

Empirical Validation of Outcomes from Training Programs: A Case Study

Gregory Kenneth Laing (Corresponding author)

Faculty of Business

University of the Sunshine Coast

Maroochydore DC Qld 4558 Australia

Email: глаинг@usc.edu.au Telephone: +61 7 5459 4675

Paul Andrews

Applied Consulting Pty Ltd

Email: paul@appliedconsulting.com.au

Accepted: August 19, 2011

DOI: 10.5296/ijhrs.v1i1.860

Abstract

The purpose of this paper is to demonstrate a method for the validation of training programs. This paper empirically tests variables to provide evidence in support of the outcomes from a training program in a case study. Whilst the ability to identify and measure the outcomes from training programs is highly relevant to all organizations it should be noted that no one set of variables will be applicable to all circumstances. The implications from this paper are that statistical analysis can be used to validate measured outcomes of a training program. The method used in this paper provides a technique that has applicability for use over a wide range of training programs. This method is especially significant for Strategic Human Resource Management as a means for providing supporting evidence to justify the outcomes from training programs.

Keywords – Training programs, human resource management.

1. Introduction

Human resource management plays an important role in an organisation's ability to remain competitive (Schuler & MacMillan, 1984; Klienman, 2000). A major contributing factor to the success of the human resource development is the provision of staff training. For a training program to be successful in achieving the goals of the organisation there needs to be pre-established performance objectives which are congruent with the strategic organisational goals. Consequently, it is essential that the outcomes of the training program be measured in terms of its contribution to the organisation's performance. However, this has been identified (Swanson, 1998; Cifalino & Baraldi, 2009) as a problem in the human resource management literature.

There are a number of approaches that have evolved in an attempt to address the evaluation of training programs from different perspectives (Ostroff, 1991). The most commonly used method is the Kirkpatrick Model (Kirkpatrick 1978, 1994), which adopts a four level approach to the evaluation of training programs (Swanson, 1998). However, Kirkpatrick's model has been criticised as being simplistic and lacking in any underlying economic or psychometric theory that can claim to provide valid measurements (Swanson, 1998). Typically, training programs have been evaluated from the perspective of the participants' post-training behaviour and generally compared to the original objective of the training with reliance upon the observations reported by their immediate supervisors.

The use of cost benefit analysis and return on investment (ROI) have been recognised as a more appropriate method to measure the performance of training programs (Nechvoglod, Karmel & Saunders, 2009; Doucouliagos & Sgro, 2000). Evaluations undertaken using these methods represent the outcomes of the investment in training from an economic and financial perspective. Whilst this is generally recognised as providing a more meaningful analysis the calculation of ROI presents a problem in most organisations. Specifically, the problem is associated with the identification of the relevant variables and the allocation of dollar values in terms of costs and benefits and this in turn causes management to doubt the validity of the claimed returns for the training dollars. This paper presents a framework that utilises a quantitative formula based on dollar values to mathematically determine the relative return on investment for a training program. These financial approaches overlook the importance of validating the outcomes from training programs. There is a common misconception that money spent on training is simply a cost.

Organisations choose to undertake training for a range of reasons, and not all training objectives are amenable to assessment of the return on investment. For example an organization seeking to ensure compliance with legislation will not necessarily be expecting to demonstrate a financial return over time. The literature provides some generalisations about when ROI may be most suited as an evaluation tool:

Table 1. *When an ROI study is and is not indicated*

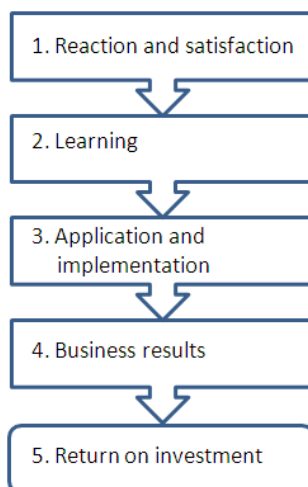
ROI study most suited when	ROI study least suited when
<ul style="list-style-type: none"> • A specific organisational performance indicator associated with financial benefits is targeted for improvement • Data is known (or able to be discovered) about current performance • There is agreement among the key stakeholders of the study as to the appropriateness of the measures • There is choice about the investment – ie it is not a legislated requirement. That is, the study will inform decisions about training. 	<ul style="list-style-type: none"> • The change sought from the training is unrelated to a financial return • Data about current and past performance is unknown, difficult to collect and or of doubtful validity and reliability • The data will make no difference to the decision (for example, it is mandated and will continue regardless of the evaluation conclusions)

Subsequently not every training program requires the application of the ROI model for the purpose of evaluation. However, from human resource management perspective validation of the outcomes is something which does require closer scrutiny.

2. Literature Review

Alampay and Morgan (2000) identified that the Kirkpatrick model (1994) was the most commonly used framework for evaluating training. The model does not extend to the identification of the benefits of the training program. Phillips (1996) addressed this short coming by adding a fifth level which requires that the benefits be converted into monetary value thus allowing for comparison against the full costs of the training program. Figure 1 provides an overview of the extended model that accommodates the Phillips extended model dealing culminating in the calculation of ROI.

Figure 1. *Extended Model*



The ROI process imposes the need to undertake five steps in order to derive a cost-benefit analysis. Adapting the ROI to the evaluation of training programs the steps are reframed to reflect the nature of the training evaluation:

1. Identify the training benefits;
2. Convert the benefits to a monetary value;
3. Calculate the total costs of the training;
4. Identify any intangible benefits (i.e. benefits which may or may not be converted to a monetary value); and
5. Compare the monetary benefits to the monetary costs.

The relationship between monetary benefits and monetary cost was further defined (Doucouliagos & Sgro, 2000 as involving; collecting the data, pre and post testing, multivariate analysis, and calculation of the return on investment. This four step approach is presented in Figure 1 along with additional comments as to the underlying activities associated with the process. The model proposed in this paper takes into consideration the existing models and extends the process to better reflect the stages involved in the conducting of an ROI evaluation.

3. Method

According to Yin (1989) the case study is a useful way to evaluate theoretical issues by examining real life situations. The argument being that a case study enables the researcher to become more focused on the details under examination in the research. This in turn allows for the possibility that the focus may change as new knowledge is acquired which may challenge the propositions of the research (Yin 1989). The case study may be useful when the research is seeking to explore issues that might otherwise be missed or overlooked and this can be important for the early testing in the development stage of a theoretical model or framework.

3.1 Case Study Background

Company “A” is a large construction and engineering organisation with a workforce of over 5,000. In 2003 Company “A” developed a Leadership Program, which they referred to as Quality in Leadership (QIL). The program had as its objective the desire to better prepare new leaders at Superintendent and Supervisor levels in line with the organisations aim to achieve the development and promotion of staff within the organisation. Specifically the six values that the company sought to encourage and promote were:

- Safety and the environment
- Integrity
- High performance
- Win-Win relationships
- The courage to lead change
- Respect for each other

The QIL program was delivered via a range of strategies including off job training, self paced learning materials, coaching and mentoring and on job assignments and projects. Participants

in the program could choose to aim for a recognised qualification within the Australian Vocational Education Framework. The program operated from 2003 until June 2009 at which time a review of the outcomes was undertaken.

The management of Company “A” were interested in understanding the extent to which the program had influenced performance against the specific values that had been established at the commencement of the training program. Management required the analysis to provide evidence that the measurement of performance outcomes relative to the training program were valid indicators. To address this requirement statistical testing was determined to be the most appropriate approach to provide validation.

3.2 Data Collection

The staff performance appraisal system in Company “A” focused on the four outcomes that the training was intended to address. That is improved levels of performance by staff in the areas of general rating; behavioural rating; supervisor rating; and overall performance rating. Subsequently, the data collected through the staff appraisal system was used to operationalise the outcomes as variables to be tested. The appraisal was conducted annually and included peer and supervisor ratings. Success of the training was determined to be assessable in terms of the extent of improvement in the rating scores over the duration of the program. An additional variable collected was time in role and this was done for application in a later stage of the validation process.

4. Case Study Results

4.1 Pre-test post-test

A pre-test post-test analysis was conducted to provide evidence as to whether the observed changes in performance levels after the training were of a significant difference or not. A t-test was used and as reported in Table 2 the result indicates that the post-test performance scores were significantly different. This test does not provide conclusive evidence that the training was the sole reason for the changes.

Table 2. *Statistical Results for Pre-test Post-test*

<i>Variable</i>	<i>Value of t-test</i>	<i>Comment</i>
<i>Overall performance</i>	0.0107	Significant at the 0.01 level!

Note: Two tail distribution, two sample equal variance (homoscedastic) assumed.

4.2 Control Group vs Subjects

Having established that the improved performance scores reported after having completed the training were significantly different the next test is to examine whether the performance scores were significantly different between those staff members who had undertaken and completed the training against those who had not. To address this stage of testing the overall performance rating data was collated according to whether the staff member had undertaken the training program or not. The staff that had not undertaken the training was used as a control group to statistically test whether any change that had been identified was significantly different between the two groups. That is whether any differences that might be

observed were due to mere chance or the result of the training. To examine this aspect the data were analysed using the student *t*-test the results are presented in Table 3.

Table 3. *Statistical Results for Control Group versus Subjects*

<i>Variable</i>	<i>Value of t-test</i>	<i>Comment</i>
<i>Overall performance</i>	0.0453	Significant at the 0.05 level!

Note: Two tail distribution, two sample equal variance (homoscedastic) assumed.

The result from the *t*-test indicates that the positive difference between the results of the group that did complete versus the other two groups was significant. That is the *t*-test indicates that this difference is significant at the 5% level – indicating that the difference is due to (or the result of) the completion of the training program. However, this test does not provide evidence that the training was the only reason for the change that will be examined in the next stage.

4.5 Multivariate Analysis

While the *t*-test was used to determine the level of significance of any change, that is, that such change was not due to some random chance. Improvement in ratings may be the result of other factors the most obvious is the growth of experience in the role over time. To examine this aspect a multiple regression was conducted using time in role as the dependent variable and the ratings as the independent variables. This test was undertaken to determine the extent to which the training program as opposed to additional experience may have been responsible for the observed improvement in the rating scores. The results of the multiple regression analysis are presented in Table 4.

Table 4. *Multiple Regression Analysis Results*

<i>Regression Statistics</i>					
Multiple R	0.481313662				
R Square	0.231662841				
Adj R Square	0.156703119				
Standard Error	1.598695471				
Observations	46				

	<i>Coefficients</i>	<i>Standard Error</i>	<i>t Stat</i>	<i>P-value</i>	<i>Lower 95%</i>
Intercept	2.2256891	2.2504113	0.989014	0.328457175	-2.3191087
Overall rating	0.2772447	0.7086365	0.391237	0.697648	-1.1538757
General rating	0.6145306	0.2163144	2.840913	0.006974912	0.1776747
Behavioural rating	0.3090658	0.6681855	0.462545	0.646136118	-1.0403621
Supervisor rating	-1.2170471	0.6441106	-1.8895	0.06591042	-2.5178550

In three of the four performance rating scores the *p*-values are not significant. However, in

regards to the general rating scores there was a statistically significant relationship found. Upon further investigation it was discovered that this aspect in the staff appraisal was closely linked to the time in role and this anomaly was raised for review of the design of the appraisal form. The results from the regression analysis indicate that experience gained from time in the role did not have a significant contributory effect on all the improved levels of performance.

4.6 Case Summary

This case study has been concerned with providing validation of the outcomes from the QIL training program which had been developed and implemented to achieve improved levels of performance by the staff. Based on the validation provided by the statistical analysis the following conclusions were able to be made:

- The QIL program made a statistically significant impact on the performance of the participants especially when contrasted against prescribed outcome measures of the staff performance appraisal system.
- A possible explanation for improved performance against the prescribed measures (specifically length of time in role) was found to have a less than significant contribution to the overall improved performance.

5. Discussion

Perhaps the most significant assumption that underlies this case is that the rating scores derived from the staff performance appraisals were an acceptable proxy measure for performance against the established values identified by the company. The justification for this assumption was based upon the argument that the ratings provide the only readily available data that had been consistently undertaken each year within the confines of the company itself. Certainly within Company “A” the outputs of the performance appraisal process were viewed by management as being a valid way of demonstrating such progress of staff on an annual basis.

The approach employed in this paper has practical implications for human resource professionals. Firstly, it provides an empirical method for the testing and validating of claims that training programs have a positive impact in terms of measurable outcomes. Secondly, the statistical techniques used in this approach can now be done using the Microsoft Excel spreadsheet software which is readily available and more user friendly than the complex statistical packages.

References

- Cullen, J., Sawzin, S., Sisson, G. & Swanson, R. 1978. Cost Effectiveness: A Model for Assessing the Training investment, *Training and Development Journal*, January, pp.24-29.
- Doucouliafos, C. & Sgro, P. 2000. *Enterprise Return on a Training Investment*, National Centre for Vocational Education Research Ltd: Adelaide.
- Kirkpatrick, D.L. 1978. Evaluating in-house training programs, *Training and Development Journal*, Vol. 38, pp.32-37.
- Kirkpatrick, D.L. 1994. *Evaluating training programs: the four levels*, Berrett-Koehler: San Francisco.
- Kleiman L.S. (2000). *Human Resource Management: A Managerial Tool for Competitive Advantage*, 2nd Edn., South- Western College Publishing: Cincinnati, OH.
- Nechvoglod, L., Karmel, T. & Saunders, J. 2009. *The Cost of Training Apprentices, Occasional Paper*, National Centre for Vocational Education Research Ltd: Adelaide.
- Ostroff, C. 1991. Training Effectiveness Measures and Scoring Schemes: A Comparison, *Personnel Psychology*, Vol.44, No.2, pp.353-374.
- Schuler R.S. & MacMillan I.C. (1984). Gaining Competitive Advantage through Human Resource Management Practices, *Human Resource Management*, 23(3), pp. 241-255.
- Swanson, R.A. 1998. Demonstrating the Financial Benefit of Human Resource Development: Status and Update on the Theory and Practice, *Human Resource Development Quarterly*, Vol.9, No.3, pp.285-295.
- Yin, R. (1989). *Case Study Research Design and Methods*, Sage Publications: Beverly Hills.