribute to improved performance, it has not yet been conclusively shown how this method can yield monetary returns on education or training programs. However, it may

have some useful application in education. For example, suppose teachers are evaluated by administrators, peers, and students. While it does not readily provide monetary data, this approach might put in perspective, or balance out, negative evaluations by any of these sources.

### **Performance-Learning-Satisfaction Evaluation System (PLS)**

Another model thought to be useful in evaluating how well training has improved executive or supervisor performance, the PLS model embraces the domains of performance, learning, and satisfaction. It is thought to be both rigorous and flexible: rigorous in terms of core questions, techniques, and reporting results, and flexible enough for most applications. It also includes an auxiliary data processing program (Swanson & McClernon, 1996). When using this system, performance is considered in terms of two factors: (a) business results at the organizational, process, or individual levels; and/or (b) financial results or benefits in terms of money or monetary ratios. Swanson and McClernon asserted that most business results can be "monetized" (p. 719) and expressed as financial data. The general financial results goal of the model is to have benefits exceed the costs by a 2:1 ratio. Once the 2:1 return on investment goal is achieved, it can be said that the program achieved 100% of its goal. In that sense it appears to have more application value than 360-degree feedback.

In the PLS System, learning has two components: (a) knowledge demonstrated on tests and other measurements, and/or (b) expertise in demonstrating simulated or actual workplace environments. Satisfaction is seen in terms of perceptions of the behaviors that are demonstrated. Trainees are expected to meet learning and expertise standards. Swanson and McClernon (1996) further state that if the standard is a rating of two on a one-to-three scale, an average score of two by the trainees would represent 100% attainment of the goal. Attention to raising the goal should occur only after there is clear evidence that existing performance and learning goals are being reached. Although the PLS Evaluation System depicted in the 1996 manuscript resulted in an 8:1 financial return on investment in less than a year, this method does not seem to have been used much in recent years.

### **The Balanced Scorecard**

The balanced scorecard, developed by Kaplan and Norton, has been getting serious looks by players in the financial- and strategic-planning world (Willyerd, 1997). This method tracks a company's strategy and helps link its financial budgets to its strategic goals (Kaplan & Norton, 1996). Could it be adapted for use in educational institutions?

The balanced scorecard deals with four key areas of performance and provides answers to the following questions.

1. How do we look to our shareholders?
2. How do customers see us?
3. What must we excel at?
4. Can we continue to improve and create value?

This approach helps ensure that all of the critical performance measures are evaluated in addition to return-on-investment issues. It serves the purpose of a check and balance so that one area is not overemphasized at the expense of another. Unlike the PLS System, this method has enjoyed numerous recent applications in strategic planning (Kaplan & Norton, 2001; Kaplan & Norton, 2000; Anonymous, 1999) and in the evaluation of program effectiveness (Novak, 2000; Abernathy, 1999). However, it has been criticized by Forbes (2000) for not addressing the external and internal indicators that measure the value-added elements of the system. Forbes apparently thought the addition of these factors would create a more robust system.

### **HRD Benefit-Forecasting Model**

The basic HRD benefit-forecasting model uses three components to determine financial value: the performance value to result from the HRD program, the cost of the HRD program, and the benefit resulting from the HRD program (Swanson & Gradous, 1988). Performance value is the dollar value of the performance units resulting from an HRD program. Benefit is calculated by multiplying the total number of units expected to result from the program by the dollar-value amount of one unit. Cost is an item of outlay that is incurred in the operation of HRD programs; it may include such items as direct and indirect costs, fixed and variable costs, and amortization. HRD cost is any expenditure that the organization chooses to attribute to an HRD program.

This approach is useful in that it contributes useful metrics for concepts like performance unit, performance value, and benefit. It is not a post hoc method of evaluation, but it does not seem to suffer much from this condition. Forecasting benefits can yield a realistically derived dollar value expected from an HRD or training program. From the numbers on a forecasting worksheet, other measures such as expected ROI and payback time can be calculated. The projected numbers can then be compared to the actual numbers. This kind of comparison can make forecasting efforts much more accurate and believable. Could this be used in education? Possibly, if performance unit and performance values can be quantified for teachers and students.

### **The Relative-Aggregate-Scores Approach**

This approach, like the time-value-of-money method previously discussed, is borrowed from production operations management (POM). Gaither (1996) used it to help compare alternative locations for a production facility. It can be used to compare the relative value of training for different tasks or duties within a job, and/or for assessing the gain in value for that job after training. When the relative aggregate scores approach is used to assess the benefits of training, characteristics such as the difficulty and the importance of various job tasks are assessed, usually on a 1-4 scale. For each job task, weights are derived by multiplying difficulty ratings times importance ratings times frequency figures that are expressed as a percentage of time spent on a task. These products are converted to a percentage of the total, and the percentages become the weights for each job task. Then, actual data are used to show a performance value for each task. The performance value for each task is calculated by multiplying each weight times the total salary for the job. Ratings by supervisors before and after training enable the computation of pre-and post-performance values for each task, as well as a performance gain for each task and for the job as a whole (Brauchle, 1992).

The results of this approach can be used in several ways. The weights for each job task can help in creating a kind of a Pareto analysis of the job so it can be seen which tasks are the most significant. They can also indicate where it may be desirable to place training or HRD dollars in order to obtain the best results. The relative aggregate scores approach necessitates weighing and considering what various portions of the job are actually worth, and provides actual figures for a gain in performance value. It can help in monitoring the results of training and in planning training so that limited resources are invested in the most important areas. Technical educators could use it to assess the value of mastery of various tasks in a program. Administrators could use it to assess the value of teacher improvement in various categories from in-service training programs.

### **Unemployment Insurance (UI) Wage Studies**

Most of the preceding methods for evaluating HRD programs are good for microanalysis, focusing on just one individual, program, or job; but macro-analysis methods, those that cast a much wider net, are also available. One of these is known as the UI Wage Study (Kornfeld & Bloom, 1999). In a UI Wage Study, blocks of independently gathered data are merged and are

subjected to post hoc analysis to ascertain whether individuals' post-training salaries showed a gain from their pre-training salaries: The use of salaries as a measure of program effectiveness is appealing for several reasons.

1. Salaries use dollars as the unit of measure for the analysis.
2. The dollar is a ratio measure. For example, \$100 is twice as much as \$50. An interval measure, like that used in the questionnaire, does not permit such statements. Just because a thermometer uses an interval scale, it cannot be said that a temperature of 100 degrees is twice as hot as 50 degrees. The ratio scale enables the expression of differences with more precision, therefore making more meaningful statements possible.
3. The use of dollars as a measurement is usually acceptable to managers because they understand that metric and frequently use it to compare options.

The UI Wage Study is based on the notion that most organizations have some kind of computerized record system for tracking their people. For example, every educational institution has some sort of number assigned to each student. Linked to that number are the data on what courses were taken and when, as well as other information. Other sources of data are the Unemployment Insurance Wage Reports that are posted quarterly in each state. If the Social Security number of an individual is known and can be matched to educational records, the gains in salary for individuals after training can be computed by comparing educational records with Unemployment Insurance Wage Reports. Salary gains can be calculated for individuals, by program, or for completers and non-completers.

This kind of study has several advantages over other commonly used methods of evaluating training programs. First, it is less expensive than costly and cumbersome surveys. Second, it is as good or better at obtaining high-quality data and not nearly as costly (Kornfeld & Bloom 1999). Third, it appears to be able to much more accurately represent a population than traditional survey based approaches. Sanchez, Laanan, and Wisely (1999) found that from 71% to 84% of occupational program completers had been found in UI Wage data bases. This makes them very useful for analyzing the effects of occupational programs and should be beneficial for other programs as well. For example, the City Colleges of Chicago studied student earnings in a 1997 cohort of students from Truman College (Brauchle & Hastings, 2003). The district was able to track changes in earnings for students who had taken individual courses, received certificates, or obtained associate degrees. They were also able to track earnings by gender, by age, for students from an economically disadvantaged background, and by vocational major. It appears that this method of establishing the monetary benefits of training or HRD programs has enormous potential to provide compelling evidence for learner and program achievement.

### **A Perspective on Evaluation Methods**

Many HRD practitioners believe that the ultimate level of organization, return on investment, shows the true contribution of training (Phillips, 1996). The process is not complete "...until the results have been converted to monetary values and compared with the cost of the program" (p. 20). Four distinct and important benefits can be derived from the implementation of effective evaluation measures in an organization. First of all, ROI will measure the contribution the program made to the organization and will determine if it was a good investment. Second, this calculation will determine which programs contribute the most to the organization and allow priorities to be established for high-impact training programs. Third, the evaluation brings a focus upon the results of all programs, not just those targeted for the financial evaluation. Fourth, this process can help convince management that training or education is an investment and not just an expense (Phillips, 1997b). Another benefit, this one not mentioned by Phillips, is that most of these methods express the results of training programs in terms of dollars, a metric that is of common interest to managers and decision makers.

### Conclusion

"One of the most common criticisms of trainers is that they do not measure the organization's return on investment in training" (Mendoza, 1995, p. 66). "The statistics everyone wants, those that would tell us the return on training dollars spent, has proven to be stubbornly elusive" (Fagiano, 1995, p. 12). A number of methods have been described in this manuscript that help organizations evaluate the benefits of training, including several models that have been used, or are currently being used, to evaluate return on investment. These methods are generally

**Table 2**  
*Summary of Assessment Methods*

Methodology	Data	Results	Applied	Level	Value
Benefit/cost ratio	Hard	Gives ratio of cost to benefits	Before or after	4	High
Payback period	Hard	Provides time to payback of initial investment	Before or after	4	High
Return on true value of dollars	Hard	Given % return on initial investment	Before and after	4	High
Present value of dollars and future value of dollars	Hard	Accounts for value of future savings in present dollars	Before	4	High
Utility analysis	Hard	Net value of HRD	After	4	Med
360-degree feedback	Soft	Provides estimates of performance improvement from various sources	After	3/4	Low
Performance team satisfaction	Semi-soft	Estimates knowledge and expertise developed from HRD program	After	3/4	Low
Balanced scorecard	Soft	Relates various performance measures to strategic issues	Before and after	3/4	Low
HRD benefit forecasting	Hard	Estimates relative value of HRD program approaches	Before	3/4	High
Relative aggregate scores	Hard	Computes relative values of job and task value	After	4	High
UI Wage Studies	Hard	Provides actual earnings one, two, and three years out	After	4	High

thought to fit within the "Results" level of Kirkpatrick's four-level evaluation model. Most HRD, training, and educational professionals agree with Parry (1996) that "...training doesn't cost ...it pays, and HRD is an investment, not an expense" (p. 72). These approaches are being used more and more by companies to show that there is a high return on the investment made in training and development (Purcell, 2000). Education might be well served to take note of this and to become better at using these techniques. As Table 2 indicates, a wide variety of methods exist to show the value of education and training; and each has its advantages and disadvantages. A careful mixture of these approaches is likely to be useful for most evaluation problems.

As early as 1985, Laird had stated that "client-managers need measurement as an indication that they're solving or eliminating performance problems...getting something back on the training investment" (p. 241). In today's economy with down- sizing and world-wide global competition among public and private organizations, it is very important to be able to justify training expenses, as well as all expenses of these organizations. All of these expenses

must relate to organizational performance growth, profit, market share, etc.; and the dollars must contribute to this bottom-line measurement.

### References

- Abernathy, D. (1999). Thinking outside the evaluation box. *Training and Development*, 53(2), 18-23.
- Alliger, G., and Janak, E. (1989). Kirkpatrick's levels of training criteria: Thirty years later. *Personnel Psychology*, 42(2), 331-342.
- American Society for Training and Development (1996). How to conduct a cost-benefit analysis. *Info-Line*, #007. Alexandria, VA: Author.
- Anonymous. (1999) Strategy-focused groups at all levels are embracing the balanced scorecard. *The New Corporate University Review*, 7(6), 16-17.
- Antonioni, D. (1996). Designing an effective 360-degree appraisal feedback process. *1996 Conference Proceedings of the Academy of Human Resource Development*, Minneapolis, MN, February 29-March 3, 1996.
- Berntal, P. (1995). Evaluation that goes the distance. *Training and Development*, 49(9), 41-45.
- Blank, W. (1982). *Handbook for developing competency-based training programs*. New Jersey: Prentice Hall.
- Blickstein, S. (1996). Does training pay off? *Across the Board*, 33(6), 16-20.
- Bracken, D., Summers, L., & Fleenor, J. (1998). High-tech 360. *Training and Development*, 52(8), 42-45.
- Brauchle, P. (1992). Costing out the value of training. *Technical and Skills Training*, 3(4), 35-40.
- Brauchle, P. (1995). Interpreting a compact cost/benefit analysis that assesses the performance value of training. *1995 Conference Proceedings of the Academy of Human Resource Development*, St. Louis, MO, March 2-March 5, 1995.
- Brauchle, P. E., & Hastings, J. H. (2003). Using unemployment wage data to assess educational and economic outcomes in a multicultural inner-city community college. *Journal of Applied Research in the Community College*, 10(1), 57-69.
- Brown, S., and Seidner, C. (1998). *Evaluating corporate training: Models and issues*. Boston/Dordrecht/London: Kluwer Academic Publishers.
- Charney, C., and Conway, K. (1998). *The trainer's tool kit*. New York: American Management Association.
- Church, A. H., & Bracken, D. W. (1997). Advancing the art of 360-degree feedback. *Group and Organization Management*, 22(2), 149-161.
- Dionne, P. (1996). The evaluation of training activities: A complex issue involving different stakes. *Human Resource Quarterly*, 7(3), 279-286.

- Dixon, N. (1996). New routes to evaluation. *Training and Development*, 50, 82-85.
- Ellis, R. (2001). Software roundup. *Training and Development*, 55(6), 94.
- Fried, N. (1998). 360 software shootout. *HRMAGAZINE*, 43(3) supplement, 8-13.
- Fagiano, D. (1995). Making training quantifiable. *Supervision*, 56(6), 12-13.
- Forbes, R., (2000). HPI soup: A response. *Training and Development*, 54(6), 26-36.
- Gaither, N. (1996). *Production and Operations Management* (7th ed.). Belmont, CA: Duxbury Press.
- Geber, B. (1996). Does training make a difference? Prove it! *Training*, 32(3), 27-34.
- Godkewitsch, M. (1987). The dollars and sense of corporate training. *Training*, 24(5), 79-81.
- Hendrick, T., & Moore, F. (1985). *Production/Operations Management* (9th ed.). Homewood, IL: Richard D. Irwin, Inc.
- Hartley, D. (2001) E-evaluation: Pricing e-learning. *Training and Development*, 55(4), 24-27.
- Holton, E. (1996). The flawed four-level evaluation model. *Human Resource Quarterly*, 7(1), 5-20.
- Holton, E. (1995). In search of an integrative model for human resource development evaluation. *1995 Conference Proceedings of the Academy of Human Resource Development*, St. Louis, MO, March 2-March 5, 1995.
- Kaplan, R., and Norton, D. (1996). Using the balanced scorecard as a strategic management system. *Harvard Business Review*, January-February, 75-85.
- Kaplan, R., & Norton, D. (2000). Having trouble with your strategy? Then map it. *Harvard Business Review*, 78(5), 167-176.
- Kaplan, R., & Norton, D. (2001). *The strategy-focused organization*. Boston, MA: Harvard Business School Press.
- Kornfeld, R., & Bloom, H. (1999). Measuring program impacts on earnings and employment: Do unemployment insurance wage reports from employer agree with surveys of individuals? *Journal of Labor Economics*, 17(1), 168-97.
- Kirkpatrick, D. (1994). *Evaluating training programs: The four levels*. San Francisco: Berrett-Koehler Publishers, Inc.
- Kirkpatrick, D. (1996). Great ideas revisited. *Training and Development*, 50(1), 54-59.
- Laird, D. (1985). *Approaches to training and development* (2nd ed.). MA: Addison-Wesley.
- Meade, J. (1999). Visual 360: A performance appraisal system that's "fun." *HRMAGAZINE*, 44(7), 118-122.

- Mendosa, R. (1995). Is there a payoff? *Sales Marketing Management*, 47, 64-71.
- Moe, M., & Blodgett, H. (2000). Merrill Lynch & Co., Global Securities Research and Economics Group, Global Fundamental Research Department, in *ASTD Annual Report*. Alexandria, VA: Author.
- Novak, C. (2000). HPI balanced scorecard. *Info-Line*, #250010. Alexandria, VA: American Society for Training and Development.
- Parry, S. (1996). Measuring training's ROI. *Training and Development*, 50(5), 72-77.
- Parsons, J. (1995). The impact of values on the financial analysis of human resource development. *1995 Conference Proceedings of the Academy of Human Resource Development*, St. Louis, MO, March 2-March 5, 1995.
- Phillips, J. (Ed.) (1994). *In action: Measuring return on investment* (Vol. I). Alexandria, VA: American Society for Training and Development.
- Phillips, J. (1995). Return on investment - Beyond the four levels! *1995 Conference Proceedings of the Academy of Human Resource Development*, St. Louis, MO, March 2-March 5, 1995.
- Phillips, J. (1996 a). ROI: The search for best practices. *Training and Development*, 50(2), 42-47.
- Phillips, J. (1996b). Was it the training? *Training and Development*, 50(3), 28-32.
- Phillips, J. (1996c). How much is the training worth? *Training and Development*, 50(4), 20-24.
- Phillips, J. (1997a). *Handbook of training evaluation and measurement method* (3rd ed). Houston: Gulf Publishing Company.
- Phillips, J. (1997b). *Return of investment in training and performance improvement programs*. Houston: Gulf Publishing Company.
- Phillips, J. (Ed.). (1997c). *In action: Measuring return on investment* (Vol. II). Alexandria, VA: American Society for Training and Development.
- Pine, J., & Tingley, J. (1993). ROI of soft-skill training. *Training*, 30(2), 55-58.
- Purcell, A. (2000). 20/20 ROI. *Training and Development*, 54(7), 28-32.
- Riggs, J. L. (1987). *Production systems: Planning, analysis, and control*. (4th ed.). New York: John Wiley & Sons.
- Sanchez, J., Laanan, F., & Wiseley, W. (1999). Post college earnings of former students of California community colleges: Methods, analysis, and implications. *Research in Higher Education*, 40(1), 87-113.
- Shelton, S., & Alliger, G. (1993). Who's afraid of level 4 evaluation? *Training and Development*, 47(6), 43-46.
- Sugrue, B. (2003). *State of the industry: ASTD's annual revue of U.S. and international trends in workplace learning and performance*. Alexandria, VA: American

Society for Training and Development.

Swanson, R., & Gradous, D. (1988). *Forecasting financial benefits of human resource development*. San Francisco: Jossey-Bass Publishers.

Swanson, R., & McClernon, T. (1996). PLS evaluation system: Sales communication case study. *1996 Conference Proceedings of the Academy of Human Resource Development*, Minneapolis, MN, February 29-March 3, 1996.

Thompson, C., Koon, E., Woodwell, W. H., Jr., & Beauvais J. (2002). *Training for the new economy: An ASTD state of the industry report*. Alexandria, VA: American Society for Training and Development.

Training Magazine Staff (2000). Industry report 2000: A comprehensive analysis of employer-sponsored training in the United States. *Training Magazine*, 37(10), 45.

Williams, L. (1996). Measurement made simple. *Training and Development*, 50, 43-45.

Willyerd, K. (1997). Balancing your evaluation act. *Training*, 34(3), 52-58.

---

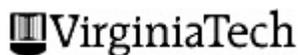
Brauchle is Professor and Schmidt is Assistant Professor in the Department of Technology at Illinois State University in Normal, Illinois. Brauchle can be reached at [pebrauc@ilstu.edu](mailto:pebrauc@ilstu.edu). Schmidt can be reached at [kschmid@ilstu.edu](mailto:kschmid@ilstu.edu).

---

[DLA Ejournal Home](#) | [JITE Home](#) | [Table of Contents for this issue](#) | [Search JITE and other ejournals](#)

---

Send questions or comments to:  
[DLA, University Libraries](#)  
[Virginia Tech](#), P.O. Box 90001,  
 Blacksburg, VA 24062-9001



<a href="#">Digital Library &amp; Archives</a>	<a href="#">ETDs</a>	<a href="#">ImageBase</a>	<a href="#">Ejournals</a>	<a href="#">News</a>	<a href="#">EReserve</a>	<a href="#">Special Collections</a>
------------------------------------------------	----------------------	---------------------------	---------------------------	----------------------	--------------------------	-------------------------------------

URL: <http://scholar.lib.vt.edu/ejournals/JITE/v41n3/brauchle.html>  
 Last modified on: 11/04/05 16:36:22 by Daniel Culpepper