

INFLUENCE OF PRE AND POST TESTING ON RETURN ON INVESTMENT  
CALCULATIONS IN TRAINING AND DEVELOPMENT

Calvin H. Hiraoka, B.Ed., M.Ed., M.B.A.

Dissertation Prepared for the Degree of  
DOCTOR OF PHILOSOPHY

UNIVERSITY OF NORTH TEXAS

May 2008

APPROVED:

Jerry Wircenski, Major Professor

Jeff Allen, Minor Professor

Lynn Johnson, Committee Member

Jeff Allen, Chair of the Department of Learning  
Technologies

M. Jean Keller, Dean of the College of Education

Sandra L. Terrell, Dean of the Robert B. Toulouse  
School of Graduate Studies

Hiraoka, Calvin H. Influence of pre and post testing on return on investment calculations in training and development. Doctor of Philosophy (Applied Technology and Performance Improvement), May 2008, 118 pp., 22 tables, references, 107 titles.

When expenses become an issue, training is often one of the first budget items to be cut. There have been a number of evaluation studies about rates of return from training interventions. Most results are based on interviewing participants about the value of the intervention and its effect on their productivity. This often results in quadruple digit return on investment indications. Decision makers who control the budget often view these kinds of results with skepticism.

This study proposes a methodology to evaluate training interventions without asking participants their opinions. The process involves measuring learning through a series of pre-tests and post-tests and determining if scores on pre-tests can be used as predictors of future return on investment results. The study evaluates a series of return on investment scores using analysis of variance to determine the relationship between pre-tests and final return on investment results for each participant. Data is also collected and evaluated to determine if the financial results of the organization during the period of the training intervention could be correlated to the results of the training intervention.

The results of the study suggest that the proposed methodology can be used to predict future return on investment from training interventions based on the use of pre-tests. These rates of return can be used as a method of selecting between competing training intervention proposals. It is a process that is easily understood by the key decision makers who control the allocation of financial resources. More importantly, it is a process that can maximize the value of each dollar spent on training.

Copyright 2008

by

Calvin H. Hiraoka

## ACKNOWLEDGEMENTS

I would like to acknowledge the assistance of Professors Jerry Wircenski and Jeff Allen from the Department of Learning Technologies and Robin Henson from the Department of Educational Psychology, and thank them for the knowledge they shared and the long hours spent in the classroom to absorb the principles and practices related to the field of Applied Technology Training and Development.

I would also like to recognize and thank the Human Resource Departments of Eli Lilly & Company and Marriott International, Inc. for the assistance they provided in allowing me to conduct pilot studies at their respective organizations.

## TABLE OF CONTENTS

	Page
ACKNOWLEDGEMENTS .....	iii
LIST OF TABLES .....	vi
Chapter	
1. INTRODUCTION .....	1
Background and History	
Theoretical Framework	
Need for the Study	
Statement of the Problem	
Purpose of the Study	
Research Hypothesis	
Delimitations	
Limitations	
Definition of Terms	
Summary	
2. LITERATURE REVIEW .....	17
Introduction	
Training Evaluation Statistics	
Training Evaluation Models	
Financial Evaluation Models	
Return on Investment Process	
Return on Investment Studies	
Employee Attitudes	
Summary	
3. METHODOLOGY .....	64
Introduction	
Research Hypothesis	
Population	
Sample	
Instrument Validity and Reliability	
Methodology	
Data Treatment	
Summary	
4. FINDINGS .....	76

Participants in the Study	
Data Assessment and Analysis	
Effect of Training on the Financial Results of the Organization	
Effect of Training on Employee Attitudes	
Summary	
5. DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS .....	95
Overview	
Summary and Discussion of the Findings	
Limitations of the Results	
Implications for Organizations	
Recommendations for Practice	
Conclusions	
BIBLIOGRAPHY .....	111

## LIST OF TABLES

Table	Page
1. List of Research Questions and Hypotheses	66
2. Job Types of Participating Sample	68
3. Training Evaluation Methodology	71
4. Statistical Treatment of Research Questions	75
5. Job Classification of Participants	77
6. Null Hypothesis #1 – Descriptive Statistics – Underwriting Group	80
7. Homogeneity of Variance – Underwriting Group	80
8. Analysis of Variance – Underwriting Group	81
9. Null Hypothesis #2 – Descriptive Statistics – Underwriting and Non-Underwriting Groups	81
10. Independent Samples Test – Underwriting and Non-Underwriting Groups	82
11. Null Hypothesis #3 – Group Statistics – Underwriting Assistants and Other Groups	83
12. Independent Samples Test – Underwriting Assistants and Other Groups	83
13. Null Hypothesis #4 – Group Statistics – Underwriter Results by Years of Experience	84
14. Independent Samples Test – Underwriter Results by Years of Experience	84
15. Pre-Test Scores and Rates of Return	85
16. Analysis of Pre-Test ROI Score Bands	86
17. Homogeneity of Variance – Pre-Test ROI Score Bands	86
18. Analysis of Variance – ROI Score Bands	87
19. Multiple Comparisons of ROI Score Bands – Tukey HSD	87

20. Multiple Comparisons of ROI Score Bands – Dunnett T3	88
21. Multiple Comparisons of ROI Score Bands – Games-Howell	89
22. Results of Return on Investment Calculations	101



## CHAPTER 1

### INTRODUCTION

Training is a part of the activities of every business organization. All organizations support the value of training. Yet in many organizations, there is no direct link between the training and the value it returns to the organization. As a result, when an organization is not meeting its profit expectations and expenses need to be slashed, the training budget is one of the first to be cut (Schneider, Monetta, & Wright, 1992 and Lachnit, 2001). When faced with a slowdown in the economy, questions arise about the return on investment from training activities (Dionne, 1996). Arguably, such times could be when many training budgets should be increased the most to make organizations more competitive. Establishing a competitive advantage is one route to financial success. Part of the budgetary problem in the use of training interventions may be this failure to link training to the ability to enhance profits (Moller, Benscoter, & Rohrer-Murphy, 2000). Top management is often inclined to pressure human resources departments to contribute more to the bottom line process while at the same time having little confidence in their ability to do so (Swanson & Gradous, 1988). The solution lies in more aggressive training evaluation methods from a financial decision-making perspective that top management can understand (Holton & Naquin, (2005).

Evaluation is defined as a systematic process by which pertinent data are collected and converted into information for measuring the effects of training, helping in decision making, documenting results to be used in program improvement and providing a method for determining the quality of training (Basarab & Root, 1992). According to Pat Kittle, manager of the Daimler Chrysler Academy, “People say that training is expensive. If you think training is expensive, try ignorance” (Sussman, 2006, p.20.). This statement reflects a time period when Chrysler was not

doing well against any of their competitors. Yet when Chrysler established a program of dealership training, their sales increased significantly.

The evaluation process should be closely linked to the process of measuring performance within the organization. First, accurate measurement and evaluation can provide an organization with data representative of their current performance situation. That data can then be used to make strategic and managerial business decisions that can direct the future of the organization. Second, measurement and evaluation serves the organization as a means of communicating current and expected levels of performance. Employees may know what their expected levels of performance are, but without a good measure of current performance, it becomes difficult to institute corrective action. Third, both measurement and evaluation encourage employee performance improvement, which in turn will lead to organizational success (Sleezer & Gradous, 1998).

Obtaining participants' satisfaction with the training they have participated in is not uncommon. Asking the participant how they felt about the training and even asking them to place a value on that learning is an easy and cost effective way of getting data. Somewhat less common is determining whether learning has taken place by measuring the amount of change in knowledge at the end of the training intervention. Going out and seeing how that training is used on the job is even less common, and determining how the organization benefits financially from its training activities is seldom seen. The latter is referred to as the calculation of an organization's return on investment from training.

### Background and History

The creation and use of tools, weapons, shelter, and language all required some form of training (Steinmetz, 1942). As man developed in complexity in the use of his artifacts, so too did

his methods of training. The concept of the apprenticeship dates back to 2100 B.C. The fruits of master and apprentice were produced one at a time and remained so for a long time. The Industrial Revolution brought with it the institution of the factory school (Miller, 1996). Classes were held within the confines of the factory. The mechanization of the Industrial Revolution meant increased productivity per worker. It also meant that goods could be produced and sold for far less than what they cost when they were made by hand. As demand increased for cheap goods so too did the need for trained workers to run the machines.

The need for mass production accelerated with the start of World War I (Nolan, 1996). Trained workers were entering the military, leaving a significant gap in the industrial segment. Younger workers and women were entering the workforce in large numbers and needed to be trained quickly. However, systematic training began with World War II. Businesses were booming and they all wanted training that was quick and cost effective. Programmed learning was introduced followed by individual instruction. Today, both concepts are employed in the application of computer-based training. As a result, distance education over the Internet is gaining widespread acceptance, even amongst the most prestigious colleges and universities. Virtual universities are being established among larger U.S. corporations. And for the individual worker, with the continuous advances being made in technology, they, and their employers, are being faced with the reality of perpetual training situations.

The need for training has always been apparent. Establishing the actual value of training has been a point of contention. In the past, training has been accepted as an expensive “have to do” and organizations conducted whatever training they could afford to do. The other view is training is a nice to do, oftentimes wasting the resources of an organization (Swanson, 1998). Traditionally, training and development managers have been unable to connect training to

business results on other than a trust basis. These companies treat training as a necessary expense rather than as an investment (Sussman, 2006). Today, technology is replacing even greater numbers of employees needed for specific tasks than were ever contemplated by those in the Industrial Revolution. Yet even as technology increases in complexity, the training parameters are becoming equally more complex, and much more expensive. Yet just spending more money on training does not necessarily create positive change unless that training is linked to business strategies (Van Brakel, 2002). With the cost of training increasing, it becomes even more critical to make training expenditure decisions based on expected financial returns to the organization.

### Theoretical Framework

The theoretical framework of this study was based on the evaluation of training activities in business. Its specific purpose was to determine how to objectively calculate the value of training, without having to rely on the participants' opinions about what that value was. Current methods of evaluation often rely on participants' self-evaluations relative to how they view their own their productivity following training. This process has resulted in huge quadruple digit returns on investment. The participant satisfaction survey is the weakest of all training evaluation intervention methods (Swanson, 1993). And there is evidence that there is little correlation between participants' reactions and learning (Dixon, 1990). Attempts are not made to isolate the effects of training to determine if it is the cause for any change in productivity. Training is often expected to be entertaining with a perception that learning is passive. This can negate even the most well-designed satisfaction surveys (Dixon, 1994). These results often create a lack of credibility with the evaluation process (Swanson & Gradous, 1988). This study sought to establish an objective method of evaluation and to demonstrate that through the use of financial analysis, it could be used to determine what training interventions should be undertaken.

This study expands on the work of Campbell (1994) in his application of participant's salary to the training evaluation process, and Swanson (2001) and Swanson & Gradous (1998) whose premises were supporting forecasting financial benefits prior to beginning any training intervention. Their premise also extended training evaluation beyond just the allocation of resources by including a motivational component. Clark, Dobbins & Ladd (1993) followed that up with correlational studies on the relationship between training motivation and job utility. Following those concepts, this study provides data on participant attitudes in an attempt to determine if there is a relationship between training and participant's attitudes regarding the organization.

#### Need for the Study

The training function is what keeps employees up to date with change. New employees need to be trained. More experienced employees often can use refresher training. And then there are changes that require further training. Acting quickly in the application of appropriate training interventions is critical in the competitive world of business today (Korth, 2001). Established formal education is a broad-based process of general education. Training, on the other hand, is instruction in very specific knowledge and skill sets, to update an employee's ability or to retrain for a different job. Training is often seen as a method to improve the results of the organization by meeting the needs of its customers.

The goal of any organization must be to satisfy its customers. Customer satisfaction is the intangible aspects of quality the customer expresses about the global characteristics which allow a product, or a service, to satisfy the customer's explicit as well as implicit needs (Schael, 1996). Ultimately, the purpose of training is to allow an organization to increase the satisfaction level of its customers. By doing so, a business maintains its customer base, secures new customers, and

stays competitive and remains in business. Resources, whether used for production or training employees to be more productive, or for any other reason, all come with a cost. The readiness and willingness to spend these scarce resources is significantly influenced by what top-level managers believe will be a sufficient return for spending those resources. And getting some kind of measurement as to the success of any training venture prior to its start would be further convincing evidence to release the resources. When consistent and quality decisions are made on when and where to spend resources, the result will be a more cost efficient organization that ends up as a more competitive organization and a more competitive organization is one of the goals of the training process.

While most senior executives and managers would acknowledge the importance of training, the measurement of training outcomes has been significantly lacking. The American Society for Training and Development (ASTD) has captured organizational data that demonstrates the need for this study. Part of that data is based on Kirkpatrick's four levels (Kirkpatrick, 1998). The four levels represent levels of evaluation. Level 1 is reaction, which means collecting participants' opinions of the training. Level 2 is learning, which means determining the change in the knowledge or skill levels of the participants. Level 3 is behavior, which means evaluating the changes in participant behavior. Level 4 is results, which means representing the results of the organization that can be attributed to training.

In the 2000 ASTD State of the Industry Report, a survey of the 501 United States organizations that participated in the 1999 benchmarking service showed that 77% of the participants in the survey conducted level 1 evaluations, but only 36% conducted level 2 evaluations, only 15% conducted level 3 evaluations, and only 8% conducted level 4 evaluations (McMurrer, Van Buren, & Woodwell, 2000). In the 2001 ASTD State of the Industry Report, a

subsequent survey of the 365 U.S. organizations that participated in the 2000 benchmarking service showed that 77% of the participants conducted level 1 evaluations, 38% conducted level 2 evaluations, 14% conducted level 3 evaluations, and only 7% conducted level 4 evaluations (Van Buren, 2001). In the 2002 ASTD State of the Industry Report, a survey of the 367 U.S. organizations that participated in the 2001 benchmarking service showed that 78% of the participants conducted level 1 evaluations, 32% conducted level 2 evaluations, 9% conducted level 3 evaluations, and only 7% conducted level 4 evaluations (Van Buren & Erskine, 2002). This pattern continued into 2003 and 2004. In the 2003 ASTD State of the Industry Report, a survey of the 276 U.S. organizations that participated in the 2002 benchmarking service showed that 75% of the participants conducted level 1 evaluations, 41% conducted level 2 evaluations, 21% conducted level 3 evaluations, and only 11% conducted level 4 evaluations (Sugrue, 2003). In the 2004 ASTD State of the Industry Report, a survey of the 213 U.S. organizations that participated in the 2003 benchmarking service showed that 74% of the participants conducted level 1 evaluations, 31% conducted level 2 evaluations, 14% conducted level 3 evaluations, and only 8% conducted level 4 evaluations (Sugrue & Kim, 2004). The ASTD did not collect evaluation data in 2005 and 2006.

The ASTD also identified two other groups for analysis. These were called the training investment leaders and the benchmarking forum. Training investment leaders were identified as those companies that had made a dedicated financial commitment to developing employee knowledge, skills, and abilities. Benchmarking forum members were identified as having an even higher financial commitment to training. Yet even the training investment leaders and benchmark forum members did not differ from the benchmarking service members in their application of the four levels of evaluation. In the 2000 Survey, 8% of the members of the

benchmarking service (40 organizations) used level 4 evaluations while only 9% of the training investment leaders (5 organizations) and 3% of the benchmarking forum members (2 organizations) used level 4 evaluations.

In the 2001 Survey, 7% of the benchmarking forum members (26 organizations) used level 4 evaluations while only 11% of the training investment leaders (4 organizations) and 6% of the members of the benchmarking forum (3 organizations) did so (Van Buren, 2001). In the 2002 Survey, 7% of the benchmarking service members (26 organizations) used level 4 evaluations while only 7% of the training investment leaders (3 organizations) and 6% of the benchmarking forum members (3 organizations) used level 4 evaluations (Van Buren & Erskine, 2002). Even with the significantly higher training expenditures that the training investment leaders and the benchmarking forum members had, their application of level 4 evaluations still played a very minor role in the evaluation process.

The ASTD reports show that most evaluation efforts relied on participant opinions. Even larger organizations that spent millions on training did not substantially differ in their evaluation methods. For one reason or another, business organizations appear to be reluctant to use methods of evaluation to manage their training resources. One might expect to see some different results for programs involving technical training as technical workers produce physical products and have available more measurable outcomes. A study of technical training programs reported 73% of the programs used level 1, 47% at level 2, 31% at level 3, and 21% at level 4 (Twitchell, Holton, & Trott, 2000). While this suggests that evaluation is applied more at the technical training level, there is still a much more than needs to be done.

While intuition alone may suggest the positive value of training, there are many reasons that are given why evaluations are not done (Moller, Benscoter, & Rohrer-Murphy, 2000). Lack



of time, lack of resources, lack of knowledge, lack of concern for quality, lack of access to the data, and lack of credibility in the process were some of the reasons evaluations were not done. And some organizations have a culture that simply resists evaluations. In a study of 191 firms, 168 of them identified organizational resistance as the reason for the lack of evaluations (Moller & Mallin, 1996). Resistance can be found at many levels. There is the management level, the process level, and the job-performer level. High-level management comments revolve around costs, little or no perceived value, and past failures. The process level comments revolve around the existing processes being adequate whether evaluations are currently done or not. Job-performer level comments include evaluations being unnecessary, wasting time, and never resulting in any real change (Riley, Davani, Chason, Findley, & Druyor, 2003). In addition, when training focused on gains at the individual employee level, those results were often an insufficient reason for justifying resource allocation (Lyau & Pucel, 1995).

Studies are needed to demonstrate the value that training brings to the organization in a language that is used by business decision makers to make business decisions. The more convincing the evidence is beyond the first three levels of the Kirkpatrick model, the more likely it is that training interventions will be successfully justified and funded. Training costs can be significant and the greater they are, the more likely they are to not be approved, whether that training is needed or not. The answer may lie in the applicability of financial analysis to the problem (Swanson, 2001). This study created a method using the financial language of business decision-making to evaluate training interventions that was not based on participants' opinions but rather on empirical data and proven financial return models to maximize the value of each training dollar. By doing so, the effort to justify training expenditures becomes an issue of dollars and sense rather than an emotional appeal for funds.

## Statement of the Problem

Organizations in business are always searching for more and more efficient ways to accomplish what needs to get done. Training is often recognized as important but not always undertaken. Many organizations see training intervention as an optional nice to do activity or as a waste of resources where the costs far outweigh the benefits (Swanson, 1995). The problem that was under consideration was to find an objective method of evaluating training interventions so that organizations could consider the spending of resources on training as an investment and follow normal business practices that would be applied to any other investment analysis under consideration. Being able to objectively consider the expenditure of resources on training as an investment takes the decision-making for training resource allocation from a subjective process to an objective one. Minimum required rates of return could then be established against which prospective training results could be measured. The result would be a more effective use of the funds allocated for training resources.

## Purpose of the Study

The purpose of this study was to develop an objective method of evaluating training that can be incorporated into the decision-making process of an organization that does not rely on just asking the participants their opinion of the value of the training. Part of the problem is traditional return on investment (ROI) computations that use participant surveys produce such high results they raise questions about the credibility of those results. As previously cited, these surveys are suspect as to their validity. One study found a significant correlation between participant enjoyment of the training and the ratings given the intervention (Dixon, 1990). This study proposes a methodology that can be used to objectively evaluate training as an investment and estimate financial returns prior to the training intervention being undertaken to determine if the

effort should be funded at all, allowing the organization to set a minimum rate of return requirement for the funding of any intervention.

### Research Hypothesis

This study is based on two research questions and five research hypotheses. The first research question asks can an objective method for evaluating training be developed? This research question is supported by four null hypotheses:

1. There will be no statistically significant difference in ROI results of an underwriter-training program consisting of three different training interventions across a sample of underwriters.
2. There will be no statistically significant difference in ROI results of an underwriter-training program between a group of underwriters and a group of non-underwriters.
3. There will be no statistically significant difference in ROI results of an underwriter-training program between a group of assistant underwriters and a group of other employees of the same grade.
4. There will be no statistically significant difference in ROI results of an underwriter-training program between underwriters with less than 5 years of experience and underwriters with 5 or more years of experience.

The second research question seeks to determine if a pre-test can be used as a predictive tool to make a decision regarding training interventions at the individual participant level based on a future estimated ROI. It builds on the foundation set by the first research question and the supporting first four null hypotheses. The second question is supported by the fifth hypothesis:

5. There will be no statistically significant difference in ROI results of an underwriter-training program within ROI score bands.

## Delimitations

This study applied to training evaluation in an insurance services organization. It was based on pre-tests and post-tests that were being applied as part of an in-house training program. These pre-tests and post-tests were used to compute the amount of learning that had taken place and then used that information as part of a seven-step formula to complete a return on investment calculation. This can be used to determine when resources should be allocated for training interventions. A minimum acceptable return on investment can be established as a benchmark for the allocation of resources. The method could also be used to forecast return on investment on any proposed intervention based on pre-test results. High pre-test scores would suggest a lower return on investment since there is less room for an increase in knowledge as demonstrated by the application of the post-test. This kind of analysis will be the subject of some of the null hypotheses.

This study was not a study of the relationship between employee attitudes and organizational profitability nor was it a study of the relationship between motivation and the profitability of the organization. Such an inquiry belongs in a separate study. There are a number of such studies on employee attitudes, self-efficacy and motivation. The subject was touched on here because of the relationship between training and motivation, as indicated on some of the referenced research studies. Neither was this a study of the most significant contributors to the bottom line profitability of an insurance company. That would be a sizeable undertaking and if it could be done successfully, insurance companies would no longer be in the business of taking risk as they could follow the results of such a study to profitability year after year. But the study does include a section that analyzes the relationship between training and the financial results of this insurance company.

## Limitations

It was critical that the pre-test be a comprehensive test of the learning objectives of the selected course of study. The results of the pre and post-tests would depend on the quality of the training the organization was conducting. Because these were actual training classes in a major corporation, there was no random selection of participants. The process of validating that learning was occurring after the intervention was done through an existing monthly file evaluation process for all participants that could be flawed as different individuals do these evaluations. The validity of the test questions themselves was dealt with by using the questions prepared by the Insurance Institute of America. The Insurance Institute has been providing a variety of insurance training materials and conducting tests for certification for over forty years. This material is considered a standard for the insurance industry. The Institute also conducts an item analysis of each test question making needed revisions as necessary to ensure test question reliability.

One of the most significant limitations of the study was the difficulty in trying to control the testing scenario to resemble an ideal experimental situation. The organization had a significant need for this training with a limited window of opportunity to complete the training. Having identical groups for each test event would have been ideal but not possible as this was a real training situation required of all of the underwriters. For example, some of them could have been out on sales calls the day of the specific intervention. As a result, the groupings are uneven and in each case the Levene's test for unequal variances was applied in the statistical analysis.

## Definition of Terms

Combined ratio – an insurance industry term that defines profitability. The combined ratio is a percentage that determines profitability using 100% as a baseline. It consists of adding

together a loss and loss adjustment ratio and an expense ratio component. The loss ratio consists of all of the incurred losses and loss adjustment expenses divided by the earned insurance premiums for the year. The expense ratio consists of all of the incurred expenses, except for loss adjustment expenses, divided by the earned insurance premiums for the year. Expenses mean all of the expenses that are paid out to run the organization on an annualized basis. A percentage below 100 points indicates profitability. A percentage above 100 indicates how poorly a company is performing. Insurance companies have assigned financial ratings and the combined ratio over time is a key component of the financial ratings that are assigned.

Evaluation – a systematic process by which pertinent data is collected and converted into information for measuring the effects of training, helping in decision-making, documenting results to be used in program improvement and providing a method for determining the quality of training.

Job relevance factor – this is a factor that represents the percentage of knowledge and skills of the job covered by the designated training intervention. The specific value represented by the factor should be determined by the organization. For example, a supervisory course may be assigned a value of 100% if the individuals taking the course are all supervisors and the material taught in the course is all the position requires for full competency. A leadership course may be assigned a value of 20% for the managers taking the course if there are other skill sets that are needed in the job that are not included in the leadership course. This variable was not found in other established methodologies. By not having such a factor, the results could be considered to be double dipping if participants go through a series of training interventions in a given year. This process allows the organization to be the only decision-maker on how relevant any training proposal is by assigning factors for each position that is involved.

Post-training assessment factor – an adjustment factor applied to the results of the post-test based on an evaluation of the participant following the training intervention, in this case it was based on written evaluation of underwriter files per quarter.

Workflow – a unit of work within a process generating products or services related to, or resulting in, customer satisfaction.

### Summary

Training has always had a major role in the business development history of the United States. From individual apprenticeships to mass production to fully automated processes, training has been instrumental in effectively making the transition of change from one process to another. However, a major component of the success of any business organization attempting to institute a new business process is its ability to deal with expenses associated with that change. Most disciplined organizations have budgetary procedures that require that expense initiatives be justified whether it is through some form of return on investment or internal rate of return calculation. Training has been an area that has lagged behind in the use of financial models of return to justify training expenditures. Training is often seen as a necessary evil that must be done, no matter what. But when organizations are not meeting their business plans, the usual response is some form of expense control resulting in budgetary restrictions. Those line items that have no direct connection to the bottom line tend to be given extra scrutiny and often end up on the chopping block, and they are often in the area of training and development.

This study has developed an objective method of calculating the value of training interventions. This method was based on the use of an analytical measure of the value of training without having to resort to using participants' opinions on how they valued their training. By using financial return models to prove the value of training, training venues can be undertaken

knowing that its justification has already been established. By having such a procedure as part of the process of determining what training to provide, business organizations can maximize the value of each training dollar and take some comfort in knowing that the anticipated rates of return are based on hard data rather than soft.



## CHAPTER 2

### LITERATURE REVIEW

#### Introduction

The first section is a review of the state of the industry relative to training evaluation. It is built around data that has been gathered by the American Society for Training and Development since 1999. The data was very revealing with regards to training evaluation methods in use today by some of the largest organizations in the U.S. These studies are very comprehensive and provided valuable information on the state of the art of training evaluation. This information was used to validate the need for this study. The second section is a discussion of the various training evaluation models that exist, including the original breakthrough four levels by Kirkpatrick, and models proposed by Phillips, who introduced the fifth level to Kirkpatrick's four levels, and Brinkerhoff, and others.

The third section is a review of the different financial analysis models and their key components. The work of Swanson, (2001), Swanson (1998), Campbell (1994) and Swanson & Gradous (1988) was used to support and analyze the propositions in this paper. Campbell's work was important in the framing of the proposed methodology. Swanson (2001), Swanson (1998) and Swanson & Gradous (1988) all did work on the positive impact of financial modeling relative to training evaluation. Their work was used to help evaluate the results of this study. The fourth section is a discussion of return on investment as a competitive advantage. It specifically deals with research based on using return on investment as a method of evaluating training. The fifth section of this study covers the results of a number of return on investment (ROI) case studies. These are documented examples that provide return on investment results as part of the evaluation process. The sixth section deals with employee attitudes. It covers some of the more

significant studies that have been completed on employee attitudes, including motivation, self-efficacy and organizational commitment and how they may relate to training.

While there are several training evaluation models in existence today, the National Human Resource Development Executive Survey on Measurement and Evaluation (1997) found that 67% of organizations that did training evaluation used the Kirkpatrick model. Other models that were mentioned included the Brinkerhoff model, the Phillips model, and those that were developed in-house. The survey also found that the greatest challenge for human resources executives was determining the impact of training interventions on the financial performance of the organization.

#### Training Evaluation Statistics

The ASTD benchmarking service was established in 1997, and was open to all any organization that sought to assess the strengths and weaknesses of its training efforts (Bassi & Van Buren, 1999). The benchmarking service was a major advance in establishing a process to collect data from major organizations about their training and development activities and spending patterns. The service provided useful data on the state of training within the business community on an annualized basis. The application of evaluation techniques using any kind of analysis on the effects of training on the financial results of the organization was still a minority measure in the process of evaluation. In the 2000 ASTD State of the Industry Report (McMurrer, Van Buren, & Woodwell, 2000), a survey of the 501 U.S. organizations that participated in the 1999 ASTD benchmarking service showed that Kirkpatrick's level 4 evaluations did not play much of a role in the training evaluation process. The survey indicated that 77% of the participants conducted level 1 evaluations, consisting of asking for trainees' reaction to the training. At the other levels, 36% of the participants conducted level 2 evaluations, measuring

actual learning that took place at the end of the training, and 15% of the participants conducted level 3 evaluations, going to the job site to measure behavior change. Only 8% of the participants conducted level 4 evaluations, measuring the impact of the training to the organization itself.

In the 2001 State of the Industry Report (Van Buren, 2001), a survey of the 365 U.S. organizations that participated in the 2000 benchmarking service showed that level 4 evaluations still does not play a major role in the evaluation process. The survey indicated that 77% of the members conducted Kirkpatrick's four level evaluations. The results at level 1, 38% at level 2, 14% at level 3, and 7% at level 4. The ASTD also established two other groups called training investment leaders and the benchmarking forum. Training investment leaders were identified as those organizations that had made a dedicated financial commitment to developing employee knowledge, skills, and abilities. The ASTD identified 37 training investment leaders. Benchmarking forum members were identified as having an even higher commitment to training. The ASTD identified 42 benchmarking forum members. On average, in 1999, benchmarking service members spent \$1.2 million on training. On average, training investment leaders spent \$33.0 million, and benchmarking forum members spent on average \$75.7 million (Van Buren, 2001).

In the 2002 State of the Industry Report (Van Buren, & Erskine, 2002), a survey of 367 U.S. organizations that participated in the 2001 benchmarking service showed a slight drop in the use of level 4 evaluations compared to the prior year. This may be due to who was sampled since the groups are not the same year to year. The survey indicated that 78% of the members conducted evaluations at level 1, 32% at level 2, 9% at level 3, and 7% at level 4. The ASTD identified 39 training investment leaders and 45 benchmarking forum members. On average, in 2000, benchmarking service members spent \$2.0 million on training. On average, training

investment leaders spent \$31.2 million, and benchmarking forum members spent on average \$63.4 million (Van Buren & Erskine, 2002).

The organizations described as training investment leaders and benchmarking forum members did not differ from the benchmarking service members in their application of level 4 evaluations. In the 2000 survey, 8% of the benchmarking service members (40 organizations) used level 4 evaluations while only 9% of the training investment leaders (5 organizations) and 3% of the benchmarking forum members (2 organizations) used level 4 evaluations (McMurrer, Van Buren, & Woodwell, 2000). In the 2001 survey, 7% of the benchmarking service members (26 organizations) used level 4 evaluations while only 11% of the training investment leaders (4 organizations) and 6% of the benchmarking forum members (3 organizations) did so (Van Buren, 2001). In the 2002 survey, 7% of the benchmarking service members (26 organizations) used level 4 evaluations, only 7% of the training investment leaders (3 organizations) and 6% of the benchmarking forum members (3 organizations) used level 4 evaluations (Van Buren & Erskine, 2002). Even with the significantly higher training expenditures that the latter two groups had, that money was not spent on level 4 evaluations. Their application of level 4 evaluations still played a very minor role in the overall evaluation process. Even today, level 4 evaluations are not an integral part of the training evaluation process, further confirming the need for studies such as these. Level 4 evaluations, using techniques such as ROI, demonstrate the value of training investment analysis to the organization. Yet even the training investment leaders and benchmarking forum members, with their significant expenditures on training, still confined their evaluations for the most part to level 1 questionnaires.

In the 2003 State of the Industry Report (Sugrue, 2003), the 276 benchmarking service participants indicated 75% participated in level 1 training, 41% participated in level 2 training,

21% in level 3 training, and 11% in level 4 training. The 28 training investment leaders indicated 72% participated in level 1 training, 59% participated in level 2 training, 33% participated in level 3 training, and 14% in level 4 training. The 21 benchmarking forum members indicated 84% participated in level 1 training, 38% participated in level 2 training, 25% participated in level 3 training, 9% participated in level 4 training and 10% participated in level 5 training. These results were similar to prior year reports.

In the 2004 State of the Industry Report (Sugrue & Kim, 2004), the 213 benchmarking service participants indicated 74% participated in level 1 training, 31% participated in level 2 training, 14% participated in level 3 training, and 8% participated in level 4 training. In the 2005 State of the Industry Report (Sugrue & Rivera, 2005), evaluation methods data was not collected from benchmarking service organizations in 2005. However, it was collected from the 18 benchmarking forum members. Of that group, 91% participated in level 1 training, 54% in level 2, 23% in level 3, 8% in level 4, and 2% in level 5 training.

One might expect to see different results for programs involving technical training. Technical workers produce physical products and as such have available more measurable outcomes. It is thus easier to measure accurately the outcomes of training. One study of technical training programs reported 73% of the programs used level 1, 47% level 2, 31% level 3, and 21% level 4 (Twitchell, Holton, & Trott, 2000). While the results would indicate that evaluation is applied more at the technical training level than in general training, there is still a much more than can be done to apply level 4 evaluations.

There are many reasons for conducting level 4 training evaluations. Return on investment calculations can be used to justify future expenditures. These calculations can also be used to measure training department efficiency in controlling costs. And prospective calculations can be

used to make decisions between training alternatives. However, there are equally as many reasons that are given for not doing evaluations why evaluations are not done. Lack of time, lack of resources, lack of knowledge, lack of concern for quality, and lack of access to the data were some of the reasons evaluations were not done (Moller, Benscoter, & Rohrer-Murphy, 2000). Some organizations have a culture that simply resists evaluations. In a recent study of 191 organizations, 168 identified organizational resistance as the reason for the lack of evaluations (Moller & Mallin, 1996). Clearly, training evaluation is a subject of significant research. The ASTD does an annual survey of major organizations, some with very large training budgets, to determine what is being done in the business world. Yet despite the large amounts being spent, there is little to show in terms of significant increases in the application of the 2<sup>nd</sup>, 3<sup>rd</sup> and 4<sup>th</sup> levels of Kirkpatrick, much less the 5<sup>th</sup> level of Phillips.

#### Training Evaluation Models

Training can be divided into two groups (Lewis, 1996). One type of training is proactive. Proactive training does not necessarily relate to the current needs of the organization. It is visionary and is concerned with the lifelong learning needs of employees. Oftentimes, this type of training is harder to justify in any organization where such justification is required. The other type of training is reactive. Reactive training is based on the immediate identified needs of the organization. The training is needed either because of performance deficiencies or lack of knowledge. This may be an easier type of training to justify because the issues are so visible. Training also provides two types of outcomes for the organization. One type is tangible and the other is intangible. Tangible outcomes are observable and measurable. Examples include increased productivity a reduction in errors and a decline in absenteeism. Intangible outcomes are difficult to measure. Examples include teamwork, leadership skills and a commitment to the

organization. There are numerous evaluation models that attempt to measure these kinds of outcomes.

The training process has a number of goals (Brinkerhoff, 1987). The first goal is to provide employees with the skills and knowledge needed to perform their current jobs. The second is to provide employees with what they need to achieve promotion and attain other career advancement objectives. The third is employee development, providing what they may need for future roles in the organization. The fourth is to provide employees with the knowledge needed to understand the organization, the business, its culture, and its rules and policies. The fifth is to help employees increase their ability to handle stress and cope with change. The sixth goal is leadership development among the employees of the organization. The final goal is to help employees to learn and grow for personal development.

There are a number of ways to evaluate training. There are ten guidelines that aid the evaluation process (Parry, 1997). First, evaluation must start before you train. It is too late after the fact. Second, evaluate the entering behavior of the trainee, the needs and expectations of the organization and establish the terminal behavior you expect of trainees after training. Third, evaluation must be an integral part of the instructional process. Fourth, conduct formal evaluation by someone other than the trainer. This keeps everything at a level of impartiality needed to get the most effective results. Fifth, an up-front performance contract always makes the evaluation easier. Sixth, delayed evaluation is usually better. Seventh, the higher up on the organizational ladder we train, the harder it is to evaluate the results. Eighth, we all need to remember that we don't know what we don't know. Ninth, respondents will often say what they think you want to hear. This is a significant problem with smile sheets. And tenth, evaluation should yield cost and benefit data.

Shandler (1996) describes five steps for measuring human resource development performance. The first step is to define and document the business issues that need to be addressed. The second step is to conduct a situational analysis to establish a benchmark of current performance. The third step is to initiate the training intervention. The fourth step is to re-measure the same performance benchmarks that were identified in the second step after the training intervention has been effectively completed. The fifth and final step is to measure and document the effect of the training intervention to the bottom line of the organization. This latter step involves calculating ROI. Shandler then provides a five-step process for calculating a return on investment. First, identify the desired outcomes of the training. Second, attach a value based on increased revenues or decreased costs to these outcomes. Third, measure the change in the outcomes following the training. Fourth, multiply the value of the outcome by the change that occurred following the training. This gives you the total return from training. Fifth, calculate the total costs of the training. Then subtract the cost of training from the dollar benefits of training calculated in the previous steps. Then divide the benefits by the total costs. These theories were used as a benchmark to compare the results of this study.

Fitz-enz (2000) provides a four-step process for calculating the impact of ROI in training. The first step is to conduct a situational analysis. This consists of identifying the problem in terms of service, quality or production, or SQP. Then quantitatively identify the current SQP problem level. Then determine the effect of the current SQP performance level on the organization's competitive advantage. Finally, identify the work processes that are the critical components to this situation. The second step following the situational analysis is the actual intervention. The choice of intervention is based on identifying the source of the problem, determining the best solution to the problem, agreeing on a specific solution, planning



implementation and then conducting the intervention. The third step is to determine the impact of the intervention. Determine if performance has changed. If so, determine if the direction was positive or negative. Determine how much change occurred. This should again be done in a quantitative format. Ascertain if the change was due to the intervention or to other extraneous factors. The fourth and final step is to determine the value of the change that was brought about by the intervention. A process should be created to measure the effects of the change on service, quality and production. Then determine if the change enhanced the organization's competitive advantage.

Performance value analysis (PVA) is based on the principle that if work contributes to the performance goals of the organization, then it can be valued in dollars and cents (Swanson & Gradous, 1988). Determining the value of work is based on securing four pieces of information. The first determines the unit of work performance. The second determines performance levels, both for the present level of performance and for the future desired level of performance. The third determines the value of each unit of work in dollars and cents. The last piece determines the actual performance value, which is computed by multiplying the value of a single unit of work times the total number of work units produced that can be attributed to the training intervention. Tied to performance value analysis is the process of forecasting financial benefits. The forecasting financial benefits model allows the user to make accurate investment decisions based on a forecast of the financial value of the improved performance. It is one of just many financial analysis methods reported on by Mosier (1990).

There are thirteen steps in calculating the value of performance in performance value analysis. The first step is to select the unit of work performance that will be measured. The second step is to determine the performance goal per worker at the end of the training

intervention. The third step is to determine the actual performance of the worker prior to the training intervention. The fourth step is to assign a dollar value to each performance unit. The fifth step is to determine the amount of time required to reach the level of desired performance. The sixth step is to determine the length of the period of evaluation. The seventh step is to determine how many employees will participate in the training intervention. The remainder of the steps involves the actual performance value calculations. The eighth step is to determine if employees will produce units of work during the training intervention. If the answer is “yes,” the method of computing the average rate is work performed prior to the intervention added to work performed after the intervention, the sum divided by two. The ninth step is to determine the total units of work that will be produced during the training intervention. The tenth step is to determine how many units of work will be produced per worker during the period of evaluation. The eleventh step is to determine what will be the value of worker’s performance during the period of evaluation. The twelfth step is to determine the value of the performance gain for each worker. And the final step involves calculating the total performance gain.

All organizations, when faced with the challenge of scarce resources, will eventually have to make critical decisions about how those resources are to be allocated. Any organization that is to survive must evaluate all of its business components on a return-on-investment type of framework. It will be forced to do so, with or without valid data (Swanson, 1998). This situation would also include training as one of those components that will require an ROI decision. It would be much more effective if that decision on training ROI was based on an objective method of evaluation.

Dixon (1990) proposes three cost models, one of which is return on investment. The cost comparison model compares the cost of one method to the cost of another. Examples include the

cost of on-site training compared to distance education, the cost of an event a year ago to the same event today, and the cost of developing a program in-house compared to purchasing it from a vendor. Also considered is the cost of a training event compared to the cost of doing nothing based on the anticipated benefits to be derived. And this latter point is often ignored as a viable alternative. The causal model is used to ascertain what behaviors actually will produce the desired result and only those behaviors should be addressed by training. This method can be very subjective in terms of exactly what behaviors really will produce the desired results and can the results be measured with accuracy. The testing method is pre-test post-test control group design. The causal model is useful for business goals, such as absenteeism and morale, which are hard to place a dollar value on. The final model was return on investment. ROI is particularly beneficial when the unit of measure for both the cost and the benefit are the same.

Cost consequences analysis is a process that organizations use to approximate the return on investment for any selected project (Muir, Watkins, Kaufman, & Leigh, 1998). Cost consequences analysis involves the application of cost-utility analysis, cost-benefit analysis, and cost-effectiveness analysis. Cost consequences analysis is an estimate and therefore does not require the time and rigor of an ROI analysis. The process is based on estimating the cost of current interventions, estimating costs to reach the desired level of performance, and the anticipated results of such an intervention. An estimate of the results of the intervention is based on evaluating prospective changes in revenue, expenditures, and employee productivity. This method provides an estimate of what goes into an intervention and what is expected to come out of it. It is used to make decisions when there is neither the time nor the necessity or resources for a true ROI analysis (Kaufman, & Watkins, 1996). Organizations often have alternative proposals that seek to accomplish the same thing. Cost consequences analysis provides a method of

evaluating each alternative to determine which interventions are estimated to produce the best return for the investment.

Utility analysis is the most detailed of the analytical methods available to evaluate training programs (Cascio, 1989). It expresses the outcomes of an intervention in dollar terms. Utility refers to expected dollar returns. Utility analysis can be used to evaluate different competing programs to determine which ones provide the largest dollar return to the organization. Capital budgeting methods are used to determine the minimum annual benefits required as the return on investment of any prospective program. Break-even analysis estimates the minimum effect size needed for a program to produce the necessary results. Benefits from any intervention must be stated in terms of measurable changes in the cash flows of the organization. The information needed includes the cost of the program, the additional benefits derived from instituting the training intervention, the length of time of the benefits, and the discount rate that represents the organization's minimum expected rate of return on the intervention. Discounting represents the present value of future cash flows, recognizing that a dollar received in the future is worth less than a dollar received today. Net present value is equal to the dollar benefit of training times the present value factor for an annuity at a selected discount rate, minus the cost of the training intervention. When the equation is solved for the dollar benefit level, that amount is the minimum additional amount that must be generated from the training intervention for the intervention to have a positive net rate of return. Break-even analysis is then applied to determine what would be the minimum additional change in performance required from each participant to reach the necessary dollar benefit level change.

Another training evaluation model is based on the value of human capital (Stolovitch & Maurice, 1998). This model consists of seven major steps. Step 1 involves the calculation of the

potential for improved performance. A front-end analysis is conducted that determines the purpose of the training intervention, identifies actual and desired performance, identifies feelings associated with the desired performance, identifies the causes for not achieving the desired performance, and identifies solutions that are economically feasible and acceptable to both the organization and the prospective trainees. This front-end analysis identifies the magnitude, urgency, and economic value of the gap in performance. Potential for improved performance (PIP) is calculated by dividing exemplary performer production ( $W_{ex}$ ) by typical performer production ( $W_t$ ).

The calculation of training costs is found in step 2. Training costs are broken out into development costs, implementation costs, and maintenance costs. Step 3 involves calculation of worth analysis. Worth analysis verifies that the worth of the intervention exceeds the anticipated costs of the training. An estimate is made of highest value of improvements, given the training intervention. Another is made of the lowest value of improvements. These values are multiplied by the expected life of the effect of the training. Each number is divided by the estimated cost of the training. The benefit to the organization is determined by subtracting the low worth value from the high worth value. This net value is the benefit cost ratio to the organization. Step 4 is training. This includes design, development, implementation, and evaluation.

Step 5 involves the calculation of the true cost of the training intervention. Step 2 provided an estimate of training costs. Step 5 represents the actual training costs. Step 6 involves calculating the organizational return on investment. This step is divided into two sub-steps. Separate calculations are made for tangible monetary improvements and intangible non-monetary improvements. Monetary performance indicators can include average size of sale, increased volume, increased output, reduced absenteeism, repeat business and transactions per

day. Non-monetary performance indicators can include call to close ratio, reduced number of grievances, and number of referrals. Step 7 involves the calculation of individual increased value of human capital. This step differs from most training evaluation methods. Each employee is assigned a human capital account that initially consists of his or her base salary. As the employee's skills increase through training, the value of the account increases also. If comparing an employee's pre-training performance to post-training performance is valued at \$10,000 in step 6, then the value of that individual's human capital account increases by \$10,000. When the employee's salary and the value of his or her human capital account begins to significantly differ, the result may be the employee leaving the organization for more money. This last step produces a separate benefit outside of the traditional ROI evaluation in that it keeps a running total of the value of each employee to the organization based on the training received. This process differs quite a bit from more traditional methods of evaluating training interventions.

Probably the most widely known method of training evaluation is the four levels (Kirkpatrick, 1998). Kirkpatrick may be the most quoted individual in the literature when it comes to training evaluation. The four levels forms the basis for a number of other methods that have been proposed and are mentioned in this survey. Kirkpatrick proposed a sequence of four steps in the training evaluation process. Each level was a prelude to the next level, meaning that none of the levels should be skipped just to get to the next level. Successful training evaluation involved the completion of all four levels. They consisted of reaction, learning, behavior and results.

Level 1 – reaction is an indication of customer satisfaction. It involves finding out how the participants felt about the training session. A distinction should be made here between reactions to out-sourced training sessions and in-house ones. With out-sourced training, an

organization has chosen a particular topic and is paying fees to have employees participate. If the evaluation results are poor, they are not likely to use that out-source again and the fee company may go out of business. With in-house training, the measure of customer satisfaction takes on a different meaning. Employees usually have no choice in their attendance. And reports of their progress may get back to their supervisors, a kind of motivation that usually does not exist in out-sourced training events.

Level 2 – learning is the extent to which participants change as a result of the training. That change can be a change in attitude, such as diversity training in the workplace. It can be an increase in knowledge, such as legal seminars on employment discrimination practices, or an increase in skills, such as training in auto repair. Other subjects, such as leadership, motivation, or communication training can cover all three areas. The goal of learning is to effect a change in the participant's behavior. The goal of this level is to improve skills, increase knowledge, or change attitudes.

Level 3 – behavior is defined as a change in behavior that can be attributed to the results of a training program. A level 3 evaluation cannot be completed unless an evaluation has been made at levels 1 and 2. Kirkpatrick identifies the possibility that reactions may be favorable and learning objectives may have been met, but a change in behavior is not observed. For a change in behavior to occur, there must be a desire to change, an awareness of how to apply the new knowledge, a suitable work environment, and some method of rewarding the change in behavior.

Most training programs complete level 1 and 2 evaluations. Level 3 is the first of the four levels that can present difficulties based on conditions outside of the training environment. Training can have a positive effect on a participant's desire to change and the skills to apply this new found knowledge. The climate of the working environment however is something that only

the participant's supervisor can provide. There are five different types of climate that can exist: preventing, discouraging, neutral, encouraging, and requiring. Ideally, an encouraging supervisor who discusses the purpose of the training beforehand, discusses the results following completion, and encourages the participant to make use of what was learned on the job, provides the most ideal ground for a change in behavior to occur.

In addition to a suitable climate within the participant's working environment, some form of reward system is also helpful to encourage a change in behavior. Rewards can be intrinsic or extrinsic. Intrinsic rewards include job satisfaction, personal pride, and a sense of achievement. Extrinsic rewards include recognition of the individual's accomplishments through praise, salary increases or bonuses. The degree to which training will translate into changes in job behavior will depend on a combination of the climate of the job environment and the organization's reward system.

Level 4 – results can be defined as the final results that occurred to the organization that can be attributed to the training. Such results can be based on increased productivity, improved quality, reduced costs, and a higher profit margin. However, not all training can be measured in these terms. While assembly line skills training can be measured in terms of any increased productivity, training programs in leadership, motivation, communication and decision-making are often difficult to evaluate in terms of tangible changes to productivity, quality, and cost. Such skills are often described as soft skills. Level 4 evaluations are the ones least frequently pursued when evaluating training programs. Yet these evaluations are the ones that can be the most beneficial in terms of justifying the expenses incurred or about to be incurred.

Kirkpatrick recommends five guidelines to level 4 evaluations. The first is the use of a control group whenever possible. A randomly selected control group, if all extraneous variables



are suitably accounted for, would indicate that any resulting change in the group that received training would be due solely to the training that was provided. The second is to allow time to pass for results to be achieved from training before conducting a level 4 evaluation. How much time should elapse depends on the type of training that was conducted. The third is to measure results both before and after the training. This step involves a lot more than just a pretest and a posttest. It involves measuring the prospective participant's contribution to the organization and then measuring that same contribution after the training has occurred. Like the use of a control group, this step may not always be possible. The fourth is to repeat the measurement more than once at various times following the training. This would provide a better indication as to the results of the training. The last guideline is to consider a cost benefits analysis. Measuring the costs of training can often be accomplished with little subjective judgment. However, measuring the benefits of training often require a significant amount of subjective judgment. Calculating level 4 value can only properly be evaluated by collecting data from the previous three levels (Long, 1999).

Yet even the four levels have come under some criticism. It does not go far enough to be a model by not including the role of financial forecasting. Financial forecasting can be used to evaluate an intervention before that intervention is carried out (Holton, 1996). This can be critical to any decision-making process that involves choosing between many alternatives. The argument that organizations should not engage in interventions unless the payoff justifies the expenditure is something that can make sense to top management as the decision-making process is phrased in a language they can understand.

Another model that bears some similarities to the Kirkpatrick model is the five-level ROI framework model (Phillips, 1997). The first four levels are similar to the four levels of

Kirkpatrick. The first level is reaction and planned action. At this level, the participant's satisfaction with the training intervention is measured. The second level is learning. At this level, measurements are taken to determine what the participants have learned. Tools for measurement at this level include tests, skill practices, role-playing, simulations, and group evaluations. The purpose of this level is to measure changes in knowledge, skills or attitudes. The third level is job applications. At this level, measurements are taken to determine behavior change on the job and the specific application of the training material to the job tasks at hand. Measurements are usually applied to the frequency with which the former participants are using their newly acquired knowledge on the job. The fourth level is business results, the impact the changes in participants' behavior has had on the business results of the organization. Measurements are taken of items such as production, quality, costs, time, and customer satisfaction. The fifth level is ROI. At this level, measurements are taken of the financial value of the results and divided by the total costs of the training to come up with a percentage that represents the benefit to cost ratio. Very few conduct ROI evaluations.

Another evaluation method is the five levels for evaluation of interventions for human performance improvement model (Kaufman, 1994). The first level is split into two parts. Level 1a is enabling. At this level, a determination must be made to ensure that enough resources are available to undertake the intervention. Level 1b is reaction. This step measures the participants' reaction to the training. This is identical to Kirkpatrick's level 1. Level 2 is acquisition. This step measures the knowledge that was gained from the training intervention. Level 3 is application. This step measures how much of what was learned is being applied on the job. Level 4 is organizational output. This step measures the contributions and payoffs the organization received as a result of the training intervention. Level 5 is called societal outcomes. This step seeks to

measure how the training intervention has improved society and the environment surrounding the organization. This level was not addressed in the model proposed by Kirkpatrick nor any of the other models.

Another approach is the training systems life cycle model (Kearsley & Compton, 1981). True returns can only be computed by taking into account the entire life cycle of the training intervention. This includes research and development, the start-up period, the operational period, and the transition period into an updated intervention. Costs are high during the research and development and start-up periods. Costs are for the most part constant over the operational period. Within the operational period is a steady state period, usually the one-year organizational budgetary period. During the transition period, costs go down for the present intervention and then start to rise for the replacement intervention, beginning with research and development and so forth as another life cycle begins again. Replacement interventions can be due to updating the curriculum or changing the method of instructional delivery or both.

Core competency models clarify knowledge, skills and behaviors that are critical to employees achieving the organization's goals and objectives (Montier, Alai, & Kramer, 2006). Core competencies have become a more recent process used within human resource departments. It addresses the type of competencies that are needed to get the job done in different job grades. Competencies include items such as demonstrable creativity, continuous learning, effective communication, high standards, use of information technology, partnering, knowledge of the organization, driving change, customer focused, global business knowledge, strategic thinking, authority delegation, coaching skills, leadership. These competencies often fall under what are described as soft skills. The process involves analysis of all levels of the organization and the incorporation of best practices. Evaluation of individual performance is based on an individual

behavioral assessment within the assigned core competency arena. Measurement is based on the individual's strengths and gaps compared to the requirements of the job.

The critical outcome technique (COT) seeks to assess critical organization outcomes to allow decision-makers to evaluate a variety of human resource interventions (Mattson, 2005). A critical outcome evaluates business results at the organizational level and measures the financial benefits in monetary terms. It consists of five steps. First is the post-hoc determination of the initial intended outcome. The second step is to collect data from participants to determine if they attained the intended outcome identified in the first step. Third is to validate the participant information from sources other than the participants. Fourth is to determine the performance value of the outcome. The fifth step is to produce a summary evaluation report of the findings for key decision-makers.

The last training evaluation model to be considered is called the six stages of effective HRD evaluation (Brinkerhoff, 1987). This model is a blend of formative and summative evaluation methods. Stage I is called goal setting and needs analysis. The primary task at this stage is to determine what changes in skills, knowledge and attitudes would lead to job behavior changes that would most benefit the organization. Stage II is called program design. The primary task at this stage is to determine what inputs are needed to carry out the intervention and what planned activities can best accomplish the goal of the desired change in behavior. Stage III is called implementation and operation of HRD. This represents the actual training intervention. Stage IV is called immediate outcomes. The primary task at this stage is to obtain reactions from the participants regarding the intervention and to determine what level of new knowledge has been acquired. Stage V is called endurance and application of immediate outcomes. The primary task at this stage is to determine if the participants' behavior at work has changed due to the

intervention. Stage VI is called organizational benefits. The primary task at this stage is to determine if the organization is receiving sufficient benefits from the intervention and to evaluate those benefits against the cost of the intervention.

While there are many good reasons to conduct sound training evaluation, most of them consist of students filling out a questionnaire at the end of the course. However, such reactions can have very limited value (Lott, 1975). First, results may have been influenced by the popularity of the instructor. That can interfere with the participant's ability to objectively evaluate the instructor's performance. Second, the participants will rarely have a background in training and not have a standard frame of reference from which to evaluate the course. For training evaluation to have a benefit to the organization, there must be objective measurement of learning having taken place. Then that learning must be valued in terms of the benefits it has brought to the organization. One study found a significant correlation between participants enjoyment of the training and how the instructor was evaluated (Dixon, 1990). The results may also be clouded if participants expect the training intervention to have entertainment value (Dixon, 1994).

#### Financial Evaluation Models

Evaluations can be completed on training endeavors to determine the financial benefits to an organization. The use of financial analytically tools and models offer a number of benefits. First, they raise the issue of the specific costs of individual training programs. Second, they require the determination of financial benefit to the organization. And third, the process requires in depth contact with stakeholders. Evaluations can serve other useful purposes. They can be used to make decisions about maintaining current training materials, determining which of the existing programs are the most effective and which ones need to be improved (Rohrer-Murphy,

Moller, & Bencoter, 1997). Regardless of how they are justified, the evaluation process improves the effectiveness of training programs, improving the bottom line of the organization. However, organizations often avoid level 4 evaluations because they are time consuming and costly (Shelton & Alliger, 1993). And when organizations do evaluate the impact of training they often do so indirectly by relying on participants' opinions. And as previously mentioned, participants' opinions may not be relevant to what was learned (Dixon, 1990). What is needed are level 4 evaluations if training interventions are to be linked to specific business outcomes.

There are a number of financial analytical methods available. They include payback, average rate of return, present value, internal rate of return, cost benefit ratio. Each method can be used to calculate a return on the organization's training investment (Mosier, 1990). Payback is the time period in which the amount invested in training is recovered by financial returns. It is not affected by questionable long-range projections and the size of the endeavor does not affect the calculation. Payback period is used to evaluate capital budget expenditures. The amount of savings is measured against the original cash outlay in terms of years and months. Savings of \$12,000 with a cash outlay of \$1,200 would have a payback period of one month. Savings of \$12,000 a year with a cash outlay of \$24,000 would have a payback period of two years. However, it disregards the time value of money and the effect of returns that still come in after the payback period has ended. Average rate of return determines which training alternatives will produce a higher rate of return to the organization. Usually decisions are made to accept or reject an intervention based on a minimum required average rate of return. However, averaging does not adequately value early financial returns and the ability to reinvest those returns. Present valuing discounts all future cash expenditures to their present value using a common interest rate. Profit or loss can be determined based on the cost of capital compared to anticipated

financial gains all based on present values. Discounted cash flow makes a net present value assumption based on selected interest rates and that future dollars are not worth as much as those dollars today and valuing what those future dollars are worth today.

Internal rate of return calculates the interest rate needed to make the present value of the required cash expenditure back to zero. It determines the maximum interest rate that could be paid for the intervention to break even. Like payback, it is not affected by the size of the intervention. An internal rate of return method determines what interest rate is required to make the present value of the cash flow outlay equal to zero. This is not a recommended method for evaluating training since all projects will not have the same internal rate of return. And oftentimes, the assigned internal rate of return may be very subjective. Cost-benefit ratio is the ratio of benefits to costs. A number greater than one indicates a positive return. This method allows for comparisons to be made among alternatives. Phillips (1997) recognizes a number of cost benefit methods that fall into the ROI category. There is the traditional ROI formula of net program benefits divided by program costs. There is a benefit cost ratio based on program benefits divided by program costs.

Bottom line evaluation is another method of measuring the cost effectiveness of training. In addition to the other common methods of return on investment, cost-benefit ratio and payback period (Campbell, 1994), bottom line evaluation looks at the value training provides to the organization's financial bottom line. Each participant's productivity is measured as is the total value added for the entire program. This value is then compared to the total value of the training. This process uses a questionnaire for participants based on collecting Kirkpatrick's level 1 and level 4 data at the same time. The formula is based on percent of job time spent on performing the new task times the change in productivity as a result of training (level 1) time the total annual

compensation package. Campbell's work is the benchmark against which much of the proposed methodology will be measured against. His work on both the costs and benefits portions of the evaluation process will be used as a reality check after this methodology is tested.

All of these methods require specific information about what measures will be used to do the evaluation. These measures include employee productivity, costs, and sales. A primary ingredient of any level 4 evaluation is the cost of the training. These costs include personnel, facilities, equipment, course materials, and travel. The second primary ingredient is the benefit the organization derived from the training. This step usually requires a pre and post measure of the standard being used to determine the benefit. A cost benefit analysis is simply a ratio between total costs and total benefits. Return on investment is calculated by subtracting the training costs from the benefits derived from training, then dividing the result by the training costs.

The costs portion of the cost benefit equation can be divided into two groups (Dixon, 1990). Each-time costs are associated with the specific training activity and such costs are incurred each time training is held. One-time costs are associated with the design, development, and revision of the materials for the training activity. Each-time costs include participant salaries, participant travel expenses, participant lost opportunity costs, instructor salary, instructor travel, facilities, equipment, materials, and keeping records and other administrative costs. One-time costs include needs analysis, subject matter expertise, development of materials, development of evaluation tools, pilot and revisions costs, graphic artists' time, clerical and other administrative costs. Participant lost opportunity costs represent the value of what the participants would be contributing to the organization if they were not participating in the training. This incorporates a dollar value of lost productivity as a result of being away from the job to participate in the



training intervention. This step is often missing from the studies of other researchers. And for participants not directly involved in quantifiable productivity efforts, such as middle managers, the valuation of lost productivity is very subjective.

Training costs can be divided into two major categories. The first category is cost benefit, the analysis of the cost of training in dollars compared to the benefits of training in non-dollar terms. The latter includes changes in attitudes. The second category is cost effectiveness, the analysis of the cost of training in dollars compared to the benefits of training in dollar terms (Cullen, Sawzin, Sisson, & Swanson, 1978). Calculating costs must be both thorough and exact. A major source of expense is course design and development. It should include costs for material as well as direct labor. Hours spent on development should be electronically tracked and charged back to the departments benefiting from the training intervention. Material costs come from audiovisual aids and production costs. Audiovisual expenses include the purchasing or renting of equipment, the purchasing of tapes or cassettes. Production costs include word processing, copying, typesetting, and printing.

Development costs include developer's time, the conducting of pilot tests, maintenance and revision costs, and the use of subject matter experts. There is also a cost for inventory maintenance. The largest single training expense is for the participants taking the course. The cost of instruction includes the instructors' time and the lost opportunity cost of the participants not being able to work. Costs include airfare, lodging, salary, meals, local transportation, and telephone charges. In addition to the above, there are fixed plant costs such as classroom use, equipment, utilities, janitorial services, and the like. Other miscellaneous costs include coordinating and scheduling registration, oversight expenses, and recordkeeping costs (Spencer, 1984). Other methods include breaking costs out into direct costs, such as travel and personnel,

and indirect costs, such as facilities and administrative costs, and targeted costs that are specific to the training population and the subject matter being taught (Carnevale & Schulz, 1990).

#### Return on Investment Process

Assessing training's impact to the financial results of the organization and making future training decisions in recognition of those impacts is a sustainable competitive advantage. Part of the reason for it being a sustainable competitive advantage is that so many organizations do not conduct training investment analysis. A recent American Society for Training and Development report (1999) provides a definite link between workforce training and future financial performance. The study involved the outcomes of 575 publicly traded U.S. companies during 1996 – 1998. The study found that companies who invested \$680 more in training per employee than the average company in the study improved their total stockholder return the following year by six percentage points, after controlling for other variables. When dividing the group in half, based on training expenditures, the study found that the top half had an average stockholder return of 36.9% while the bottom half had an average return of only 19.8% while the Standard and Poor's 500 had an annual return of 25.5% during the same time period. Similar patterns were found when using criteria such as gross profit margin, income per employee, and price to book ratios. The use of training as a competitive advantage and the use of evaluation of results to the bottom line, is clearly a winning strategy.

It is important to remember that while training has the potential to produce significant benefits to any organization, it is not the answer to all of the problems of the organization. It is not the answer to all the problems of every organization. Training can make a difference in three particular instances (Rossett, 1989). First it is required either by law or regulation. Second, either a new system is being introduced to employees or the existing system is being introduced to new

employees. Third, a shortcoming in performance has been identified. In the first two instances, training is necessary regardless of the return on investment of the training program. It can be calculated but does not necessarily have any influence in the decision-making process. While the third instance provides more potential for significant return on investment, it is also important to remember that a gap between actual performance and desired performance may not be due to lack of skills or knowledge. In many cases, these gaps may be due to other than skill and knowledge deficiencies (Stolovitch & Maurice, 1998). They can be due to inadequate information such as lack of clear expectations, poor feedback, or incomplete documentation. They can be the result of insufficient tools and resources, such as staff and equipment. They can be caused by inadequate or counterproductive incentives. These can range anywhere from inadequate bonuses to fear of punishment. And there can be task interference where there are obstacles in the path of employees being able to do their jobs. Where the issue is a gap in skills or knowledge, a properly constructed training intervention can provide the solution and a significant return on the training investment.

The key to the calculation of return on investment rests in the degree of transfer of training that occurs. Transfer of training is the amount of what was learned that is taken back to the workplace and applied. If there is little or no application, there can be no return on investment since the material learned is not being applied to the job. This also means that if there was little to be gained in the first place, that information could be used to prioritize the use of resources. That is why increase in knowledge by itself does not necessarily mean anything. And there can be a significant amount of wasted expenditures. As was already mentioned, the most obvious waste arises from applying training when the problem did not lie in a lack of training. Other wasted expenditures include poor selection of employees to attend training, employees

enrolled in training with no clear expectations from managers, lack of support to apply training on the job, lack of performance evaluation following the training, lack of resources to implement what was learned in training, and lack of incentives to apply new skills and knowledge that was acquired in training (Stolovitch & Maurice, 1998).

The return on investment process must be effectively applied within the organizational structure (Phillips, 1997). It must be simple with an easy to understand methodology. It must be economical, quick and easy to implement. Above all, the methods and assumptions underlying the methodology must be credible and theoretically sound. The method must also account for other extraneous variables that could have had an impact on the results of the training. The process and method should be applicable to a variety of human resource situations. The process should be flexible enough to apply to pre and post-test scenarios over any time period. The process should be applicable to hard as well as soft data. The method must include all relevant costs associated with the program. Finally, based on the prior items, that process must then have a track record of success.

There are four ways to measure return on investment from training (Parry, 1997). The first preference is to use hard data whenever it exists. Hard data tends to exist more readily for skilled, non-managerial positions, for hard skills rather than soft ones. This is especially true in the manufacturing assembly line type of process. Where such data exists, it is important to establish the benchmark level of what the performance was prior to the training. A second method is to obtain estimates from trainees and managers. While this is the easiest method to estimate return on investment, it is also the most subjective, hence the most prone to errors. The third method is the use of action plans. Trainees prepare an action plan that details how the knowledge that is being acquired through training is to be applied at work. These plans are

shared with management and at a later date, both can evaluate the benefits to the organization from the action plan. The fourth method is a cost benefit analysis that is the most accurate method to compute return on investment. This requires the determination of the dollar value of costs of the training and the benefits received by the organization.

Costs can be of three types. There are one-time costs, such as needs analysis and design, cost per offering, such as facility rental and instructor's salary, and cost per participant, such as meals and training materials. Benefits fall into four categories. They are savings in time, increased quantity, improved quality, and personnel data. The latter proposes that as a result of a training intervention, items such as lower absenteeism, fewer medical claims, and reduced grievances need to be taken into account to properly value the effects of the training. Another benefits schema that can be found involves using revenues per employee, profits per employee, sales per employee, changes in market share, changes in customer satisfaction and retention, stock price, earnings per share, and return on equity (Ford, 1999). The specific items that are the most appropriate to be applied are based on the actual type of employee function and the type of training involved.

#### Return on Investment Studies

There have been numerous ROI studies done over the years on training interventions. These studies may use a variety of different analytical tools but they all have one thing in common. They all go far beyond Kirkpatrick's level one and go into some form of level four return on investment calculation. While more examples could have been obtained, the sixteen that are provided are sufficient to demonstrate that organizations see a value in doing return on investment calculations for training interventions. The key issue here is the method in which the data is obtained to determine the components for the return on investment calculation. Many still

make use of asking the participants the value of the training intervention to their performance on the job.

The following sixteen case studies represent efforts that have been undertaken to demonstrate the value of training to the financial bottom line of the organization. These cases represent attempts at level 4 evaluation using a return on investment model and beyond. These are only a few of the return on investment case studies that can be found in the literature but they do serve as examples of the kinds of evaluative analysis that are being done by a variety of organizations. With the methods in use today, return on investment projections can be staggeringly high, sometimes implying that if an organization did only training, it might make the returns required by an organization's business plan. The issue is not necessarily the accuracy of any particular study since they are all based on a set of different assumptions their sponsors have about measuring the value of training. The more significant point is that processes can be developed to measure the impact of training to any organization's bottom line. Each of these studies is trying to get to an answer about the value of training using different variables and different assumptions. The results show extremely high returns that question the validity of such numbers.

Magnavox Electronic Systems, a manufacturer of satellite communications equipment, sought to reduce its illiteracy problems with its 250 hourly employees (Ford, 1994). The company conducted an assessment of 109 volunteers using the Test of Adult Basic Education procedure. Average reading ability was found to be at the eighth grade level, and math was at the seventh grade level. The results were obtained for hourly as well as salaried workers. These results dropped one grade level when the results of salaried employees were removed from the evaluation. The company began a program of assessment and analysis of problems workplace

illiteracy was causing. This was followed by the creation of a training program designed to address the issues uncovered by this study.

An 18-week program, divided into three six-week terms was begun. Classes were in an informal lecture-discussion format. The material included sample work documents that were used for reading material. Discussions of the content followed. Evaluations were based on a post course questionnaire, pre and post-tests to measure learning, and a time series analysis to determine bottom line benefits to the organization. Reading skill improvement was measured at an average of a 15% gain. Math skills improved an average of 21%. Productivity changes were measured for the group that attended training and another group that had not. Productivity rose 45% for the participant group. The control group rose only 5%. The computational process consisted of taking the productivity increase of 45% multiplied times the average hourly wage. This resulted in a monthly labor savings of \$24,503. Total costs for the training were established at \$38,233. The return on investment was calculated at  $(\$321,600 - \$38,233) / \$38,233$  for an ROI of 7.41 or 741%.

The Coca-Cola Bottling Company of San Antonio conducted a survey among its distributors and found a consistently identified need for supervisory training (Parry, 1994). At that time, there was no formal training program for supervisors. Supervisors were also the primary source of providing training for all other employees. The company hired a local vendor to develop a series of supervisory workshops. Instructors were chosen from the senior staff of the organization. The participants attended a class for a week followed by discussions with their managers the following week to decide how to implement changes they had decided to make as a result of their prior class session. A new module was presented every other week followed by these managerial discussions. Each manager and supervisor calculated the expense of the

changes they had implemented as a result of the eight modules. Benefits were converted into estimated dollar values. Some of the results of the changes included a reduction in delinquent accounts, increase in route sales, and a reduction in dispatching errors, the latter resulting in savings in fuel consumption, lost hours, and wear and tear on the vehicles. The costs and benefits were totaled for each of the participants. The organization had realized benefits of \$526,000 and costs of \$34,000, which produced a return on investment of 1,447%.

Litton Guidance and Control Systems sought to implement self-directed work teams in a plant that assembles, tests, and delivers inertial navigation systems for the military (Graham, Bishop, & Birdsong, 1994). Senior management reviewed policy and procedural changes that would enhance employee productivity. Eventually, there were 133 teams covering all three eight hour shifts. Each employee received three days of off-site training. Supervisors attended workshops on conducting effective meetings, problem solving, communications, generating ideas, and group dynamics. Each factory team held weekly meetings where representatives of member teams shared information and worked to solve problems common to all teams. Savings that could be attributed to other than work teams were factored out. Then 40% of the remaining savings was removed as an allowance for unknown factors, making the ROI calculation process very conservative. Productivity, attributable directly to the teams, increased 46%. The dollar amount of savings equaled \$7.5 million while costs were estimated at about \$1 million. The return on investment was then calculated at  $(\$7,500,000 - \$1,000,000) / \$1,000,000$  or 6.5 or 650%.

International Oil Company sought to improve the customer service skills of their dispatchers (Payne, 1994). Eleven dispatchers handled 600 calls per shift, eight to ten calls per dispatcher per hour. Dealers made up 60% of the calls. The other 40% was made up of drivers



calling for assistance. Problems included not working as a team, poor telephone manners, absenteeism, and poor rapport with the dealers. This situation led to numerous customer complaints. Managers, dispatchers, and dealers were all interviewed to get their perspectives on the problem. A training program was designed with seven objectives in mind. They consisted of telephone etiquette, rapport building, developing listening skills, improving problem solving skills, defusing anger, working as a team, and dealing with stress. An evaluation was conducted upon completion of the training based on pullout records, dealer complaints, and the results of dealer surveys, before and after the training. Each pullout cost the company \$250. A pullout occurs when the driver cannot unload the gasoline due to some problem at the station. A reduction in the number of pullouts saved the company \$354,750. Each dealer complaint cost the company \$50 of a manager's time. Prior to training the cost of dealer complaints was \$2,200 and had dropped to \$300 after the training. Absenteeism cost the company \$5,200 prior to the training and had dropped to \$1,200 after the training intervention was completed. The total cost of training was calculated at \$60,000. The return on investment to the organization was calculated at  $(\$360,650 - \$60,000) / \$60,000$  or 5.01 or 501%.

Midwest Banking Company, seeking to increase their competitiveness in the area of consumer loans, sought to teach their loan officers techniques in sales (Phillips, 1994). A three-day program was designed called the consumer lending seminar. It focused on prospecting, customer needs analysis, product benefits, handling objections, sales closing techniques, and cross-selling. Also introduced were ten new strategies to attract new consumer loan customers. Before and after loan volume was used to measure improvement as a result of training. Adjustments had to be made to eliminate the effects of other variables on increasing loan volume. After the program began, interest rates fell by one percent. Statistics were available for

consumer loan increases in volume for every percentage point decrease in the interest rate. An adjustment was made for this variable. Also, some increases would be realized without the training. An estimate of 10% was used as a normal increase in loan volume regardless of training. The bank conducted a loan profitability analysis. After deducting for the cost of funds and the direct costs for consumer lending, as well as corporate overhead, the net profit per loan was determined to be 1.82%. The average loan amount was \$15,500. Adjusted loans increased by an average of six per participant totaling \$20,088,000 a year for all 18 participants. This sum multiplied by the 1.82% profit margin produced a profit of \$365,601 as a result of the training. Costs were calculated at \$17,510. Return on investment was calculated at  $(\$365,601 - \$17,510) / \$17,510$  or 19.88 or 1,988%.

Information Services conducted an interpersonal skills training program in order to bring about a cultural change to establish a more supportive environment within the organization (Russ-Eft, Krishnamurthi, & Ravishankar, 1994). The program consisted of 14 segments. The skills that were taught included effective communication, adjusting to change, self-management skills, and self-worth. The 85 participants in the program were professional staff and support personnel such as business analysts and data entry operators. Ratings were assigned to each participant for pre-training and post-training skills. The calculated difference represented the benefit derived from the training. This benefit was multiplied times an assessment that indicated how much of an employee's time was spent using those skills. That amount of time was converted to salary dollars and the increase in skills was multiplied times that benefit. Significant differences appeared between the group before the training and after the training as well as differences between the trained group and a control group. Return on investment was calculated for a subset of the original group. Those trainees numbered 42. Therefore the resulting ROI

calculation to the organization was understated. The return on investment was calculated at  $(\$305,000 - \$70,000) / \$70,000 = 3.36$  or 336%.

Financial Services, Inc., a diversified financial services firm, experienced an unacceptable branch manager trainee turnover rate between 48% and 63%. The organization sought to reduce that rate through an enhanced screening and selection program (Schoeppel, 1994). This is a good example of return on investment that was also tied back to the expenses of high employee turnover. Turnover costs for the company in terms costs of screening, interviewing, training, and salaries were approaching \$10 million. A number of changes were made to their human resource practices. Changes applied to recruiting strategies, interviewing guidelines, evaluation guidelines, and individualized feedback. Even recruiting procedures were revised to appeal to proactive, sales-oriented candidates. Interviewing questions were changed to better target individuals with the characteristics desired for the job. Evaluation guidelines were changed to include a rating scale to make the evaluations as objective as possible. Finally, the process was changed to provide detailed feedback to each district manager as to how each candidate could best be trained based on the information contained in the application. One-third of the company participated in the new training. The other two-thirds did not.

Terminations, separations, and turnover rates were monitored for both groups. The cost of the training program itself was calculated at \$125,000. The cost of turnover to the company was calculated at \$10,070,780. This included the salaries and benefits of trainees who had left, and the cost of managers' time hiring replacements for the trainees who had left. Following the training, turnover was reduced from 54% to 39%, a reduction of 28%. Savings was represented by multiplying \$10,070,780 by 28%. The return on investment was calculated at  $(\$2,800,000 - \$125,000) / \$125,000 = 21.4$  or 2,140%. The return on investment figures rise significantly when

the effects of employee turnover are added to the computation. The ability of training to reduce employee turnover is an entirely separate subject deserving of its own study. Undoubtedly, there is some positive effect that certain types of training can have on employee turnover. When an estimate is made of the value of training relative to employee turnover reduction, the returns on investment being calculated tend to rise significantly.

North Country Electric and Gas decided to tie their management development program to organizational transformation and improvements to quality of service and customer satisfaction (Wescott, 1994). Existing management development programs dealt more with management theory rather than practicing the skills needed in management. The organization was faced with outdated work practices, aging equipment, poor service, low morale, and declining profits. Eight weeks of training were held, four hours a day, in the principles of applied behavior management. The program targeted eight key areas: customer service and satisfaction, service quality, productivity, cost containment, employee morale and motivation, public and stockholder relations, and profit. Different departments produced different levels of improvement. Workflow improvements averaged an increase in productive time to 6.8 hours per day, resulting in savings of \$56,600. New service installations went from 3.83 per day to 6.0 per day, resulting in savings of \$25,220. Other changes in work scheduling resulted in less rework time, resulting in savings of \$27,900. Total benefits were calculated at \$1,077,750. The costs to implement the program were \$215,500. The return on investment was calculated at  $(\$1,077,750 - \$215,500) / \$215,500 = 4.0$  or 400%.

When deregulation began, Yellow Freight System found itself facing heated competition, reduction in profits, and a number of entities being forced out of business (Zigon, 1994). Yellow Freight System concluded that only the largest companies in their industry with the best

performing employees would survive the competition brought about by deregulation. The company decided that the entire workforce needed to be trained. A new improvement process based on updated job models was begun to bring the organization back to profitability. One goal was to reduce the high cost per bill due to poor planning on pickups and deliveries. A new job model was instituted that reduced the cost per bill by \$1.30 from scheduling alone. Another goal was to increase productivity at all employee levels using a system of coaching and rewards. Average terminal profits increased from \$43,253 to \$49,024. A third goal was to take terminals with low bill counts and high costs per bill to improve efficiency using new job models. The cost per bill decreased an average of \$1.79. Another goal was to take terminals with low bill counts and increase their volume based on new job sales models. Bill count increased on average from 7,765 to 9,405. The total value of improvements came to \$20,791,000 at a cost of \$1.7 million. Return on investment was calculated at  $(\$20,791,000 - \$1,700,000) / \$1,700,000 = 11.15$  or 1,115%.

Healthcare, Inc., a regional provider of health care services through hospitals, clinics, and health maintenance organizations, sought a sexual harassment prevention workshop for first line managers (Hill & Phillips, 1997). The organization was concerned about the on-going problem of sexual harassment complaints and the cost of defending against such claims. The one-day workshop had a number of objectives. First, be able to administer the company policy on sexual harassment. Second, be able to recognize inappropriate and illegal behavior. Third, be able to investigate and discuss sexual harassment charges. Fourth, be able to conduct meetings with the staff to discuss sexual harassment issues. Fifth, reduce the number of sexual harassment complaints with the goal of keeping the workplace free of sexual harassment. Pre-tests and posttests were administered to measure knowledge of company sexual harassment policy and

inappropriate behavior. After the program, data was reviewed using formal internal complaints of sexual harassment, external complaints of sexual harassment, and employee turnover. Seventeen training sessions had been conducted involving 655 managers. A year after the program, internal complaints had dropped from 55 to 35, external complaints had dropped from 24 to 14. Total savings were estimated at \$3,200,908 that included savings from litigation and settlement costs. The cost of the training program was \$277,987. Return on investment was calculated at  $(\$3,200,908 - \$277,987) / \$277,987 = 10.52$  or 1,052%.

Apex Corporation, a manufacturer and distributor of high tech leading edge solutions, sought to establish a consultative relationship between its sales force and its customer base. Its sales force exceeded 7,000 employees, producing sales in excess of \$10 billion. A needs analysis identified seven critical gaps in the abilities of the sales force. An appropriate training program was developed to fix those seven gaps. Sales personnel were expected to clearly understand a customer's business requirements. This was critical to meeting the needs of their customers. They were to be solutions oriented rather than product oriented. They were to understand how the company solutions solve the problems of business. They were to meet the customer's productivity needs. They were to be able to form strategic plans with their customers. They were to effectively link technology to business solutions. And they were to consistently raise the level of contact within an account. The solution was seen as a customer centered sales strategy that focused on understanding the customer's business goals and requirements and how Apex could help them meet those goals and requirements. The company established an advanced sales training program referred to as customer centered selling. The actual sales effects of the training program was based on interviews with participants who were asked to determine if each sale had been due to the techniques they had learned through the program. Based on participant input,

sales due to training were calculated at \$215,719 per participant, for each of the 7,000 participants. The return on investment was calculated at  $(\$3,451,504 - \$112,000) / \$112,000 = 29.817$  or 2,982%.

First Union National Bank sought a training program to train its managers to change their approach with customers and prospects and to introduce new products as solutions for the issues facing their customers (Stone, Steele, Wallace, & Spurrier, 1997). The goal was to become a bank that was proactive, not reactive. The bank sought to build their business on relationships rather than just transactions. Sales efforts should be based on solutions rather than products. The staff would be made up of financial consultants rather than sales vendors. Operations would be involved in delivering an entire banking process rather than just providing credit. Finally, new sales would be based on a team effort rather than recognizing any single individual. Data was collected in terms of number of new relationships, loan growth, deposit growth, sales of capital market products, sales of other products, sales of cash management services, improved customer satisfaction, and decreased employee turnover. Data collection methods included monitoring of business measures, employee questionnaires, and customer surveys. Questionnaires were sent to all 92 participants, including 33 managers.

The questionnaire listed a number of variables that would have had an impact on branch performance. Individuals were asked to assign a value for each of the variables. Based on participant reaction, on average, participants assigned 32.5% to training and 23.6% to the reengineering process. Managers, on average, assigned 22.7% to training and 23.4% to reengineering. A list of 21 improvements was evaluated along with the annual contribution of each category in terms of income for the bank. Feedback was obtained as to the percent of each category that was due to the training that had taken place. The total benefit increase came to

\$1,028,250 against training costs of \$698,725. Return on investment was calculated at  $(\$1,028,250 - \$698,725) / \$698,725 = .472$  or 47.2%.

Bell Atlantic Network Services sought to develop a computer-based training program for telephone maintenance employees (Hodges, 1997). Bell Atlantic used a subscriber loop carrier system to reduce the number of facilities needed to service a certain geographic area. Loop electronics allowed multiple simultaneous conversations over limited facilities. Maintenance administrators and repair service clerks were responsible for handling trouble calls from customers. The purpose of the training was to help employees recognize subscriber loop carrier problems and properly interpret test results.

The organization expected the training to increase customer satisfaction by taking care of the trouble the first time and eliminating repeat calls. Training focused on getting employees to properly analyze the problem and take the proper corrective action the first time. This would also reduce the number of dispatched repair calls by diagnosing the problem correctly the first time, eliminating dispatched calls to fix a problem that had not been diagnosed correctly. To determine the impact of training, participants were asked to complete a perception questionnaire. It was sent to all of the participants, basically asking them for their opinion of an estimate of the impact of training. The mechanized trouble analysis system would be used to determine the average number of repeat trouble calls. Repeat trouble calls would be a measure of the success of the training. The benefits of the training were determined based on the average number of repeat calls times the percentage reduction. The percentage reduction was equal to the number of calls saved times the value of each call. The value of each call was determined by the hourly rates of the maintenance administrator or the repair service technician. The total cost of the training was calculated at \$39,790. The value of the benefit to the company after training was calculated at



\$166,680. Return on investment was calculated at  $(\$166,680 - \$39,790) / \$39,790 = 3.66$  or 366%.

Global Technology was interested in evaluating leadership training for new managers (Phillips, 1997). The study of the management leadership program had three objectives. The first was to assess the specific impact of the program's contribution to the bottom line of the organization. The second was to identify specific barriers to successful implementation. The third was to recommend specific changes to the program. The evaluation methods used for the first objective included questionnaires and focus groups. Isolating the effects of the training and converting the effect into a dollar value were based on participants' estimates. Participants and their managers rated improvements in skills as a result of taking the course. Identified changes in skills and knowledge that were rated as significant included personal development planning, developing teamwork, valuing diversity and human rights, conflict management, team leadership, objective setting, coaching, developing employees, performance reviews, and providing rewards and recognition. Skills such as managing change, managing time, strategic planning, employee selection, and ethics were not rated as having as significant an impact on the participants. The total cost of the program was \$3,540 per participant. The business impact of the program to the organization was computed by multiplying the percent of improvement to the organization times the dollar value of the change. This amount was multiplied by a confidence level adjustment for participant's uncertainty of their estimates. The benefit per participant was estimated to be \$20,500. The return on investment was calculated at  $(\$230,625 - \$88,500) / \$88,500 = 1.60$  or 160%.

Dell Computer Corporation sought to determine the impact of its sales negotiation course to the bottom line of the organization (Tesoro, 1998). The Dell Sales Negotiation Course is a

two-day interactive training program. The course helps employees to improve their skills in negotiating agreements, obtaining needed resources within the organization, shorten sales cycles, being better prepared for the negotiation, and increasing customer loyalty. The measurable goals for the participants were to exceed product quotas, exceed total revenue, increase product volume sales, increase total revenue sales, exceed profit margin quota, and increase profit margin volume sales. Pre-training and post training metrics were collected for both the training and control groups. A total of 57 employees participated in the program. They had their sales quotas and performance-based compensation tracked by the sales information database. The control group was another 57 sales employees who were randomly selected to participate in the control group. The average improvement of the participants in the eight business metrics was 17.63%. The control group experienced a 13.59% improvement. The total benefits for the participant group was \$271,302 with costs of \$51,849. Return on investment was calculated at  $(\$271,302 - \$51,849) / \$51,849 = 4.23$  or 423%.

All of the above case studies used a return on investment model to calculate the value of training to the organization. The most difficult part of the calculation is isolating the effects of training from any and all other variables that may influence the financial results of the organization. In many cases, isolating the effects of the training was based on estimates of either the participants or their supervisors. Then second problem was valuing the effects of the change. In some cases, the valuation was based on time savings, representing quantifiable data. In many other cases, it was based on the opinions of the participants. For the most part, these case studies rely on individual opinions, usually those of the participants, to complete the determination of what benefits the organization derived from the training. While the computation of return on investment is based on numbers representing costs and benefits, the methods in which some of

those numbers may have been determined can be very subjective. Herein lies the skepticism of the evaluation process in training and development. ROI calculations that are over a thousand percent seem hard to fathom. Although the methods rely on subjectivity, the important thing to remember is that subjectivity can be replaced with objective measures.

Every attempt should be made to isolate and eliminate the effects of other variables that may influence changes in the organization other than from the effects of training. This is often difficult to do. However, to justify the spending of resources on any training intervention, the specific changed behaviors on the job attributable to the intervention should be quantified. Pre-tests and post-tests should be applied to accurately measure change both before and after the training. There are a number of financial analysis methods that can be applied to evaluating training interventions (Mosier, 1990). Methods such as trend analysis, analysis of variance and covariance can help to determine what effects may reasonably be assigned to the effects of the training and what others should not.

#### Employee Attitudes

A human resource development evaluation model can go beyond just resource allocation. It should also consider a motivational component (Swanson & Gradous, 1988). Business results are often heavily influenced by attitudes involving employee satisfaction (Harter, Schmidt, & Hayes, 2002). This satisfaction can arise from factors such as the job, the organization, fellow workers, salary, and many other variables. Employee attitudes can also influence motivation. Motivation, in turn, can have a significant impact on training interventions in terms of how much individuals learn. Training can play a key role in affecting employee attitudes. Training results are especially influenced by employee choice when it comes to participation (Hicks & Klimoski, 1987). Three areas that can be impacted by employee attitudes are self-efficacy, job satisfaction

and organizational commitment. All three of these areas can have an impact on employee turnover and absenteeism. If employee turnover and absenteeism can be reduced through an increase in self-efficacy, job satisfaction, and organizational commitment, this would represent an additional cost savings that could be attributed to training.

Self-efficacy describes an individual's belief that he or she has the skills necessary to do a specific job or task. Self-efficacy is often validated by some form of evaluation process. Self-efficacy often arises as a result of training (Stevens & Gist, 1997). There is also pre-training self-efficacy. Pre-training self-efficacy is the belief that an individual can acquire skills and knowledge during a training intervention. Job involvement, the degree to which an individual is committed to his or her job, is a key antecedent to pre-training motivation. In addition, pre-training self-efficacy is closely tied in to feelings of organizational commitment (Tracey, Hinkin, Tannenbaum, & Mathieu, 2001). So another benefit to training is not only an increase in self-efficacy but an increase in commitment to the organization as well.

Other studies have found that pre-training self-efficacy leads to motivation to do well in the training intervention (Quinones, 1995). Self-efficacy is positively correlated with motivation. It is also a predictor of performance and can be used as selection criteria in the hiring process. It is also a major antecedent of leadership training. Individuals with low self-efficacy may be ideal candidates for training (Gist, 1987). Self-efficacy has been shown to not only have a positive effect on initial performance levels, but it is also is a prime determinant of performance measured much later after the training intervention (Gist, Stevens, & Bavetta, 1991). Training also is influential on post-training self-efficacy and can create a positive attitude towards the job (Saks, 1990). Increased self-efficacy arising from training can not only lead to participants' improved performance, but effect an attitudinal change in favor of the organization as well.

Job satisfaction reflects the degree to which an employee likes doing his or her job. There are three components of job satisfaction. One is values, including the subjective requirements an employee has regarding his or her job. The second is importance. This relates to the importance they place on the values they have. The third is perception, relating to their present job and how it meets their values. While job satisfaction does not necessarily result in organizational commitment (Bateman & Strasser, 1984, and Vandenberg & Lance, 1992), other studies have shown that job satisfaction is an antecedent to organizational commitment. Most personal and organizational characteristics influence job satisfaction directly. They influence organizational commitment indirectly through its impact on job satisfaction that in turn impacts organizational commitment (Williams & Hazer, 1986). Job satisfaction has also been shown to have a negative relationship to employee turnover and absenteeism (Zellars, Hochwarter, Perrewe, Miles, & Kiewitz, 2001). In one meta-analysis, correlation between turnover and job satisfaction averaged  $-.40$  while correlation between absenteeism and job satisfaction averaged  $-.25$  (Muchinsky, 2000). And employee turnover and absenteeism are concepts that have been well documented as to the negative impact they have on the organization.

Organizational commitment is the attitude an employee has toward the organization that links the identity of the individual to that organization (Sheldon, 1971). It can be defined as the strength of an individual's identification with and involvement in the organization. It can be characterized by a belief in the organization's goals and values, a willingness to work hard on behalf of the organization, and a desire to belong to the organization (Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991). All workers have some amount of organizational commitment and those with a high degree of commitment are the most likely to be affected in training interventions. The single largest predictor of organizational commitment is the length of time the

employee has been with the organization (Stevens, Beyer, & Trice, 1978). Age is another such factor. Employees who have been with an organization for a long time have reasons for staying. These reasons can be very different for different employees.

Participant expectations and desires prior to training and their experiences during training, can affect post-training self-efficacy and commitment (Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991). This, in turn, can affect employee motivation. Ultimately, organizational commitment can have an effect on an organization's profit level (Benkhoff, 1997). Expectancy theory states that individuals will try harder and apply what they've learned if they expect that such behavior will lead to a desired reward (Tollefson, 2000). Exchange theory suggests that training can be viewed as an investment in the relationship between an organization and the employee and can contribute to organizational commitment (Farrell & Rusbult, 1981). The investment model proposes that organizational commitment is often a predictor of employee turnover.

Organizational commitment is a function of three items. First, it is a function of costs and the rewards provided by the job. Costs include inadequate resources, travel time to work, unfair promotion practices, and undesirable shifts. Rewards include salaries and promotions. Second, it is a function of the quality of job alternatives. It makes a difference if there is a shortage of jobs with high unemployment or if there are more jobs than there are people available to be hired. Finally, organizational commitment is a function of the magnitude of the employee's investment in the job. This investment can include length of service, significant non-portable skills, and retirement programs. Positive training experiences can increase the employee's perceived value of training. That heightens the value of the organization's training investment in each employee. The more positive the training experience, the greater can be the commitment to the

organization. Ultimately, organizational commitment is positively related to the employee's attitude towards training and training motivation (Carlson, Bozeman, Kacmar, Wright, & McMahan, 2000).

### Summary

There is no shortage of reviews of historical training evaluation methodologies (Swanson, 1998 and Wang, Dou & Li, 2002). A review of the literature indicates that training evaluation, especially at Kirkpatrick's levels 3 and 4, is not widely conducted in businesses today. Despite the numerous training evaluation models that stress the benefits of levels 3, 4, and return on investment calculations, businesses today still have not embraced these processes to evaluate their training expenditures. There are a number of examples that indicate the use of ROI is a competitive advantage. The literature is full of examples of case studies that indicate training interventions that have provided significant returns on investment. Part of the problem is a lack of credibility in an ROI calculation process that produces quadruple digit rates of return, based on participant opinions.

## CHAPTER 3

### METHODOLOGY

#### Introduction

The process involved applying the proposed training method, and doing a statistical analysis to determine if there was a significant difference between the return on investment (ROI) results of the different groups of participants across different training interventions. There were four groups of participants. They were underwriters, non-underwriters, assistant underwriters and other employees. The first null hypothesis asked if there was a difference in the ROI results for underwriters who participated in three different training interventions. The second null hypothesis asked if there was a difference in ROI results between the underwriters and non-underwriters who participated in the training interventions. The third null hypothesis asked if there was a difference in ROI results between the assistant underwriters and other employees in similar grades. The fourth null hypothesis asked if there was a difference in ROI results between underwriters based on years of experience. The fifth null hypothesis asked if there was a difference in ROI results within test scoring bands.

The review of the literature revealed attempts to connect the results of training to the financial results of the organization. That effort was continued in this study. The results of the department to the bottom line of the organization prior to the training intervention and after the training intervention was analyzed. These results are referred to as the combined ratio. Swanson & Gradous (1988) reported that a change in motivation was a possible outcome of training interventions. This study continued that effort on a post-hoc basis by obtaining employee participant attitudinal survey data for the time period prior to and after the intervention. That data was also analyzed.



## Research Hypothesis

This study is based on two research questions and five research hypotheses, as described in Table 1. The first research question involves the creation and application of a formula for objectively evaluating training interventions without the use of participant opinions. This resulted in the creation of a seven-step formula that evaluates training using such unique features as salary, salary grades and job factor analysis. This process places the burden for determining the skills needed in each job position with the organization's leadership staff. It does not ask participants how well their needs were met by the intervention nor how much their newly acquired skills contributed to the bottom line of the organization.

The first four null hypotheses support the first research question. The first null hypothesis tests for a statistically significant difference between three training interventions administered to a group of underwriters. The second null hypothesis tests for a statistically significant difference between training interventions administered to a group of underwriters and a group of non-underwriters. The third null hypothesis repeats the second except the intervention is administered to a group of underwriting assistants and other employees of a similar grade level. The fourth null hypothesis tests for a statistically significant difference between underwriters with the results grouped by levels of experience. The second research question seeks to determine if a pre-test can be used as a predictive tool to make a decision regarding training interventions at the individual participant level. It builds on the foundation set by the first research question and the supporting first four null hypotheses. The second research question is tested by the fifth null hypothesis that looks for a statistically significant difference in ROI score bands based on pre-test scores of the individual participants. Using pre-test scoring bands is vital in predicting the value of training prior to that training taking place.

Table 1

*List of Research Questions and Hypotheses*

---

Item	Questions and Hypotheses
Research Question #1	Can an objective method for evaluating training be developed?
Null Hypothesis #1	There will be no statistically significant difference in return on investment results of an underwriter-training program consisting of three different training interventions across a sample of underwriters.
Null Hypothesis #2	There will be no statistically significant difference in ROI results of an underwriter-training program between a group of underwriters and a group of non-underwriters.
Null Hypothesis #3	There will be no statistically significant difference in ROI results of an underwriter-training program between a group of assistant underwriters and a group of other employees.
Null Hypothesis #4	There will be no statistically significant difference in ROI results of an underwriter-training program between underwriters with less than 5 years of experience and underwriters with 5 or more years of experience.
Research Question #2	Can a pre-test be used as a predictive tool in the selection process of participants based on a future estimated ROI?
Null Hypothesis #5	There will be no statistically significant difference in ROI results of a training program within ROI score bands

---

## Population

This study was done using a sample of commercial underwriters from a single insurance company. However, that sample represented all of the underwriters from within the company. The population to which this research applies can be extended to any group of underwriters within the insurance industry. The estimated size of the population of underwriters is estimated at 100,000. It is suggested that this methodology can be applied to any participants in any industry or organization because the process that was applied to training evaluation could adjust itself to any situation or organization. This is accomplished by the application of two adjustments for job relevance and the post-training assessment factors. Because they are unique to each training intervention and controlled by the sponsoring organization, the methodology could be extended to any company outside the test sample. The success of the methodology would depend on the accuracy of the adjustment factors established for each training intervention.

## Sample

This study was conducted at a major financial institution in the western U.S. The participants were all members of the company's commercial underwriting department, consisting of underwriting managers, underwriting supervisors, underwriters and assistant underwriters, as shown in Table 2. In addition, there was a group of non-underwriters who also participated in the training. The organization required this training of all of its underwriting department personnel. Commercial underwriters make decisions about new businesses seeking insurance coverage in terms of risk acceptability and pricing. Based on education, experience and delegated authority, underwriters are expected to review the incoming applications for new and renewal business and make decisions on accepting or declining individual risks, and for those that are accepted, determine terms and price.

Table 2

*Job Types of Participating Sample*

Job Title	# of Employees	# of ROI Scores
Underwriters	27	319
Other Non-Underwriting Staff	18	168
Underwriting Assistants	6	33
Other Non-Underwriting Assistant Staff	19	70
Total	70	590

If successful across an entire book of business for a prolonged period of time, these decisions will result in increased profitability for the organization. The underwriting decision-making process is very judgmental. It is difficult to train individuals in this subject. It is typical of other soft skills training programs. These underwriters all worked in the area of business insurance. Underwriters evaluate applications, often getting dozens in any given day, and must make a decision on the acceptability of the risk, as well as the terms and conditions, and price that is to be offered. These decisions must be made quickly during a relatively short period of time as many applications can be received at once.

#### Instrument Validity and Reliability

The actual instrument used for the pre-tests and post-tests were questions taken from materials purchased from the Insurance Institute of America. The tests consisted of ten, fifteen or twenty multiple-choice questions, depending on the length and complexity of the reading material that accompanied each line of commercial insurance. Each training session was based on a single line of insurance. There were 36 training sessions conducted with the staff over a two-year period.

Validity involves a testing instrument that measures what it should measure within a given discipline. The training intervention used materials made available through the Insurance Institute of America. The materials consist of textbooks and workbooks for each course. The tests were based on questions provided as part of the training materials that were used by the participants. The Institute's Associate in underwriting certification program consisted of three parts; Insurance 23, Introduction to Commercial Insurance; Associate in Underwriting 65, Property Underwriting; and Associate in Underwriting 66, Casualty Underwriting. This program is in use by most insurance companies in the U.S.

Reliability involves obtaining the same results over time with the instrument. The Insurance Institute of America materials and exams have been in use for over 40 years. The contents of the programs are considered a standard within the industry. Content is update every two years or so and is reviewed by a number of industry personnel. The IIA also conducts an item analysis of each test question used in their national examinations. Questions and responses are adjusted as needed on a regular basis.

The company's cost of training was based on \$140 per participant. This cost was provided by the company's human resources department. This cost includes training materials, instructor's time, and physical costs such as room rental and utilities. The cost does not include lost opportunity costs for the participants' time. Based on weekly classes and the number of participants, the lost opportunity costs can be very significant.

### Methodology

The process began with the creation of a formula that could be used to evaluate any kind of company-elected training intervention. The formula did not include a component for using the opinions of participants. The organization was required to make a decision on the importance of

the training intervention compared to other required training interventions through the use of the job relevancy factor. This required in turn that the organization had highly defined their job descriptions such that they knew what training interventions were needed in total for each position. The job relevancy factor should total 1.00 when all jobs have been evaluated. The entire formula is shown in Table 3.

Pre-tests and post-test scores were collected for every participant from every training intervention. Salary mid-point data was obtained from the human resources department for each of the participant job grade categories across all participant groups. Information was also obtained on the cost of training. However, lost opportunity costs due to training were not included in the calculation. This information is difficult to compute because the lost opportunity costs include the opportunity of writing additional amounts of premium. As this activity is never a certainty despite the amount of time expended, the lost opportunity costs of the underwriters were not included.

In addition to the pre-test and post-test data, data was collected on underwriter performance for six months after the end of each training intervention. That data was necessary to calculate the post-test adjustment factor. This involved the application of the training to on the job activities. Data was also collected on the financial results of the organization prior to the training intervention and after the training intervention. This profitability information is commonly published information and covered a seven-year period from 2000 through 2006. Data was also collected from company employee surveys to address attitudinal issues. This information was used to conduct an evaluation of employee attitudes relative to their satisfaction with the training intervention. Comparative data between the participants and non-participants in the training intervention was applied.

Table 3

*Training Evaluation Methodology*

Step	Description
1	Participants take a scored pre-test prior to the training intervention.
2	Participants take a scored post-test following the training intervention.
3	Compute the difference between scores to measure learning.
4	Multiply each score by a Job Relevancy Factor.
5	Multiply each adjusted score by a Post-Test Adjustment Factor.
6	Multiply each score by the mid-point of the participant's salary grade.
7	Take the dollar benefit minus the cost and divide the results by the cost.

The proposed return on investment methodology was supported by the research described by Campbell (1993). Training costs, described as element number one were not applied as the organization had a pre-established amount calculated for cost of training. Element number two, training effectiveness, is based on Kirkpatrick's four levels. Level 2 was represented by the post-test. Level 3 was represented by the application of the post-test adjustment factor. Level 4 was the return on investment calculation. Campbell's (1994) element number three – methods used to determine the cost effectiveness of training, discusses the pros and cons of four different methods of evaluation. They are return on investment, cost benefit ratio, bottom-line evaluation and payback period. This study on return on investment advanced Campbell's application of bottom-line evaluation. He proposed a formula that takes the participant's total annual compensation package times the percentage of time spent on performing the task the training was designed for. That amount was referred to as component pay. A gain in productivity was computed based on post-training productivity minus pre-training productivity. That productivity

gain was then multiplied times the component pay to produce the value attributed to the training intervention.

Campbell's application of salary to the evaluation process was a significant improvement to training evaluation methodologies. By using component pay, he was tying the benefits of training to one of the basic components of any organization, its salary structure. The current methodology proposed took Campbell's process one-step further by recognizing that performing a task may not be able to be solely assigned to one training intervention. For example, within the underwriting discipline, it could be described as one task that is 100% of the job but requires a variety of different training interventions to complete. The job relevancy factors used were .025 and .010 for underwriters and non-underwriters respectively. The value of Campbell's method was the limitation of the salary component to a percentage, so as to not overstate the value of any single training intervention. While Campbell's methodology was well suited to technical skills training it would be difficult to apply in soft skills training. The application of the job relevance factor in the proposed methodology adjusted for this variation because it readily made allowances for multiple training interventions at once on a percent of salary basis and the process could be applied to other soft skills training as well.

A second citation used to evaluate and support the results of this study was the forecasting financial benefits method proposed by Swanson (2001) and Swanson & Gradous (1988). The forecasting financial benefits model evaluates human resource developmental training in terms of dollars and cents. A portion of the model uses performance value analysis. This evaluation method identifies three major components: the performance value resulting from the training, the cost of the program and the benefit of the program. It specifically focuses on quantitative measurement of the results of training. By forecasting the benefits of any



prospective training intervention in advance, using financial analysis methods, senior executives have a process of making a decision that is more aligned with their normal decision-making process.

A basic underlying concept behind the proposed methodology was that the goal of any organization should be to get back at least 100% of what it pays for in salary for each of their employees. Most employees are hired at or below the mid-point of their salary grade. Growth and performance justifies future salary increases as the employee approaches and surpasses the mid-point salary level. In time, experience and skill sets dictate promotion to a higher-grade level and the process starts all over again. However, it is doubtful that a single training course could represent the total skills and knowledge of any given job. The job relevance factor adjusted for the applicability of the training intervention compared to the total skills and knowledge required for the job. This factor should always be based on a complete analysis of the skill sets and competencies that are required by each job and should be established by the participant organization. The natural outcome of this process is an entire training program outlined for each position in the organization. In any organization, it would be appropriate to have these competencies prioritized based on the objectives of the organization and possibly used to determine training schedules (Mirabile, Caldwell, & O'Reilly, 1987).

Data was also obtained and presented that addressed changes in the financial bottom line profitability of the organization. A term unique to insurance companies, the combined ratio, was used to represent bottom line profitability. The combined ratio is called that because it adds together the loss ratio for the organization and the expense ratio for the same period of time. However, insurance company profitability is based on a number of often contradicting variables. A clearly defined link was not established for company bottom line profitability, other than a

demonstration of trend. Data was also reviewed involving company employee attitudes taken from an annual employee survey and compared with the results throughout the training intervention period.

#### Data Treatment

The collected data was run through the seven-step process described in Table 3 to calculate the subject company's return on investment. The means of the results of the different participant groups were tested using both analysis of variance (ANOVA) and a number of *t* tests with the five null hypotheses proposing that there would be no significant difference ( $H_0: P_1 = P_2$ ) between the means of the groups participating in the intervention. The analysis of variance was conducted against three different training interventions for underwriters and against three different ROI scoring bands. The *t* tests were used to compare the results of groups of underwriters against non-underwriters.

#### Summary

The collected data was analyzed and using a series of ANOVA and *t* tests of five separate null hypotheses, as detailed in Table 4. These tests were used to answer the research questions for the review of a new methodology for evaluating training interventions. The tests were based on a variety of combinations of test subject groups. Results were used to ascertain the validity and reliability of this process. It also made use of a financial analysis method with which most decision-makers in organizations would be comfortable. In addition, a post-hoc study was made of an attitudinal survey of the department covering the last three years to investigate any possible relationships that may be produced. This information is compared to the data gathered from the survey of the literature on employee attitudes and the positive and negative effects it can have on training interventions.

Table 4

*Statistical Treatment of Research Questions (Null Hypothesis = H<sub>0</sub>)*

Issue	Research Questions and Hypotheses	Treatment
Question #1	Can an objective method for evaluating training be developed?	
H <sub>0</sub> #1	There will be no statistically significant difference in ROI results of an underwriter-training program consisting of three different training interventions across a sample of underwriters.	ANOVA
H <sub>0</sub> #2	There will be no statistically significant differences in ROI results of an underwriter-training program between a group of underwriters and a group of non-underwriters.	<i>t</i> Test
H <sub>0</sub> #3	There will be no statistically significant difference in ROI results of an underwriter-training program between assistant underwriters and a group of other employees of similar grade.	<i>t</i> Test
H <sub>0</sub> #4	There will be no statistically significant difference in ROI results of an training program between underwriters with less than 5 years of experience and underwriters with 5 or more years of experience.	<i>t</i> Test
Question #2	Can a pre-test be used as a predictive tool in the selection process of participants based on a future estimated ROI?	
H <sub>0</sub> #5	There will be no statistically significant difference in ROI results of an U/W training program within ROI score bands.	ANOVA

## CHAPTER 4

### FINDINGS

Data was collected from a mandatory company-training program to evaluate return on investment (ROI) results for four groups of employees in a major western company. The training consisted of commercial underwriter training conducted in 2004 and 2005. The underwriting training was based on material purchased from the Insurance Institute of America (IIA). This material has been in use throughout the industry for the last 40 years and is an industry standard in the training of underwriting personnel.

The study is based upon two research questions and five null hypotheses. The study was completed based upon an analysis of variances (ANOVA) and *t* tests. The first research question asks can an objective method for evaluating training be developed? This question is supported by four null hypotheses:

1. There will be no statistically significant difference in ROI results of an underwriter-training program consisting of three different training interventions across a sample of underwriters.
2. There will be no statistically significant difference in ROI results of an underwriter training program between underwriters and non-underwriters.
3. There will be no statistically significant difference in ROI results of an underwriter-training program between a group of assistant underwriters and a group of other employees of the same grade.
4. There will be no statistically significant difference in ROI results of an underwriter-training program between underwriters with less than 5 years of experience and underwriters with 5 or more years of experience.

The second research question asks if a pre-test can be used as a predictive tool to make a decision regarding training interventions at the participant level based on future estimated ROI. It builds on the foundation set by the first research question and the first four null hypotheses. The second research question is supported by the fifth research hypothesis:

5. There will be no statistically significant difference in ROI results of an underwriter-training program within ROI score bands.

#### Participants in the Study

The participants were employees from a large financial services firm located in a western state. The participants were underwriters and employees from other departments. There were a total of 590 tests administered to four participant groups, shown in Table 5.

Table 5

#### *Job Classification of Participants*

Group Code	Participants	Number of Tests
Test Group A	Commercial Underwriters	319
Test Group B	Non-Underwriters	168
Test Group C	Assistant Underwriters	33
Test Group D	Non-Assistant Underwriters	70
	Total Test Scores	590

The first group consisted of commercial lines underwriters and they were administered a total of 319 tests. The second group of participants consisted of non-underwriting employees who participated in the same training interventions as the underwriters. The non-underwriting employees were administered a total of 168 tests. The third group of participants consisted of underwriting assistants and they were administered a total of 33 tests. The fourth group consisted

of other employees who were matched up with the underwriting assistants in terms of grade and experience. They were administered a total of 70 tests. All of the commercial lines underwriters and underwriting assistants in the organization participated in this training program. The non-underwriters came from other departments represented by risk control specialists and premium auditors but they were still part of the commercial underwriting reporting structure. The fourth group consisted of support staff from throughout the company. The training that was administered consisted of three different applications of underwriting decision-making training.

Commercial insurance underwriters are hired to make decisions about accepting new business applications. These applications are requesting insurance coverage and the underwriters must make decisions regarding risk acceptability and price. Based on their education, experience and delegated authority, underwriters are expected to review these applications and make decisions on accepting or declining individual risks. For those risks that are being accepted, they must determine terms and conditions that are to be offered and then determine an appropriate premium charge for the risk.

The underwriting decision-making process is highly judgmental. It is difficult to train individuals in this subject because the decision-making that is required is dependent on so many variables that in some cases are very subjective. The underwriters all work in the area of business insurance. The different applications include requests for property, liability, automobile, crime, workers compensation, and many other lines of insurance. The commercial underwriters evaluate these applications, often receiving dozens in any given day, and must quickly make a decision on acceptability, terms and conditions, and price. They all undergo training in the beginning of their careers to provide them with the knowledge needed to do this job well. The training intervention was based on three formal training programs offered by the Insurance Institute of America and is

used to provide professional certifications in underwriting. The training focused on the decision-making process for accepting and rejecting applications based on the exposures being presented, by lines of business.

All of the participants were provided training based on the Insurance Institute of America's associate in underwriting (AU) program. This program consisted of three separate parts. The associate in property underwriting program was conducted in 2004 and the associate in casualty underwriting program was conducted in 2005. The participants who were non-underwriters of similar grade level who participated in the same training formed a separate group for analysis purposes. A separate general commercial insurance program was conducted in 2004 at a simpler level for assistant underwriters and non-assistant underwriters of similar grade levels.

#### Data Assessment and Analysis

This section addresses each of the results of the null hypotheses in turn. For the first null hypothesis, the proposed methodology was applied to only the participants who functioned as commercial underwriters using three different training interventions. The three interventions were the three parts of the AU program. An analysis of variance (ANOVA) was conducted based on the training interventions for the underwriting group. The first null hypothesis stated that there would be no significant difference in ROI results of an underwriter-training program consisting of three different training interventions across the same group of underwriters. The training results were divided into three sections. They were basic underwriting, part one of AU 65, property underwriting, part two of AU 65, miscellaneous underwriting issues and casualty underwriting, represented by AU 66. The results of the analysis of variance are shown in Table 6.

Table 6

*Null Hypothesis #1 - Descriptive Statistics – Underwriting Group*

Group	N	M	SD	SE	Lower 95% CI	Upper 95% CI
Basic	81	.9700	2.29759	.25529	.4620	1.4780
Property	128	1.6427	2.48097	.21929	1.2087	2.0766
Casualty	110	1.8117	2.91545	.27798	1.2608	2.3627
Total	319	1.5302	2.61001	.14613	1.2426	1.8177

The Levene's test for homogeneity of variance was conducted using the above data. The significance at .087 was greater than .05 meaning that the results were not significant. The Levene  $F$  statistic of 2.456 was also less than the critical value of  $F$  ( $F_{CV} = 3.0$ ). Thus there is no reason to believe that the variances of the basic, property and casualty interventions were different from one another. The assumption of homogeneity of variance of the analysis of variance has not been violated. The results of the Levene's test are shown in Table 7.

Table 7

*Homogeneity of Variance – Underwriting Group*

Levene's Test	df1	df2	Sig.
2.456	2	316	.087

The first null hypothesis states there would be no statistically significant difference between the means of the three training interventions for underwriters. Each intervention consisted of a different Insurance Institute of America underwriter training program. The critical value of  $F$  with 2 and 316 degrees of freedom for  $\alpha = .05$ , is  $F_{CV} = 3.0$  but the  $F$  statistic for the ANOVA ( $F = 2.652$ ) was not greater than that critical value. Therefore the null hypothesis could not be rejected. There was not a statistically significant enough difference between the mean ROI



scores of the three underwriter training interventions amongst the underwriters. The results of the ANOVA are in Table 8.

Table 8

*Analysis of Variance – Underwriting Group*

Source	SS	df	MS	F	Sig.
Between Groups	35.757	2	17.878	2.652	.072
Within Groups	2130.509	316	6.742		
Total	2166.265	318			

The second null hypothesis sought to determine if there was a statistically significant difference between the means of the two different groups of participants, the underwriters and non-underwriters, who took part in the identical underwriter training intervention. The means of each group are based on the individual return on investment scoring results. There were 319 underwriters and 168 non-underwriters who took part in the same training intervention. The non-underwriters came from a number of different departments throughout the organization such as claims and risk control technicians but were of similar grade levels to the underwriters. A *t* test of the means of the two groups of underwriters and non-underwriters was used to determine if there was a statistically significant difference between the results of the two groups of participants. The results of the *t* test are shown in Table 9.

Table 9

*Null Hypothesis #2 - Group Statistics – Underwriting and Non-Underwriting Groups*

Training Intervention	N	M	SD	SE
Underwriting Group	319	1.5302	2.61001	.14613
Non-Underwriting Group	168	.3818	1.33373	.10290

Just by viewing the means between the underwriting and non-underwriting groups, the average return on investment of 153% for underwriters participating in the training and an average ROI of 38% for non-underwriters participating in the training would seem to be significantly different. The Levene's test for homogeneity of variance, shown in Table 10, produced an  $F$  value of 65.386 at a level of significance of .05 so the test was significant and equal variances cannot be assumed.

Table 10

*Independent Samples Test – Underwriting and Non-Underwriting Groups*

*Equal Variances Not Assumed*

$F$	Sig.	$t$	$Df$	Sig.	MD	SE
65.386	.000	5.345	485	.000	1.14831	.21483
		6.425	484.645	.000	1.14831	.17873

The  $t$  test results demonstrated that the average means for return on investment, even with equal variances not assumed, were statistically significantly different. The results showed  $t(485) = 6.425, p < .05$ , two tailed.

A  $t$  test was for the training done with the underwriting assistants was also calculated. The insurance 23 training material that was used is at a much easier level than the AU 65 and AU 66 materials. It was more of a basic introduction to commercial insurance. The training sessions were comprised of underwriting assistants and non-underwriting employees who were at similar grade levels. The null hypothesis stated there was no statistically significant difference between the results of the assistant underwriting group and a group of non-assistant underwriters. There were 33 scores for the assistants and 70 scores for the non-underwriting assistants. The Levene's test for homogeneity of variance produced an  $F$  value of 46.118 at a level of

significance of .05 so the test was significant and equal variances cannot be assumed. The data is presented in Table 11.

Table 11

*Null Hypothesis #3 - Group Statistics – Underwriting Assistants and Others Groups*

Training Intervention	N	M	SD	SE
Underwriting Assistants	33	4.9630	8.18777	1.42531
Non-Underwriting Assistants	70	.4904	1.04231	.12458

The *t* test produced a value of 3.126 with a significance level at .05,  $t(32.49) = 3.126, p < .05$ , two tailed. In this intervention there was a significant difference between the mean scores of the underwriting assistants compared to the mean scores of the non-underwriting assistants. This was similar to the result found for the underwriters and non-underwriters. The results for the assistant underwriters are shown in Table 12.

Table 12

*Independent Samples Test – Underwriting Assistants and Others Groups*

*Equal Variances Not Assumed*

<i>F</i>	Sig.	<i>t</i>	<i>df</i>	Sig.	MD	SE
46.118	.000	4.518	101	.000	4.47260	.9900
		3.126	32.49	.004	4.47260	1.4307

The fourth null hypothesis stated there would be no statistically significant difference between two groups of underwriters divided by years of underwriting experience. The first group of underwriters was basically “trainees” who had less than 5 years of underwriting experience. The second group of underwriters was the more experienced group with 5 or more years of underwriting experience. The group with less than 5 years of experience produced 58 scores

while the more experienced group produced 261 scores. The results are provided in Table 13 below.

Table 13

*Null Hypothesis #4 - Group Statistics – Underwriter Results by Years of Experience*

Years of Underwriting Experience	N	M	SD	SE
5 or More Years of Experience	261	1.3198	2.31901	.14354
Less Than 5 Years of Experience	58	2.4769	3.51562	.46162

The Levene’s test for homogeneity of variance produced an  $F$  value of 12.527 at a level of significance .05 so the test was significant and equal variances cannot be assumed. The  $t$  test produced a value of 2.394 with a significance level at .05,  $t(68.42) = 2.394, p < .05$ , two tailed. In this intervention there was a significant difference between the mean scores of the underwriters with less than 5 years of experience compared to the mean scores of those with 5 or more years of experience. Results are shown in Table 14.

Table 14

*Independent Samples Test – Underwriter Results by Years of Experience*

$F$	Sig.	$t$	$df$	Sig.	MD	SE
12.527	.000	-3.095	317	.002	-1.15713	.37387
		-2.394	68.42	.019	-1.15713	.48343

The analysis produced a  $t$  value of 2.394 at a significance level of .05. The mean difference of  $-1.15713$  fell within the 95% confidence range. The null hypothesis was rejected, recognizing that there was a statistically significant difference between the two sets of scores. These results seem to suggest the organization got back a significantly greater return on investment from the training of the commercial lines underwriters who had less than 5 years of

experience compared to the other more experienced group of commercial lines underwriters who had five or more years of experience. The results produced a 148% return for the underwriting group compared to a 32% return for the non-underwriting group.

Data was further analyzed to determine if high underwriter pre-test scores would lead to lower ROI scores as there would be less gain available to be made during the training intervention. This constituted the second research question. The results shown in Table 15 indicate that as the pre-test scores increased, the number of positive ROI scores within that pre-test score band decreased and the average ROI score for that group decreased. This allows for the use of the methodology to function as a predictive tool. Should this or any organization have set a rate of return benchmark, such as a 12% return, the individuals scoring 70% or higher would not have participated in the intervention.

Table 15

*Pre-Test Scores and Rates of Return*

Pre-Test Scores	N	Positive ROI	Average ROI
70% - 99%	33	4	5.6%
60% - 69%	52	16	50.9%
50% - 59%	63	23	58.5%

An analysis of variance was completed to determine if the differences by pre-test score band were statistically significant. The bands chosen were 70% to 99%, 60% to 69% and 50% to 59%. The 50% threshold was selected as a benchmark for pre-test scores below which are considered appropriate for continuing on with a training intervention.

Only the bands that were 50% or greater were used in the comparison. A lower pre-test score should not be used to predict return on investment since its validity would depend on the

results of the actual training intervention. Higher pre-test scores were used to determine predictability value with the analysis of variance being applied to determine significance of the difference in results. The analysis of the bands are shown in Table 16.

Table 16

*Analysis of Pre-Test ROI Score Bands*

ROI Score Bands	N	M	SD	SE
70% - 99%	33	.0564	.16902	.02942
60% - 69%	52	.5092	.88669	.12296
50% - 59%	63	.5849	1.06182	.13378
Total	148	.4405	.89345	.07344

The Levene's test for homogeneity of variance was conducted using the above data. The significance at .05 meaning that the results were significant, failing the test of equal variances. The Levene  $F$  statistic of 19.246 was greater than the critical value of  $F$  ( $F_{CV} = 3.07$ ). Equal variances cannot be assumed between the variances of the means of the three return on investment score bands. The assumption of homogeneity of variance of the ANOVA cannot be made. The results are shown in Table 17.

Table 17

*Homogeneity of Variance – Pre-Test ROI Score Bands*

Levene's Test	$df1$	$df2$	Sig.
19.246	2	145	.000

The null hypothesis states there would be no statistically significant difference between the means of the three ROI bands. The critical value of  $F$  with 2 and 145 degrees of freedom for  $\alpha = .05$ , was  $F_{CV} = 3.07$ . The  $F$  statistic for the ANOVA ( $F = 4.203$ ) was greater than the critical

value of 3.07. Therefore the null hypothesis was rejected. The ANOVA indicates there was a statistically significant difference among the groups of underwriters for each scoring band but not necessarily among all three groups.

Table 18

*Analysis of Variance – ROI Score Bands*

Source	SS	df	MS	F	Sig.
Between Groups	6.429	2	3.215	4.203	.017
Within Groups	110.913	145	.765		
Total	117.342	147			

The comparisons between the three ROI scoring bands indicated that there was a statistically significant difference between the 50% - 59% band and the 70% - 99% band at the .05 level. However, the results were not statistically significant at the .05 level between the other pairs of bands. The difference between the other bands was large but not statistically significant using the Tukey HSD post hoc test. The results of the post hoc test are shown in Table 19.

Table 19

*Multiple Comparisons - ROI Score Bands – Tukey HSD*

Group 1	Group 2	MD	SE	Sig.
70% - 99%	60% - 69%	-.45287	.19465	.055
	50% - 59%	-.52856	.18794	.015*
60% - 69%	70% - 99%	.45287	.19465	.055
	50% - 59%	-.07569	.16386	.889
50% - 59%	70% - 99%	.52856	.18794	.015*
	60% - 69%	.07569	.16386	.889

\*The mean difference is significant at the .05 level.

The analysis of the three ROI score bands had failed to meet the requirements of the Levene's test for homogeneity of variance. As a result, harmonic means were applied to the analysis. Harmonic means are used in post hoc multiple comparison procedures when the group sizes are unequal. The harmonic mean of the group sizes was used. In this case, the harmonic mean sample size was 45.867. This causes there to be a threat of increased type I error levels or rejecting the null hypothesis when in fact it was true. Besides the Tukey HSD post hoc test, two other post hoc comparison tests were completed based on equal variances not assumed. They were the Dunnett T3 and Games-Howell multiple comparison tests. Both are pair-wise comparison tests for use with unequal variances. The results for both the Dunnett T3 and Games-Howell post hoc tests were very different in comparison with the Tukey HSD test. The results using the Dunnett test are shown in Table 20. The results of the Games-Howell post hoc test are shown in Table 21.

Table 20

*Multiple Comparisons - ROI Score Bands – Dunnett T3*

Group 1	Group 2	MD	SE	Sig.
70% - 99%	60% - 69%	-.45287	.12643	.002*
	50% - 59%	-.52856	.13697	.001*
60% - 69%	70% - 99%	.45287	.12643	.022*
	50% - 59%	-.07569	.18170	.966
50% - 59%	70% - 99%	.52856	.13697	.001*
	60% - 69%	.07569	.18170	.966

\*The mean difference is significant at the .05 level.

The comparisons between the three ROI scoring bands indicated this time that there was a statistically significant difference between the 50% - 59% band and the 70% - 99% band at the



.05 level and a statistically significant difference between the 60% - 69% band and the 70% - 99% band. There was still not a statistically significant difference between the 50% - 59% band and the 60% - 69% band. Note that the only difference between the two comparison tests was the significance of .966 for the Dunnett T3 post hoc comparison test versus .909 for the Games-Howell post hoc comparison test for the comparison between the 50% - 59% band and the 60% - 69% band. Dunnett's T3, designed for unequal variances and unequal sample sizes, is useful for its strict control over the alpha level of significance. Games-Howell, also designed for unequal variances and unequal sample sizes, can be liberal for sample sizes less than 5, should not be a problem with this study. And the fact that both were used and produced identical results for significant differences aided in validating the results of the post hoc comparisons. The results based on Games-Howell are shown in Table 21. Both of these post hoc comparison procedures, based on unequal variances, indicated that the ROI mean scores in both the 50% - 59% and 60% - 69% bands were statistically significantly different from the ROI mean score in the 70% - 99% band.

Table 21

*Multiple Comparisons – ROI Score Bands – Games-Howell*

Group 1	Group 2	MD	SE	Sig.
70% - 99%	60% - 69%	-.45287	.12643	.002*
	50% - 59%	-.52856	.13697	.001*
60% - 69%	70% - 99%	.45287	.12643	.022*
	50% - 59%	-.07569	.18170	.909
50% - 59%	70% - 99%	.52856	.13697	.001*
	60% - 69%	.07569	.18170	.909

\*The mean difference is significant at the .05 level.

## Effect of Training on the Financial Results of the Organization

The survey of the literature showed studies that also attempted to see if training intervention results and return on investment calculations from training could be tied in to the bottom line financial results of the organization. Post-hoc data was reviewed for this organization. There are two basic ways to influence the profitability of an insurance company. One way is to keep losses low, having a very positive effect on the combined ratio. Training can assist in this area. More knowledge can assist underwriters in avoiding loss-producing scenarios. The other means is to increase written premium to lower the average expense ratio. Training can also assist in this area. The expenses are what it takes to run an insurance company. Increased premium has the effects of lowering average fixed costs. Whether a company sells \$1 million in premium or \$10 million, there is still a fixed expense structure that must be dealt with just to keep the lights on in the company. This study sought to see if training is one of the variables that can affect the bottom line of this organization.

The generally accepted method of measuring the financial success of an insurance company is called the combined ratio (Tuckey, 2006). The benchmark of the combined ratio is 100%, considered to be the breakeven point, not including the results of financial investments. The components of an insurance company's combined ratio consisted of losses, expenses and profit. Added together, the loss ratio percentage and the expense ratio percentage, the percentage remaining is called an underwriting profit, or if greater than 100%, an underwriting loss. The loss ratio component was calculated by dividing incurred losses in a given calendar year by premium earned in that same period of time. The expense ratio component was calculated by dividing incurred expenses by earned premium for the same time period. Fundamentally, the lower the combined ratio is, the greater will be the profit for the organization. Numbers in the low 90-

percentile range are considered very good and few companies achieve these results over any prolonged period of time.

A comparison of the combined ratio results prior to training in 2000, 2001, 2002, and 2003 and the results after training in 2004, 2005, and 2006 were revealing. The combined ratio results in 2000, 2001, 2002 and 2003 were 100.6%, 100.1%, 97.9% and 93.2% respectively. The combined ratios during and after the training interventions were 91.3% in 2004, 91.9% in 2005, and 91.5% in 2006. A correlation analysis was done comparing the combined ratio results between the two time periods. The years 2004, 2005, and 2006 were set up as post-training years. The other years were not. The result was a correlation coefficient of .819, statistically significant at the .05 level.

These results suggested that there was a strong relationship between company results as measured by combined ratios and the training interventions. The above is an indication only. There are too many uncontrollable variables that have an effect on the loss side of the equation, as mentioned in the Limitations section of this study. There are many variables that can influence the financial outcome of an insurance company. The environment an insurance company plays in is not static. Rather it is dynamic and often there are changes that affect the variables that influence organizational profitability that are out of the control of the insurance company. There are government controlled rate changes, changes in insurance regulations, new entrants into the marketplace, growth rates of the customer base, changes in the business failure rate and general upturns and downturns in the economy. The organization itself can change through the loss of staff, an unforeseen increase in the frequency and severity of losses, and internal changes in growth and pricing strategies. Any of these factors can have a significant positive or negative effect on the bottom line of any insurance company depending on the strength of each variable.

A detailed analysis of the effects of these variables is far beyond the scope of this study and would require a separate lengthy analysis devoted to that subject. But it should be mentioned that many of these factors are beyond the control of a single insurance company. They are often beyond the control of the entire industry. Each company exists within the existing financial environment and must make as many of the right decisions as it can to sustain itself. The industry often follows up and down cycles over the last fifty years, but this should not negate the positive impact training can have on the bottom line results of an organization.

#### Effect of Training on Employee Attitudes

Did the formal training intervention programs conducted in 2004 and 2005 affect employee attitudes within the organization, or more specifically, within the underwriting department? Prior research studies (Carlson, Bozeman, Kacmar, Wright & McMahan, 2000, Tannenbaum, Mathieu, Salas, & Cannon-Bowers, 1991 and Williams & Hazer, 1986) indicated some degree of correlation between training and employee motivation and feelings of commitment to the organization. This company began conducting extensive employee attitudinal surveys beginning in 2004. So can an extensive training program affect the attitudes of employees who participate in the training? Three categories of questions were chosen out of the eight categories that were administered to employees. The three categories consisted of employee commitment, operational efficiency, and overall satisfaction. The results, while anonymous, were broken out by department and provided to each department head for discussion with the staff. The following results were for the underwriting department only, representing the individuals who participated in the intervention. These results represented three years of data.

The first category selected for the comparison was “employment commitment.” The scores for the participants in this category were 70% in 2004, 80% in 2005, and 88% in 2006.

The second category selected for comparison was in the “operational efficiency” section. The scores in this category were 60% in 2004, 70% in 2005, and 82% in 2006. The last category selected for comparison is in the “overall” section. Only one statement was selected from this category. It said: “Overall I am satisfied working for the company.” This statement is a critical one and is an important measure of employee attitudes. The question is used by the organization as an overall benchmarking question in evaluating employee satisfaction for the prior annual reporting period. The scores were 76% in 2004, 82% in 2005 and 92% in 2000. These results seem to indicate an increasing satisfaction with the organization while working within the underwriting department over the last three years. Employee turnover was low enough as to be considered as having a negligible effect on these results.

Each year that training was done and in the last year following the impact of training, the scores continued to increase. While it is difficult to determine what portion of this section could be attributed to training, given the results of the above three sections, it was reasonable to believe that training did have a positive impact on employee attitudes within the organization. During this time period, there were no other human resource interventions that occurred within the underwriting department. And the employee attitudes between the underwriting group and the rest of the organization were very different, with the underwriting group being much higher. These responses were also much higher than the responses that were provided by the rest of the organization. The results do suggest that training may have had a positive role in changing employee attitudes.

#### Summary

The results seem to demonstrate the value of the proposed method of calculating return on investment without having to depend solely on the opinions of participants. Based on the use

of pre-tests, it may be possible to predict which participants can provide the organization with the highest rate of return. This would allow organizations to lower their training costs while increasing still maximizing the amount of learning that would be taking place. This process can also have a positive effect on the employees. By providing training to those who can benefit from it the most, there may be an increase in satisfaction not just with the job but with the organization as well.

## CHAPTER 5

### DISCUSSION, CONCLUSIONS AND RECOMMENDATIONS

#### Overview

This study began as an attempt to find a process to evaluate training interventions that would have greater credibility with the senior managers in organizations. The survey of the literature revealed the key issue that needed to be addressed was an objective method that would make sense to the senior level decision-makers who controlled the purse strings. Oftentimes, senior management would be given incredulous results of over a thousand percent return on investment, especially so because so much of it was based on the opinions of employees who participated in the training interventions. What employee would state that the training did not make a difference? And as the American Society for Training and Development surveys showed, there was seldom any follow-up to determine if any actual learning had really taken place, much less was that new knowledge being used on the job-site. And as the credibility of the value of training interventions came into question, so did the budgetary restrictions at times when quality training could have returned the greatest value to the organization. Understanding how senior decision-makers approach resource allocation issues was critical, especially when it comes to costly training intervention proposals.

The results seem to demonstrate the value of the proposed method of calculating return on investment (ROI) without having to depend solely on Level 1 opinions about how much value has been added to the job as a result of participating in the training. Level 1 data was not collected in this methodology. The results demonstrated by Kaufman (1994), Phillips (1997), Kirkpatrick (1998), and others all stressed the importance of measuring the value of training far beyond how much the participants enjoyed it. They all focused on determining the value of

training to the bottom line of the organization that sponsored the training. Other studies have been discussed that suggest no correlation between Level 1 data and actual learning. But the studies by Kaufman, Phillips and Kirkpatrick all included in some fashion the opinions of the participants about the effect of the training on their abilities to do their jobs and this may have heavily influenced the results of the return on investment calculations. The credibility of training intervention may become more credible when participant opinions are not included.

The process put forth in this research stressed a number of key concepts. First, there is a limit as to how much of a return on investment training activities can produce. The process recognized that by the insertion of the job relevance factor. No amount of training provided a participant could ever exceed 100% of the skills needed in any given job. By applying that percentage to the increased knowledge attained from training, the impact is reduced by the value of the job relevance factor. Second, this was applied to each participant's salary so that the dollar amount of training added value was anchored to a real number, salary, which could be compared to the other real number, the expenses incurred in delivering the training. For consistency, the midpoint of each salary grade was chosen rather than the actual salary itself. This advances the concept of component pay as recommended by Campbell (1994). Campbell defined the value of training as the total productivity gain times the component pay. Component pay was defined as the total annual compensation package times the percentage of time the employee spends doing what the training was for. The research methodology improves on the component pay concept by using the job relevancy factor to forgo the problem of having to determine Campbell's percentage of time requirement.

Quality training can be an expensive proposition. Devoting scarce resources to training must be justified in today's fast paced competitive world. Using an effective method of



allocating resources within the training budget can be a significant competitive advantage. As Porter (1985) stressed, creating and sustaining competitive advantage is critical if any organization is to survive and prosper. One of the key business strategies is effective training but the organization can only succeed tactically with such a strategy if it maximizes its return on investment from training.

Part of the benefits to well designed training is the effect it can have on an employee's self-efficacy, motivation and commitment to the organization. The results found by Harter (2002) in a meta-analysis of other organizational studies of employee satisfaction and business outcomes, and by Silver, Mitchell and Gist (1995) in relating employee self-efficacy to employee success can all lead to a significant competitive advantage if quality training not only provides an increased return on training investment but also increased employee motivation as well.

The proposed process allows for forecasting the financial benefits of any training intervention proposal. The ability to forecast can become a mainstay of how decisions are made to determine funding for a variety of training interventions. The use of a financial benefits model is supported by Swanson's (1998) position that the process of training evaluation must be similar to how other types of decisions are made within the organization. Swanson proposed a process called performance value analysis and it was the critical component of his forecasting financial benefits model. By stressing the financial benefits of a training intervention, the proponents of training can reach a meeting of the minds with those who serve as the decision-makers in the organization. Financial analysis models are used for everything else within an organization's decision-making process. By following the same process, it becomes easier for all parties to reach a meeting of the minds by sharing a common understanding of how decisions are made in the process of allocating resources.

The process should be used to estimate the financial benefits of the intervention prior to its approval. The use of the pre-test process and the results of the fifth null hypothesis demonstrate how this can be done. Pre-set required rates of return can provide an objective method of making decisions. Establishing a process of evaluating training that can be understood and embraced by senior level decision-makers can eliminate the irrelevance these decision-makers often feel towards the training intervention process. By working together with the financial decision-makers within an organization, training professionals can manage not only their own resources but can also improve the likelihood of getting other needed interventions approved by objectively demonstrating the value of those interventions to the financial bottom line of the organization.

Working together with key decision makers was supported by Nickols' (2005) proposal of a stakeholder-based approach to evaluating training. This approach focused on evaluating outcomes in ways that provide meaning, value and relevance to all stakeholders. In this case, stakeholders are those individuals without whose support any proposed intervention could not have succeeded. Referring back to Swanson's (1998) two prevalent views of human resource interventions: an optional activity that is nice to do; or a waste of business resources where costs often exceed benefits, a training intervention would be less likely to be viewed as irrelevant if the evaluation process were to be based on an accepted financial analysis model. This is even more so if it is a process that can also be used to predict the benefits of training intervention before they occur.

The ability to forecast the benefits of a training intervention prior to its inception cannot be underestimated. Business organizations view training as having four possible alternatives (Swanson & Gradous, 1988). First, training can be considered a major business process that must

be conducted if the business is to succeed. Second, training can be considered as adding value and is worth doing. Third, training is optional and would be nice to do. Fourth, training is a waste of business resources and the costs far outweigh the benefits. The latter two alternatives form the prevalent opinion. What is needed is a process that supports the first two alternatives. Businesses today find their budgets under close scrutiny as they try to improve bottom line profitability. This is especially true for publicly traded companies trying to maximize shareholder value. Asking for more money for training, and sometimes just getting what was allocated in the prior year, can be a struggle under the pressures to remain competitive. Finding a way to show those who control the purse strings that there can be a financial return that meets the organization's internal rate of return requirements can pave the way for getting approval for training intervention proposals.

#### Summary and Discussion of Findings

The purpose of this study was to develop an objective method of evaluating training that can be incorporated into the decision-making process of an organization that does not rely on just asking the participants their opinion of the value of the training. Part of the problem is traditional ROI computations that use participant surveys produce such high results they raise questions about the credibility of those results. Asking for more money for training is often received with skepticism. As previously cited, these surveys are suspect as to their validity. One study found a significant correlation between participant enjoyment of the training and the ratings given the intervention (Dixon, 1990) and this makes sense. People who sit through hours and days of training are more likely to rate it high if they found the experience enjoyable. And as Dixon suggested, enjoyment need not necessarily mean knowledge gain. This study proposes a method for evaluating training that can not only be used to objectively evaluate training as an investment and estimate financial returns prior to the training, but is easy to explain to chief financial

officers and administer as well. It allows any organization to use this method to set a minimum rate of return requirement for the funding of any intervention.

This study reviewed the literature on training evaluation and proposed a new process by which training could be evaluated. It applied a seven-step formula to create a series of individual return on investment scores for employees who took part in three training interventions. There were six proposed null hypotheses. The first null hypothesis failed to be rejected while the remaining five null hypotheses were all rejected. Data was also collected on the bottom line financial results produced by those individuals who participated in the training intervention. The issues related to doing a bottom-line analysis are detailed in the limitations section of this study. Finally, data was obtained from an annual survey of company employees, specifically for the participants in the training intervention. The results are anonymous except for the departments the employees are in and participation rates were high.

*Research Question #1: Can an objective method for evaluating training be developed?*

The results suggest that the formula for evaluating training interventions without the use of participant surveys is feasible. Participant surveys tend to focus around enjoyment of the training rather than the amount of meaningful learning attained and as such should not be used as the sole means of evaluation. This was the conclusion found on the use of participant surveys by Dixon (1990), Lott (1975), and others. The results of this study seem reasonable, especially based on the application of the Job Relevance Factor. The analysis of variance based on the results of the three training interventions produced results that were not significantly different from each other, suggesting that the underwriters benefited at about the same level for all three training interventions. From the user's perspective, the return on investment results, shown in Table 22, were reliably repeated within the three training interventions. The skeptical quadruple

digit returns found in other studies that employed participant surveys is replaced by a method that does not rely on participant surveys and produced more plausible results.

Table 22

*Results of Return on Investment Calculations*

Subject Intervention	Return on Investment
Commercial Underwriting	153%
Commercial Underwriting – Basic	97%
Commercial Underwriting – Property	164%
Commercial Underwriting – Casualty	181%
Underwriting Assistants	496%

*Null Hypothesis #1: There will be no statistically significant difference in ROI results of an underwriter-training program consisting of three different training interventions across a sample of underwriters.*

The analysis of variance completed based on the three sets of return on investment scores obtained from the three training interventions showed no statistically significant difference between the mean scores of the three training interventions. The critical value of  $F$  with 2 and 316 degrees of freedom for  $\alpha = .05$ , was  $F_{CV} = 3.0$  but the  $F$  statistic for the ANOVA ( $F = 2.652$ ) was not greater than that critical value. Therefore the null hypothesis could not be rejected. There was not a statistically significant enough difference between the mean Return on Investment scores of the three underwriter training interventions amongst the underwriters. The underwriters benefited at relatively similar levels for all three phases of their underwriter training intervention. An return on investment of 153% for the underwriter-training program was readily accepted by the organization. This result supported the first research question where not only was the

proposed methodology producing reasonable results but they were repeated for the three training interventions.

The three underwriting interventions did not produce any statistically significant differences in an analysis of the means of the results. Training in different types of underwriting skills did not produce significantly different results for the same group of commercial underwriters. Had there been a statistically significant difference, there would have been further analysis of the reasons for the differences.

*Null Hypothesis #2: There will be no statistically significant difference in ROI results of an underwriter-training program between a group of underwriters and a group of non-underwriters.*

The  $t$  test results demonstrated that the average means for return on investment, even with equal variances not assumed, were significantly different statistically between the underwriting group and the non-underwriting group. The results were  $t(485) = 6.425, p < .05$ , two tailed. The results of the  $t$  test allowed the null hypothesis to be rejected. The results between the two groups, underwriters and non-underwriters, were statistically significantly different. Given the effects of the job relevancy factor, and the fact that underwriter training should have had more meaning for underwriters than non-underwriters, it does not seem surprising that the means were statistically significantly different. This result supported the first research question where not only did the proposed method produce reasonable results but that the results were statistically significantly different when compared between underwriters and non-underwriters. However. The results were not statistically significantly different when compared for the three different interventions that were applied to the underwriters only, adding to the validity of the evaluation process.

*Null Hypothesis #3: There will be no statistically significant difference in ROI results of an underwriter-training program between a group of assistant underwriters and a group of other employees of the same grade.*

The  $t$  test results demonstrated that the average means for return on investment, even with equal variances not assumed, were statistically significantly different between the group of underwriting assistants and the group of other employees of the same grade. The results were  $t(32.49) = 3.126, p < .05$ , two tailed. The results of the  $t$  test allowed the null hypothesis to be rejected. There was a statistically significant difference between the mean ROI scores of the underwriting assistants compared to the mean ROI scores of the non-underwriting assistants. These results were expected for the same reasons that explained the difference between the results found for the underwriters and non-underwriters. This again supported the first research question. It added to the validity of the methodology and in addition added to its reliability as well.

*Null Hypothesis #4: There will be no statistically significant difference in ROI results of a training program between underwriters with less than 5 years of experience and underwriters with 5 or more years of experience.*

The  $t$  test produced a value of 2.394 with a significance level of .05,  $t(68.42) = 2.394, p < .05$ , two tailed. In this intervention there was a statistically significant difference between the mean scores of the underwriters with less than 5 years of experience compared to the mean scores of the underwriters with 5 or more years of experience. This also makes sense as underwriters with less than 5 years of experience in all probability have more to learn in a training intervention than individuals with 5 or more years of experience. In addition, the individuals in this latter group averaged over 20 years in experience, compared to the 3 years for

the other group. The experienced group has a lot less to learn compared to the group with less experience. The less experienced group produced an average ROI score of 248%. The more experienced group produced an average ROI score of 132%. The *t* test shows that these mean scores were statistically significantly different. The rejection of this null hypothesis suggests that there were factors that could be used to predict when a training intervention could produce pre-established required rates of return. In this case the factor was years of experience but there could be many others.

The data in Table 15 also supports the second research question. This shows that the higher the pre-test score, the lower the average return on investment. For pre-test scores between 50% and 59%, the average ROI score was 58.5%. For pre-test scores between 60% and 69%, the average ROI score dropped to 50.9%. For pre-test scores 70% and higher, the average ROI score dropped sharply to 5.6%. These results suggested that an organization could specify a minimum pre-test score for eligibility to participate in the training intervention, thereby maximizing their rate of return. This would also result in a reduction of training costs.

The results of this null hypothesis support Swanson's (1998) research that the evaluation process could be used as a forecasting tool of when training interventions should be undertaken. Swanson applied the benefit-forecasting model to all manner of training interventions, including customer service, project management and coaching. The process proposed by this research follows that same line of research, but with specific changes to the measurement process using salaries and job factor analysis. The results found for the fourth null hypothesis were not surprising since less experienced individuals have more to learn from a standardized training intervention than more experienced individuals. The less experienced individuals would more than likely also have a lower average salary base. Trainers can use this process to meet minimum



established rates of return before any training intervention is undertaken. In this study, pre-test scores, if they were too high, would support the premise that the potential for a high rate of return would probably be less than for other alternatives that showed low pre-test scores. In this case, underwriters with less experience had more to gain from the training than underwriters with more experience and produced a greater return.

*Research Question #2: Can a pre-test be used as a predictive tool in the selection process of participants based on a future estimated ROI?*

This research question supports the broader issue of determining if training resources can be allocated based on a minimum expected rate of return. The organization can specify a necessary rate of return as a criterion for funding any project. This research question follows up on Swanson's (1998) position that the training evaluation process should be used as a forecasting tool of when training interventions should be undertaken. The results of the first three null hypotheses provided a methodology upon which the goals of forecasting and resource allocation based on pre-set rates of return can be established.

For the corporate trainer, the idea of only training those who truly need it not only can maximize the benefit of the training, but can also provide the largest opportunity for the transfer of knowledge. It also addresses the issue of lost opportunity costs of taking participants off the line and losing valuable production time by not taking individuals who have demonstrated through this process that they don't need the training. This ability to forecast in advance of the intervention was a fundamental premise of Swanson and Gradous (1988). Their three component benefit-forecasting model was designed to give trainers the ability to forecast the anticipated returns from training. The three components were the performance value that would result from the program, the cost of the program, and the benefit that would result from the program. The

methodology being proposed here provides a detailed method of evaluating the results of the intervention, using salaries and job factors, and a method of forecasting, specifically using the financial value of pre-tests.

*Null Hypothesis #5: There will be no statistically significant difference in ROI results of an underwriter-training program within ROI score bands.*

The results of the analysis of variance allowed the null hypothesis to be rejected. There was a statistically significant difference between the mean ROI scores of the 70% - 99% band and the mean ROI scores of the other two bands, 50% - 59% and 60% - 69%, depending on which post hoc comparison test was selected. While the Tukey post hoc comparison test showed only the band pair consisting of 50% - 59% and 70% - 99% was statistically significant, both the Dunnett T3 and the Games-Howell post hoc comparison tests showed two band pairs were statistically significant, 50% - 59% and 70% - 99% as well as 60% - 69% and 70% - 99%. These results support the research question about the predictability of pre-test scores. Pre-test scores could be used to predict return on investment success. Higher ROI score bands suggested that some pre-scores participants could have had their participation waived. It is this kind of predictability adds value of the training evaluation process. It may also allow for decisions to be made without necessarily having to spend the resources first and then finding that those spent resources never provided the anticipated rate of return. For the corporate trainer, this means training only those individuals who can truly benefit from the training.

#### Limitation of the Results

As in any research study, participation is always a concern. One of the issues in this study was not all underwriting participants took part in each training intervention session because of other work related commitments. Underwriters often have external presentations that must be

done in a timely manner. In addition, there were two underwriters who left the organization. There were 319 ROI scores collected. The maximum number of scores that could have been collected had all underwriters attended every intervention was 520, resulting in a 61% participation rate.

Because of the differences in sample sizes, many of the comparisons failed the test of homogeneity of variance. All of the  $t$  tests were based on an inequality of variance. In addition, the ANOVA post hoc comparison tests had to be based on an inequality of variance amongst the groups. These sample size differences occurred because they were in place training interventions that did not afford the more desirable scenario of matched pairs from different subject groups. Neither was there an opportunity for a control group.

#### Implications for Organizations

This study provides a roadmap for all organizations that conduct training leading to a more effective method for deciding which training interventions to pursue and who should be included. It provides an objective standard since there is no reliance on participants' opinions about the value of their training experience. By connecting the value of training to the salaries of the participants, the process provides a base that is centered on real dollars. Focusing on salary as a criterion was fundamental for Campbell (1994). The seven-step calculation process takes those real dollars and merges them with the amount of learning that took place and converts the data into individual return on investment scores. This process is one that senior level managers can relate to, going forward only with those projects whose rates of return justify incurring the expenses in the first place. Attempting to convince decision-makers on the necessity of a training project becomes easier for both the proposal maker and the decision-maker when it is based on concepts of financial return.

Organizations have projects that must be done. Every department has their own projects that need to be funded. Training is just one issue amongst many. Senior-level decision-makers will find it easier to decide which interventions to approve when they are all proposed in the same financial return format. They can even institute as a control point, a minimum rate of return. This process can be used to determine not only if a training intervention should be authorized but also which potential participants can gain the most from it, thereby maximizing the return on investment to the organization. The end result is better quality training while at the same time reducing or eliminating unnecessary expenses to the organization.

For the individual corporate trainer, this is a method that can be tailored to the needs of the individual organization based on the choice of job relevance factors. It means that an organization must understand all of the knowledge needed for a particular job and be able to determine the importance of the particular skill sets that make up each job. Values must then be assigned to the jobs themselves. Swanson (2001) provides an organized approach to accomplishing this across a wide range of jobs, such as service and manufacturing industries, as well as soft skill positions.

#### Recommendations for Practice

Further studies should be done in other organizations to determine the value of this process for measuring return on investment to validate and verify its widespread applicability and acceptability. Further research should be conducted for both technical skills as well as soft skills. More knowledge needs to be gained to determine what are the most significant contributors to any organization's profitability and the role that training plays. And more research should be conducted by each organization to determine the relationship between training and the effect training has on the attitudes of their employees. It is very possible for any

given organization that not only would training increase organizational profitability but it may also enhance employee attitudes and that can further increase profitability by increasing employee motivation while at the same time lowering absenteeism and employee turnover.

What can an organization do now?

1. *Evaluate each job classification within the organization to determine what kinds of training interventions are needed for total proficiency;*
2. *Determine the job relevancy factor that is to be applied to each identified training intervention;*
3. *Establish a pre-test format for each training intervention;*
4. *Investigate each training intervention instrument for validity and reliability;*
5. *Establish a long-term continuous process to analyze the skill set needs of individual employees using the pre-test concept organization-wide;*

Organizations that choose to implement this process can do so in manageable increments to make the change process less painful. This process of making the training decision-making process more efficient can be applied company-wide or to a single training intervention. While simple in concept, it has features that can be appreciated by any financial manager who has responsibility for approving funds for training interventions.

The process starts with a quality analysis of each job to which this process would be applied. Corporate trainers are referred to the work of Swanson (2007) for assistance in the evaluation of job skills and diagnosing the needs of the organization relative to their family of jobs. Significantly greater returns on investment may be attained when training interventions are subjected to a more disciplined methodology. This proposed model provides the corporate trainer with an easy to understand framework that can applied to any organization for a single job or for

the entire family of jobs. The predictive portion of this model can be used to maximize the time of all of the employees by providing training to only those employees who need it. This is a fundamental cornerstone of this method in providing the biggest bang for each training buck.

### Conclusion

This study produced three significant results. First, the proposed methodology based on the use of participant salary, as advanced by Campbell (1994), for evaluating training interventions produced credible results. Second, it is possible to use pre-tests to forecast future return on investment, a result viewed as vital by Swanson (2001) in his research on financial benefits forecasting. Third, the pre-tests can be used as a selection process based on pre-established minimum rates of return. In addition, the study provided an indication that training may have had a significant influence on employee attitudes.

The results of the ANOVA and *t* tests added validity to the proposed methodology by recognizing statistically significant differences in the results of different groups' ROI scores allowing the appropriate null hypotheses to be rejected. There were also very positive correlations between training being provided and the financial results of the organization. Business organizations interested in establishing an ROI process to prioritize the application of training resources should give serious consideration to using a method that uses the participant's salary as the foundation upon which to establish the basis for the financial return to the organization. This method can then be used as a way to prioritize the requests for funds to pursue training agendas. By doing so, training can become recognized as a function that supports the bottom line profitability of each and every organization.

## BIBLIOGRAPHY

- Basarab, D. I., & Root, D. K. (1992). *The training evaluation process: A practical approach to evaluating corporate training programs*. Boston, MA: Kluwer Academic Publishers.
- Bassi, L. J., & Van Buren, M.E. (1999). *Sharpening the leading edge*. Alexandria, VA: American Society for Training and Development.
- Bateman, T.S., & Strasser, S. (1984). A longitudinal analysis of the antecedents of organizational commitment. *Academy of Management Journal*, 27 (1), 95-112.
- Behnkoff, B. (1997). Ignoring commitment is costly. *Human Relations*, 50 (6), 701-726.
- Brinkerhoff, R. O. (1987). *Achieving results from training*. San Francisco, CA: Jossey-Bass.
- Brinkerhoff, R. O. (1998). Clarifying and directing impact evaluation. In Brown, C. J. (Ed.), *Evaluating corporate training: Models and issues*. Boston, MA: Kluwer Academic Publishers.
- Campbell, C. P. (1993). A primer on determining the cost-effectiveness of training – part 1. *Industrial and Commercial Training*, 26 (11), 32-38.
- Campbell, C. P. (1994). A primer on determining the cost-effectiveness of training – part 2. *Industrial and Commercial Training*, 27 (1), 17-26.
- Carlson, D.S., Bozeman, D.P., & Kacmar, M.K. (2000). Training motivation in organizations: An analysis of individual-level antecedents. *Journal of Managerial Issues*, 12 (3), 271-288.
- Carnevale, A.P., & Schulz, E.R. (1990). Return on investment: Accounting for training. *Training and Development Journal*, 44 (7), 1-32.
- Cascio, W. F. (1989). Using utility analysis to assess training outcomes. In I. L. Goldstein (Ed.), *Training and development in organizations*. San Francisco, CA: Jossey-Bass.
- Clark, C.S., Dobbins, G.H., & Ladd, R.T. (1995). Exploratory field study of training motivation: Influence of involvement, credibility, and transfer climate. *Group and Organizational Management*, 18(3), 292-307.
- Cullen, J.G., Sawzin, S.A., Sisson, G.R., & Swanson, R.A. (1978). Cost effectiveness: A model for assessing the training investment. *Training and Development Journal*, 32 (1), 24-29.
- Dionne, P. (1996). The evaluation of training issues. *Human Resources Development Quarterly*, 7 (3), 279.

- Dixon, N.M. (1994). Training: Meet the training's goals without reaction forms. *Personnel Journal*, 73 (9), 1-3.
- Dixon, N.M. (1990). The relationship between trainee responses in participant reaction forms and posttest scores. *Human Resource Development Quarterly*, 1 (2), 129-137.
- Dixon, N. M. (1990). *Evaluation: A tool for improving HRD quality*. San Diego, CA: University Associates.
- Farrell, D., & Rusbult, C.E. (1981). Exchange variables as predictors of job satisfaction, job commitment and turnover. *Organizational Behavior and Human Performance*, 28 (1), 78-95.
- Fitz-enz, J. (2000). *The ROI of human capital: Measuring the economic value of employee performance*. New York: American Management Association.
- Ford, D. J. (1994). Three Rs in the workplace. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Gist, M.E. (1987). Self-efficacy implications for organizational behavior and human resources management. *Academy of Management Review*, 12 (3), 472-485.
- Gist, M.E., Stevens, C.K., & Bavetta, A.G. (1991). Effects of self-efficacy and post-training intervention. *Personnel Psychology*, 44 (4), 837-861.
- Graham, M., Bishop, K., & Birdsong, R. (1994). Self-directed work teams. In J. J. Philipps (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Harter, J. K., Schmidt, F.L., & Hayes, T.L. (2002). Business unit level relationship between employee satisfaction, employee engagement, and business outcomes. *Journal of Applied Psychology*, 87 (2), 268-280.
- Hicks, W.D., & Klimoski, R.J. (1987). Entry into training programs and its effects on training outcomes. *Academy of Management Journal*, 30 (3), 542-552.
- Hill, D. P., & Phillips, J. J. (1997). Preventing sexual harassment. In J. J. Phillips (Ed.), *Measuring return on investment, volume 2*. Alexandria, VA: American Society for Training and Development.
- Hodges, T. M. (1997). Computer-based training for maintenance employees. In J. J. Phillips (Ed.), *Measuring return on investment, volume 2*. Alexandria, VA: American Society for Training and Development.
- Holton, E. F. (1996). The flawed four-level evaluation model. *Human Resource Development Quarterly*, 7 (1). 5-25.



- Holton, E. F., & Naquin, S. (2005). A critical analysis of HRD evaluation models from a decision-making perspective. *Human Resource Development Quarterly*, 16 (2), 257-280.
- Kaufman, P. K., (1994). Levels of evaluation: Beyond Kirkpatrick. *HRD Quarterly*, 5 (4), 371-380.
- Kaufman, R., & Watkins, R. (1996). Cost consequences analysis. *Human Resources Development Quarterly* 7 (1), 21-87.
- Kearsley, G., & Compton, T. (1981). Assessing costs, benefits and productivity in training systems. *Training and Development Journal*, 35 (1), 52-61.
- Kirkpatrick, D. L. (1998). *Evaluating training programs: The four levels*. San Francisco, CA: Berrett-Koehler Publishers.
- Korth, S.J. (2001). Consolidating needs assessment and evaluation: Saving time and money. *Performance Improvement*, 40 (1), 38-43.
- Lachnit, C. (2001). Training proves its worth. *Workforce*, 52-59.
- Lewis, T. (1996). A model for thinking about the evaluation of training. *Performance Improvement Quarterly*, 9 (1), 3-22.
- Long, L.N. (2004). Analytics: The CLO's case in the CEO's language. *Chief Learning Officer*, 3 (9), 1-4.
- Lott, O. C. (1975). Evaluating to reduce training costs. In D. L. Kirkpatrick (Ed.), *Evaluating training programs* (pp. 244-246). Madison, WI: American Society for Training and Development.
- Lyau, N. M., & Pucel, D.J. (1995). Economic return on training investment at the organization level. *Performance Improvement Quarterly*, 8 (3), 68-79.
- Mattson, B.W. (2005). Using the critical outcome technique to demonstrate financial and organizational performance results. *Advances in Developing Human Resources*, 7(1), 102-120.
- McMurrer, D. P., Van Buren, M.E., & Woodwell Jr., W.H. (2000). *The 2000 ASTD state of the industry report*. Alexandria, VA: American Society for Training and Development.
- Miller, V. A. (1996). The history of training. In R. L. Craig (Ed.), *The ASTD training and development handbook* . New York: McGraw-Hill.
- Mirabile, R., Caldwell, D., & O'Reilly, C. (1986). Designing and linking human resource programs. *Training and Development Journal*, 40 (9), 60-66.

- Moller, L., & Mallin, P. (1996). Evaluation practices of instructional designers and organizational support or barriers. *Performance Improvement Quarterly*, 9 (4), 82-92.
- Moller, L., Benscoter, B., & Rohrer-Murphy, L. (2000). Utilizing performance technology to improve evaluative practices of instructional designers. *Performance Improvement Quarterly*, 13 (1), 84-95.
- Montier, R., Alai, D., & Kramer, D. (2006). Measuring and evaluating competing models. *T+D* 60 (7), 47-51.
- Mosier, N.R. (1990). Financial analysis, the methods and their applications to employee training. *Human Resource Development Quarterly*, 1 (1), 45-63.
- Muchinsky, P.M. (2000). Emotions in the workplace: The neglect of organizational behavior. *Journal of Organizational Behavior*, 53 (3), 787-790.
- Muir, M., Watkins, R., Kaufman, R., & Leigh, D. (1998). Cost-consequences analysis: A primer. *Performance Improvement*, 37 (4), 8-17.
- National HRD Executive Survey. (1997). *Measurement and evaluation*. Washington, D.C.: Society for Human Resource Managers.
- Nickols, F.W. (2005). Why a stakeholder approach to evaluating training. *Advances in Developing Human Resources*, 7 (1), 121-136.
- Nolan, M. (1996). Job training. In R. L. Craig (Ed.), *The ASTD training and development handbook*. New York: McGraw-Hill.
- Parry, S. B. (1994). Using action plans to measure return on investment. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Parry, S. B. (1997). *Evaluating the impact of training*. Alexandria, VA: American Society for Training and Development.
- Payne, R. (1994). Improving customer service skills. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Phillips, J. J. (1994). *Measuring ROI in an established program*. Alexandria, VA: American Society for Training and Development.
- Phillips, J. J. (1997). *Handbook of training evaluation and measurement methods*. Houston, TX: Gulf Publishing Company.

- Phillips, J. J. (1997). Evaluating leadership training for newly appointed people managers. In J. J. Phillips (Ed.), *Measuring return on investment, volume 2*. Alexandria, VA: American Society for Training and Development.
- Phillips, J. J. (1997). *Return on investment*. Houston, TX: Gulf Publishing.
- Porter, M. E. (1985). *Competitive advantage: Creating and sustaining superior performance*. New York: Free Press.
- Purcell, A. (2000). 20/20 ROI. *Training and Development, 54* (7), 28-33.
- Quinones, M.A. (1995). Pretraining context effects. *Journal of Applied Psychology, 80* (2), 226-238.
- Riley, T., Davani, H., Chason, P., Findley, K., & Druyor, D. (2003). Innovative evaluation: Creating speed and efficiency in your level 3 evaluation process. *Educational Technology, 43* (1), 38-45.
- Rohrer-Murphy, L., Moller, L., & Benscoter, B. (1997). A performance technology approach to improving evaluation. *Performance Improvement Journal, 36* (8), 10-15.
- Rossett, A. (1989). Assess for success. *T & D Journal 43* (5), 54-60.
- Russ-Eft, D., Krishnamurthi, S., & Ravishankar, L. (1994). Getting results with interpersonal skills training. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Saks, A.M., & Cronshaw, S.F. (1990). A process investigation of realistic job previews: Mediating variables and channels of communication. *Journal of Organizational Behavior, 11* (3), 221-236.
- Schael, T. (1996). *Workflow management systems for processing organizations*. New York: Springer-Verlag.
- Schneider, H., Monetta, D.J., & Wright, C.C. (1992). Training function accountability: How to really measure return on investment. *Performance and Instruction, 31* (3), 12-17.
- Schneider, H., & Wright, C.C. (1990). Return on training investment: Hard measures for soft subjects. *Performance and Instruction, 29* (2), 28-35.
- Schoeppel, C. (1994). Turning down manager turnover. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.
- Shandler, D. (1996). *Reengineering the training function: How to align training with the new corporate agenda*. Delray Beach, Florida: St. Lucie Press.

- Sheldon, M.E. (1971). Investments and involvements as mechanisms producing commitment to the organization. *Administrative Sciences Quarterly*, 16 (2), 143-151.
- Shelton, S., & Alliger, G. (1993). Who's afraid of level 4 evaluation? *Training & Development*, 47 (6), 43-46.
- Silver, W.S., Mitchell, T.R., & Gist, M.E. (1995). Responses to successful and unsuccessful performances: The moderating effect of self-efficacy on the relationship between performance and attributions. *Organizational Behavior and Human Decision Processes*, 62 (3), 286-299.
- Sleezer, C. M., & Gradous, D.B. (1998). Measurement challenges in evaluation and performance improvement. *Performance Improvement Quarterly*, 11 (4), 62 - 75.
- Spencer, L.M. (1984). How to calculate the costs and benefits of an HRD program. *Training*, 21 (7), 40-51.
- Steinmetz, C. S. (1942). The history of training. In Dodd, J.O. (Ed.), *How to train workers for war industries*. New York: McGraw-Hill Book Company.
- Stevens, C.K., & Gist, M.E. (1997). Effects of self-efficacy and goal orientation training. *Personnel Psychology*, 50 (4), 955-978.
- Stevens, J.M., Beyer, J.M., & Trice, H.M. (1978). Assessing personal, role, and organizational predictors of managerial commitment. *Academy of Management Journal*, 21 (3), 380-396.
- Stolovitch, H. D., & Maurice, J. G. (1998). Calculating the return on investment in training: A critical analysis and a case study. *Performance Improvement*, 37 (8), 9-20.
- Stone, R. D., Steele, D. R., Wallace, D. M., & Spurrier, W. B. (1997). Training's contribution to a major change initiative. In J. J. Phillips (Ed.), *Measuring return on investment, volume 2*. Alexandria, VA: American Society for Training and Development.
- Sugrue, B., & Kim, K.H. (2004). *State of the industry report 2004*. Alexandria, VA: American Society for Training and Development.
- Sugrue, B. (2003). *State of the industry report 2003*. Alexandria, VA: American Society for Training and Development.
- Sugrue, B. & Rivera, R.J. (2005). *State of the industry report 2005*. Alexandria, VA: American Society for Training and Development.
- Sussman, D. (2006). Daimler Chrysler quantifies training quality. *T+D*, 60 (5), 45-46.

- Swanson, R. A. (2007). *Analysis for improving performance: tools for diagnosing organizations and documenting workplace expertise*. San Francisco: Berrett-Koehler Publishers.
- Swanson, R. A. (2001). *Assessing the financial benefits of human resource development*. Cambridge, MA: Perseus Publishing.
- Swanson, R. A. (1998). Demonstrating the financial benefit of human resource development: Status and update on the theory and practice. *Human Resource Development Quarterly*, 9 (3), 285-296.
- Swanson, R. A. (1995). Human resource development: Performance is the key. *Human Resource Development Quarterly*, 6 (2), 207-213.
- Swanson, R. A. (1993). Business partner: An HRD mandate. *Human Resource Development Quarterly*, 4 (2), 121-123.
- Swanson, R. A., & Gradous, D. B. (1988). *Forecasting financial benefits of human resource development*. San Francisco, CA: Jossey-Bass.
- Tannenbaum, S.I., Mathieu, J.E., Salas, E., & Cannon-Bowers, J.A. (1991). Meeting trainees expectations: The influence of training fulfillment on the development and commitment, self-efficacy and motivation. *Journal of Applied Psychology*, 76 (6), 759-769.
- Tesoro, F. (1998). Implementing an ROI measurement process at Dell Computer. *Performance Improvement Quarterly*, 11 (4), 103-114.
- Tracey, J., Hinkin, T., Tannenbaum, S.I., & Mathieu, J. (2001). The influence of individual characteristics and the work environment on varying levels of training outcomes. *Human Resources Development Quarterly*, 12 (1), 5-23.
- Tuckey, S. (2006). Combined ratio falls but red flags keep P-C insurance carriers on edge. *National Underwriter*, 110 (38), 8 & 33.
- Twitchell, S., Holton, E. F. III, & Trott, J. W. Jr. (2000). Technical training evaluation practices in the United States. *Performance Improvement Quarterly*, 13 (3), 84-109.
- VanBrakel, R. (2002). Why ROI isn't enough. *T+D*, 56 (6), 72-75.
- VanBuren, M. K. (1999). *Investing in workforce training improves financial success*. Alexandria, VA: American Society for Training and Development.
- VanBuren, M. E. (2001). *State of the industry report 2001*. Alexandria, VA: American Society for Training and Development.
- VanBuren, M. E., & Erskine, J.J. (2002). *State of the industry report 2002*. Alexandria, VA: American Society for Training and Development.

- Vandenberg, R.J., & Lance, C.E. (1992). Examining the causal order of job satisfaction and organizational commitment. *Journal of Management*, 18 (1), 153-168.
- Wang, G.G. (2002). A systems approach to measuring return on investment for HRD interventions. *Human Resource Development Quarterly*, 13 (2), 203-226.
- Westcott, R. (1994). Applied behavior management training. In J. J. Phillips (Ed.), *Measuring return on investment* . Alexandria, VA: American Society for Training and Development.
- Williams, L.J., & Hazer, J.T. (1986). Antecedents and consequences of satisfaction and commitment in turnover models. *Journal of Applied Psychology* 71 (2), 219-231.
- Zellars, K.L., Hochwarter, W.A., Perrewe, P.L., Miles, A. K., & Kiewitz, C. (2001). Beyond self-efficacy: Interactive effects of role conflict and perceived collective efficiency. *Journal of Managerial Issues*, 13 (4), 483-500.
- Zigon, J. (1994). Yellow Freight Systems. In J. J. Phillips (Ed.), *Measuring return on investment*. Alexandria, VA: American Society for Training and Development.