The Journal of Commerce
Vol.6, No.2 pp.19-30

Methods of Training Programmes Evaluation: A Review

Surbhi Jain*

Abstract

The purpose of this paper is to review the existing literature on the methods of training programmes evaluation. Evaluation measures the extent to which programs, processes, or tools achieve the purpose for which they were intended. Phillips (1991) defined evaluation as a systematic process to determine the worth, value, or meaning of something. In this review, evaluation is defined as a study designed and conducted to assist some audience to assess an object's merit and worth (Stufflebeam, 2001). One major model of evaluation was identified. This model, developed by Kirkpatrick in 1952, remains widely used today (ASTD, 1997). The model includes four levels of measurement to assess reaction, learning, behavior, and results as related to specific training. Developing evaluation strategies based on the Kirkpatrick Model holds the greatest promise for systematic assessment of training within organizations.

Key words: Remittances Training, training programmes evaluation, cost-effectiveness

Introduction

Change is everywhere. It is inescapable. Also, today's business environment is highly competitive. In difficult economic times, when an organization is fighting harder than ever to maintain market share, it's absolutely essential to maximize employee knowledge and skills. Training is the acquisition of knowledge, skills, and competencies as a result of the teaching of vocational or practical skills and knowledge that relate to specific useful competencies. India, though a developing country, spends $50 billion on training each year whereas wealthy USA spends nearly

*Research Scholar, MDU Rohtak, India, surbhijain78@gmail.com.

twice as much. Expenditures of such magnitude call for a sharp periodic look. Thus calls for evaluation of training programmes.

Rationale of the study

Human resource is a key factor for production and hence for improved business performance. Still, in many organizations, the management of HR has not performed up to the expectations. Since most of the business organizations are spending billions of dollars for training programmes, evaluation of these programmes becomes inevitable.

Material and Methods

A review of literature on evaluation of training programmes was conducted to identify methods of effective evaluation for training programs. Five definitions of evaluation were identified in the literature.

Phillips (1991) defined evaluation as a systematic process to determine the worth, value, or meaning of something.

Holli and Calabrese (1998) defined evaluation as comparisons of an observed value or quality to a standard or criteria of comparison. Evaluation is the process of forming value judgments about the quality of programs, products, and goals.

Boulmetis and Dutwin (2000) defined evaluation as the systematic process of collecting and analyzing data in order to determine whether and to what degree objectives were or are being achieved.

Schalock (2001) defined evaluation as the determination of the extent to which a program has met its stated performance goals and objectives.

Stufflebeam (2001) defined evaluation as a study designed and conducted to assist some audience to assess an object's merit and worth.

Stufflebeam's (2001) definition of evaluation was used to assess the methods of evaluation found in this literature review. The reason for selecting Stufflebeam’s definition was based on the applicability of the definition across multiple disciplines. Based on this definition of evaluation, the Kirkpatrick Model was the most frequently reported model of evaluation.

Kirkpatrick, 1971

Kirkpatrick uses four levels of measurement for evaluating a training programme. Kirkpatrick's first level of measurement, reaction, is defined as how well the trainees liked the training program. The second measurement level, learning, is designated as the determination of what knowledge, attitudes, and skills were learned in the training. The third measurement level is defined as behavior. Behavior outlines a relationship of learning (the previous measurement level) to the actualization of doing. Kirkpatrick recognized a big difference between knowing principles and techniques and using those principles and techniques on the job. The fourth measurement level, results, is the expected outcomes of most educational training programs such as reduced costs, reduced turnover and absenteeism, reduced grievances, improved profits or morale, and increased quality and quantity of production.

Paquet, Kasl, Weinstein, & Waite, 1987

One study was found by a major corporation in USA that measured change in productivity and ROI of a training program. CIGNA Corporation's corporate management development and training department, which provides training for employees of CIGNA Corporation's operating subsidiaries, initiated an evaluation program to prove management training made a business contribution. The research question posed was, "Does management training result in improved productivity in the manager's workplace?" The team conducting the research identified that data collection needed to be built into the training program for optimal data gathering. If managers could use the evaluation data for their own benefit as part of their training, they would be more likely to cooperate. As a result, the measure of productivity was implemented as part of Basic Management Skills training throughout CIGNA Corporation.

Alliger and Horowitz, 1989

Numerous studies reported use of components of the Kirkpatrick Model; however, no study was found that applied all four levels of the model. Although level one is the least complex of the measures of evaluation developed by Kirkpatrick, no studies were found that reported use of level one as a sole measure of training. One application of the second level of evaluation, knowledge, was reported by this study. In this study the IBM Corporation incorporated knowledge tests into internally developed training. To ensure the best design, IBM conducted a study to identify the optimal test for internally developed courses. Four separate tests composed of 25

questions each were developed based on ten key learning components. Four scoring methods were evaluated including one that used a unique measure of confidence. The confidence measurement assessed how confident the trainee was with answers given. Tests were administered both before and after training. Indices from the study assisted the organization to evaluate the course design, effectiveness of the training, and effectiveness of the course instructors. The development of the confidence index was the most valuable aspect of the study. Alliger and Horowitz stated that behavior in the workplace was not only a function of knowledge, but also of how certain the employee was of that knowledge.

Bushnell (1990)

Bushnell also created a modification to the Kirkpatrick Model by identifying a four-step process of evaluation. Bushnell's model included evaluation of training from the development through the delivery and impact. Step one involved the analysis of the System Performance Indicators that included the trainee's qualifications, instructor abilities, instructional materials, facilities, and training dollars. Step two involved the evaluation of the development process that included the plan, design, development, and delivery. Step three was defined as output which equated to the first three levels of the Kirkpatrick Model. Step three involves trainees' reactions, knowledge and skills gained, and improved job performance. Bushnell separated outcomes or results of the training into the fourth step. Outcomes were defined as profits, customer satisfaction, and productivity. This model was applied by IBM's global education network, although specific results were not found in the literature.

Phillips (1991)

…stated the Kirkpatrick Model was probably the most well-known framework for classifying areas of evaluation. This was confirmed in 1997 when the America Society for Training and Development (ASTD) assessed the nationwide prevalence of the importance of measurement and evaluation to human resources department (HRD) executives by surveying a panel of 300 HRD executives from various types of U.S. organizations. Survey results indicated the majority (81%) of HRD executives attached some level of importance to evaluation and over half (67%) used the Kirkpatrick Model. The most frequently reported challenge was determining the impact of the training (ASTD, 1997).

Lookatch (1991) and ASTD (2002)

… reported that only one in ten organizations attempted to gather any results-based evaluation. In 1952, Donald Kirkpatrick (1996) conducted doctoral research to evaluate a supervisory training program. Kirkpatrick's goal was to measure the participants' reaction to the program, the amount of learning that took place, the extent of behavior change after participants returned to their jobs, and any final results from a change in behavior achieved by participants after they returned to work. From Kirkpatrick's doctoral research, the concept of the four Kirkpatrick measurement levels of evaluation emerged. While writing an article about training in 1959, Kirkpatrick (1996) referred to these four measurement levels as the four steps of a training evaluation. It is unclear even to Kirkpatrick how these four steps became known as the Kirkpatrick Model, but this description persists today (Kirkpatrick, 1998). As reported in the literature, this model is most frequently applied to either educational or technical training.

Wagner & Roland, 1992

Over 20 organizations and 5,000 participants were studied. Three measures were used to determine behavioral changes. Measure one was a questionnaire completed by participants both before and after training. The second measure was supervisory reports completed on the functioning of work groups before and after training. The third measure was interviews with managers, other than the immediate supervisor, to obtain reactions to individual and work-group performance after an OBERT (Outdoor-Based Experiential Training with the goal of team building) program. Results reported showed no significant changes in behavior.

Stephen Birch and Amiram Gafni (1992)

… said in their research paper despite the growing literature on economic evaluation of training programs, little attention has been paid to the theoretical foundations of cost- effectiveness and cost utility analyses and the validity of the decision rules adopted as methods of achieving the stated goals.

Marshall and Schriver (1994)

Another adaptation of the Kirkpatrick Model was developed by Marshall and Schriver (1994) in work with Martin Marietta Energy Systems. Marshall and Schriver suggested that many trainers misinterpreted the Kirkpatrick Model and believed that an evaluation for knowledge was the same as testing for skills. Because skills and

knowledge were both included in level two of the Kirkpatrick Model, evaluators assumed skills were tested when only knowledge was tested. As a result, Marshall and Schriver recommended a five-step model that separated level two of the Kirkpatrick Model into two steps. Only the theory of the model was presented in the article; no application of this model was found.

Clifton P. Campbell, (1994)

… discusses the need to justify training expenditures with targeted benefits. Provides details on how to calculate the direct, indirect, and full costs of a training course or program. Also describes the feasibility of linking training outcomes to organizational improvements and the selection of training outcomes (benefits) to be measured and quantified. While calculating the full cost of training is a first and critical step in determining cost effectiveness, monitoring costs is also important to planning and controlling the training budget. After training managers learn how to calculate the cost and measure the effectiveness of training, they want to know is the training effort producing benefits that are greater than the costs involved?

Clifton P. Campbell, (1995)

Following on from part one, describes four practical methods for determining the cost-effectiveness of training. Presents details and examples on how to use each method. Also identifies the advantages and disadvantages of each method. A variety of methods are available for determining the cost –effectiveness of training. Some are complex and difficult to utilize, while others are more suitable for research projects. The four methods for justifying a training investment presented in this study were selected because they are practical, relatively easy to use and generally familiar to higher management. The four methods described here are: return on investment (ROI); cost-benefit ratio; bottom-line evaluation; and payback period. Ends with a skill check which provides an opportunity to apply the content covered.

Kirkpatrick (1998)

… recommended that as many as possible of the four levels of evaluation be conducted. In order to make the best use of organizational resources of time, money, materials, space, equipment, and manpower, continued efforts are needed to assess all levels of effectiveness of training programs. Trainers from all disciplines should develop evaluation plans for training and share the results of these initiatives.

Warr, Allan and Birdie (1999)

… evaluated a two-day technical training course involving 123 motor-vehicle technicians over a seven- month period in a longitudinal study using a variation of the Kirkpatrick Model. The main objective of this study was to demonstrate that training improved performance, thereby justifying the investment in the training as appropriate. Warr et al. (1999) suggested that the levels in the Kirkpatrick Model may be interrelated. They investigated six trainee features and one organizational characteristic that might predict outcomes at each measurement level. The six trainee features studied were learning motivation, confidence about the learning task, learning strategies, technical qualifications, tenure, and age. The one organizational feature evaluated was transfer climate which was defined as the extent to which the learning from the training was actually applied on the job. Warr et al. (1999) examined associations between three of the four measurement levels in a modified Kirkpatrick framework. Warr et al. combined the two higher Kirkpatrick measurement levels, behavior and results, into one measurement level called job behavior. The three levels of measurement included were reactions, learning, and job behavior. . Findings suggested a possible link between reactions and learning that could be identified with the use of more differentiated indicators at the reaction level. Warr et al. suggested that an investigation into the pretest scores might explain reasons for the behavior and generate organizational improvements.

Abernathy (1999)

… admitted quantifying the value of training was no easy task and presented two additional variations of the Kirkpatrick Model; one developed by Kevin Oake and another developed by Julie Tamminen. However, no application of the two models was provided by Abernathy.

Phillips and Pulliam (2000)

… reported an additional measure of training effectiveness, return on investment (ROI), and was used by companies because of the pressures placed on Human Resource Departments to produce measures of output for total quality management (TQM) and continuous quality improvements (CQI) and the threat of outsourcing due to downsizing. Great debate was found in the training and development literature about the use of ROI measures of training programs. Many training and development professionals believed that ROI was too difficult and unreliable a measure to use for training evaluation (Barron, 1997).

Belfield, Hywell, Bullock, Eynon, and Wall (2001)

… considered the question of how to evaluate medical educational interventions for effectiveness on healthcare outcomes using an adaptation of the Kirkpatrick Model with five levels. The five levels were participation, reaction, learning, behavior, and outcomes. By applying the adapted Kirkpatrick Model to a meta-analysis of 300 abstracts about educational interventions within the medical profession, Belfield et al. indicated that a limited number (less than 2%) evaluated healthcare outcomes. The majority of the abstracts reviewed (70%) assessed medical educational interventions at the level of learning. Ambiguity was reported within the articles because of incorrect term usage. Of those examined, Belfield et al. indicated the authors needed to focus on clear communication of the design of evaluation methods.

Radhakrishna, Plank, and Mitchell (2001)

… used a learning style instrument (LSI) and a demographic profile in addition to reaction measures and learning measures. The three training objectives were to assess knowledge gained through a Web-based training, to determine participant reaction to Web-based material and Listserv discussions, and to describe both the demographic profile and the learning style of the participants. The evaluation of the training began with an on- line pretest and an on- line LSI. The pretest included seven demographic questions. The LSI, pretest and posttest, and LSI questionnaire were paired by the agent's social security numbers. Fifty- five agents of the available (106) agents completed all four instruments and were included in this study.

Ignace Ng, Ali Dastmalchian (2011)

The purpose of this study is to examine the link between training and the perceived contribution of training to enhanced productivity or cost reduction. Using data from 92 Canadian organizations, the results show that organizations with higher percentage of trained employees are likely to perceive training to be beneficial. In addition, the results indicate that perceived benefits of training are further enhanced by the presence of human resources management practices that either encourages employees to undertake training (the motivation bundle) and/or provides a systematic assessment of post-training effectiveness (the assessment bundle). The evidence however also shows that open climate as measured by autonomous work systems nullifies the benefits of training, suggesting that under such a structure, employees are unlikely to put in practice the skills they acquired during training.

Wilawan O., Waldemar K., Tareq Z. Ahram (2012)

Financial costs of investing in people is associated with training, acquisition, recruiting, and resolving human errors have a significant impact on increased total ownership costs. These costs can also affect the exaggerate budgets and delayed schedules. The study of human performance economical assessment in the system acquisition process enhances the visibility of hidden cost drivers which support program management informed decisions. This paper presents the literature review of human total ownership cost (HTOC) and cost impacts on overall system performance. Economic value assessment models such as cost benefit analysis, risk-cost tradeoff analysis, expected value of utility function analysis (EV), growth readiness matrix, multi-attribute utility technique, and multi-regressions model were introduced to reflect the HTOC and human performance technology tradeoffs in terms of the dollar value. The human total ownership regression model introduces to address the influencing human performance cost component measurement. Results from this study will increase understanding of relevant cost drivers in the system acquisition process over the long term.

Conclusion

The usefulness of training evaluation was demonstrated in the studies reported by many authors. The Kirkpatrick Model was assessed as a valuable framework designed with four levels of measure to evaluate the effectiveness of a training programme. Organizations recognize that training works (Skerlavaj, Dimovski, Mrvar, &Pahor, 2010) and spend billions of dollars every year (O'Leonard, 2012; Paradise, 2007) to train their employees. The problem is that training works only if the trainee transfers the training (applies the training) to improve performance on the job. It is estimated that only 10% to 30% of training transfers to on-the-job performance (Broad, 2005). Training transfer is "the extent to which the learning that results from a training experience transfers to the job and leads to meaningful changes in work performance" (Baldwin, Ford, &Blume, 2009). Investment dollars spent on training that does not transfer to on-the-job performance are a wastage of an organization's limited resources. One study reported that only one in ten organizations attempted to gather any results-based evaluation.

Value was based on the foundational ideas of Kirkpatrick and the longitudinal strength of the model. The popularity of the Kirkpatrick Model was demonstrated by the 1997 ASTD survey results; however, few studies showing the full use of the model were found. In addition to the Kirkpatrick Model, six adaptations were found; but no application was found for three of these adapted

models. One developed a confidence index to measure knowledge of employees and concluded that behavior in the workplace was not only a function of knowledge, but also of how certain the employees was of that knowledge. An additional measure of training effectiveness, return on investment (ROI), was used by companies because of the pressures placed on Human Resource Departments to produce measures of output for total quality management (TQM) and continuous quality improvements (CQI). One study discuss the need to justify training expenditures with targeted goals, provides details on how to calculate the direct, indirect and full costs of a training progrmme. Kirkpatrick (1998) recommended that as many as possible of the four levels of evaluation be conducted. In order to make the best use of organizational resources of time, money, materials, space, equipment, and manpower, continued efforts are needed to assess all levels of effectiveness of training programs. Trainers from all disciplines should develop evaluation plans for training and share the results of these initiatives.

References

Abernathy, D. (1999). Thinking outside the evaluation box. *Training and Development, 53*(2), 18-24.

Alliger, G.M., & Horowitz, H.M. (1989). IBM takes the guessing out of testing. *Training and Development, 43*(4), *69-73.*

American Society for Training and Development. (2002). *Training for the next economy: An ASTD state of the industry report on trends in employer-provided training in the United States.* Retrieved February 6, 2003 from http://store.astd.org/default.asp

Belfield, C., Hywell, T., Bullock, A., Eynon, R., & Wall, D. (2001). Measuring effectiveness for best evidence medical education: A discussion. *Medical Teacher, 23*(2), 164-170.

Boulmetis, J., & Dutwin, P. (2000). *The abc's of evaluation: Timeless techniques for program and project managers*. San Francisco: Jossey-Bass.

Bushnell, D. S. (1990). Input, processing, output: A model for evaluating training. *Training and Development, 44*(3), 41-43.

Holli, B., & Calabrese, R. (1998). *Communication and education skills for dietetics professionals* (3rd ed.). Philadelphia: Lippincott Williams & Wilkins.

Horton, W. (2001). *Evaluating e-learning*. Alexandria, VA: American Society for Training and Development.

Ng Ignace, Dastmalchain Ali, Feb 2011,'Perceived training benefits and training bundles: A Canadian study', *International Journal of human research management*, Vol.22, p829-842

Kirkpatrick, D. (1971). *A practical guide for supervisory training and development*. Reading, MA: Addison-Wesley.

Kirkpatrick, D. (1996). Great ideas revisited. *Training and Development, 50*(1), 54-60.

Lippert, R., Radhakrishna, R., Plank, O., & Mitchell, C. (2001). Using different evaluation tools to assess a regional Internet in-service training. *International Journal of Instructional Media, 28*(3), 237-249.

Lookatch, R.P (1991). HRD's failure to sell itself. *Training and Development, 45*(7), 35-40.

Marshall, V., & Schriver, R. (1994). Using evaluation to improve performance. In D. Paquet, B., Kasl, E., Weinstein, L., & Waite, W. (1987). The bottom line. In D. Kirkpatrick (Ed.), *Another look at evaluating training programs* (pp.175-181).

Phillips, J. (1991). *Handbook of training evaluation and measurement methods* (2nd ed.). Houston: Gulf Publishing Company.

Phillips, J., & Pulliam, P. (2000). *Level 5 evaluation: Measuring ROI*. Alexandria, VA: American Society of Training and Development.

Schalock, R. (2001). *Outcome based evaluations* (2nd ed.). Boston: Kluwer Academic/Plenum.

Stufflebeam, D. L. (2001). *Evaluation models.* San Francisco: Jossey-Bass. Torres, R. T., Preskill, H.S., & Piontek, M.E. (1996). *Evaluation strategies for communicating and reporting.* Thousand Oaks, CA: Sage Publications, Inc.

Warr, P., Allan, C., & Birdi, K. (1999). Predicting three levels of training outcome. *Journal of Occupational and Organizational Psychology, 72*(3), 351-375.

Wagner, R. J., & Roland, C. C. (1992). How effective is outdoor training? *Training and Development, 46*(7), 61-66.

Wilawan O., Waldemar K., Tareq Z. Ahram. (2012). Economics of human performance and systems total ownership cost. *Work: A Journal of Prevention, Assessment and Rehabilitation*, 41, 2781-2788.