Revisiting Kirkpatrick's model – an evaluation of an academic training course

P. Rajeev*, M. S. Madan and K. Jayarajan

Indian Institute of Spices Research, Marikunnu PO, Calicut 673 012, India

In scientific and research organizations, the training needs facilitator roles and methods have undergone a change necessitated by rapid information and technology boom. There is ample evidence to show that evaluation and objective assessment of effectiveness and outcomes of training programmes being implemented by organizations are not given due importance as that of their planning and implementation. An attempt is made in this communication firstly to analyse the theories of training evaluation in general; the study also illustrates a case study of training evaluation of the academic training courses being carried out at the Indian Institute of Spices Research by revisiting the popular Kirkpatricks's model. The three-step evaluation model is a combination of formative and summative approaches using multiple methods which measure reactions, perceptions, learning and behavioural components of the trainees combining quantitative and qualitative tools and aims at assessing the usefulness of the course in providing an adequate learning climate.

Keywords: Formative evaluation, Kirkpatrick's model, summative evaluation, training evaluation models.

TRAINING is an essential human resource development (HRD) function of any organization. In the Indian organizational development context, the training needs, strategies, methods and investments on training have all undergone a sea change since the last decade. Science and technology is growing faster than in the past and to tackle the global competition in many sectors, the erstwhile traditional approach in training policies of organizations has undergone a change to one which is more liberal, concept-based, comprehensive, systematic, well planned and dynamic[1]. For instance, in scientific and research organizations devoted to life sciences, these influences are due to fast technological developments in frontier areas like biotechnology, bioinformatics, biochemistry and microbiology. The challenges opening up provide a wide range of opportunities subject to the acquisition of relevant skills, knowledge and concepts. Realizing the need for HRD support, most scientific institutions have taken up a facilitating role in providing infrastructure and academic training support to clientele groups. This communication is based on the hypothesis that for the success of any academic programme, it is necessary that every course should have an in-built monitoring and evaluation system. Here we first review the contemporary theories of evaluation of training programmes in general and secondly, reveal a case study on the application of a widely accepted theoretical model to evaluate an academic training course offered to postgraduate students in life sciences.

With the changing socio-economic and technological relevance of training, the definitions, scope, methods and evaluation of training programmes have also changed. One of the earlier classic definitions of training is 'bringing lasting improvement in skills in jobs'[2]. The present-day definitions take a multi-dimensional perspective enveloping the needs of individuals, teams, organizations and the society. The steps in the training programme development are planning, programme implementation, and programme evaluation and follow-up. The evaluation of any training system helps measure the 'knowledge gap', what is defined by Riech[3] as 'the gap between what the trainer teaches and what the trainee learns'. Evaluations help to measure Reich's gap by determining the value and effectiveness of a learning programme. It uses assessment and validation tools to provide data for the evaluation. Bramley and Newby[4] identify four main purposes of evaluation:

1. Feedback – Linking learning outcomes to objectives and providing a form of quality control.
2. Control – Making links from training to organizational activities and to consider cost-effectiveness.
3. Research – Determining the relationships among learning, training and transfer of training to the job.
4. Intervention – The results of the evaluation influence the context in which it occurs.

Evaluation of training systems, programmes or courses tends to be a demand of a social, institutional or economic nature[5]. The demand of a social nature became stronger with the participation of trainees themselves and other non-traditional actors like the resource persons in the evaluation process. Evaluation is also a growing institutional demand, since the decision-makers not only need to understand the processes, but also to control and act to improve the effects of their guidelines. Evaluation is still an economic demand given the community involvement in funding. However, despite the importance of the evaluation process, there is evidence that evaluation of training programmes is often inconsistent or missing[6–8]. The literature on evaluation needs to be classified into education and training. The latter reveals many difficulties as regards evaluation. Scientific and quantitative methods are not popular. Evaluation appears to be undertaken reluctantly and with the simplest of methods. Behavioural objects are rarely even set by the trainers. Progress in the techniques of evaluation has been slow, though a good deal of research has been done. The literature is small, but growing[9].

*For correspondence. (e-mail: rajeev@spices.res.in)

Programme evaluation research involves two general approaches. Formative evaluation (also known as internal) is a method of judging the worth of a programme while the programme activities are forming (in progress). This part of the evaluation focuses on the process. This approach has been termed as process evaluation[10]. In essence, the questions being answered in formative evaluation are: 'How was the programme implemented?' and 'Was it implemented as planned?'. Results of formative evaluations are often used to improve programme implementation by providing a feedback, which can be used to modify future implementations[11]. For example, formative studies examine issues such as the consistency with which a programme was implemented, other programmes that were implemented during the study, attitude towards the programme and duration of implementation. As applied to HR programmes, formative evaluation offers a means of assessing and improving on programme validity[12]. The summative evaluation (also known as external evaluation) is a method of judging the worth of a programme at the end of the activities (summation). The focus is on the outcome. Summative evaluation assesses the extent to which the intervention achieved the outcomes described by its goals. Often summative evaluations utilize quasi-experimental research designs such as pre-test/post-test, randomized control group design, time series, or a combination of each.

The most influential framework for the evaluation of training programmes has come from Kirkpatrick[6,13–16]. Kirkpatrick's model[17] follows a goal-based approach. Even now this is the most popular model being used with required modifications and is applicable to any organizational setting. For instance, this model has been applied in evaluating training imparted to child-welfare professionals as well as entrepreneurship development training programmes[18,19]. Most of the models in use today are modified versions of Kirkpatrick's four-level framework[20,21]. Several weaknesses have been identified with Kirkpatrick's model, including overemphasis on the reactions of trainees, low correlation between reactions and performance, low correlation between measures at different outcome levels, and incompleteness of the model[22–26]. Other models for training evaluation have been presented in the literature, including contextual evaluation[27], responsive evaluation[28] and balanced score card.

Kirkpatrick's model is based on four simple questions that translate into four levels of evaluation. These are:

Level 1. Reaction: At this level, data on the reactions of the participants at the end of a training programme are gathered. This level is often measured with attitude questionnaires that are passed out after most training classes. This level measures the learner's perception (reaction) of the course.

Level 2. Learning: The intention at this level is to assess whether the learning objectives for the programme are met. This is usually done by means of an appropriate test or examination. The learning evaluation requires post-testing to ascertain what knowledge was learned during the training. In addition, the post-testing is only valid when combined with pre-testing, so that one can differentiate between what he already knew prior to the training and what he actually learned during the training programme.

Level 3. Behaviour: The intention at this level is to assess whether job performance changes as a result of training. This performance testing is to indicate the learner's skill to apply what he has learned in the classroom. This evaluation involves testing the students capabilities to perform learned skills while on the job, rather than in the classroom. Level-three evaluations can be performed formally (testing) or informally (observation and judgments).

Level 4. Results: The intention at this level is to assess the costs vs benefits of training programmes, i.e. organizational impact in terms of reduced costs, improved quality of work, increased quantity of work, etc. It measures impact, which includes monetary efficiency, moral, teamwork, etc. Collecting, organizing and analysing level-four information can be difficult, time-consuming and more costly than the other three levels, but the results are often quite worthwhile when viewed in the full context of its value to the organization[29].

After examining the methodological framework, a case study is presented on the evaluation methodology at the Indian Institute of Spices Research, Calicut. Among other programmes, training courses are organized presently at the Institute in the field of bioinformatics and biotechnology and biochemistry. The trainees are postgraduate students in life sciences and the objective of the course is to impart knowledge on concepts and methods in these fields and impart skills in using various scientific tools. The curriculum involves hands-on training on skills related to analytical techniques in biochemistry like GLC, GCMS, HPLC, etc., isolation of proteins, enzymes, DNA, RNA, plant tissue culture and micropropagation, DNA markers, preparation of molecular maps and molecular approaches in the detection and isolation of plant pathogens. The course duration is 30 days and trainees are selected based on acceptable standards of their performance in the regular courses they do in order to ensure homogeneity as far as possible.

These training programmes have an inbuilt monitoring and evaluation system; formal evaluation of participants in terms of their achievements and self-appraisal by participants in terms of their perceptions about training with the ultimate objective of assessing the usefulness of academic training in providing an adequate learning climate that could be linked with their career development. A three-stage evaluation method is followed: training orientation and pre-training evaluation, concurrent evaluation and post-training evaluation. Knowledge gain, on job performance of skills and training effectiveness index are the three major dependent variables in the method. Assessment of cognitive learning is measured as knowledge gain[30] and skill in terms of performance. Perceptions of the trainees on the organizational effectiveness are measured as the training effectiveness index.

During the training orientation session, a platform is provided in an open discussion meeting to build faculty–trainee rapport and for collecting 'open responses' on 'what are the trainees' expectations' in undergoing the course. During this session each trainee is assigned an advisor faculty for guidance throughout the duration of the course.

During pre-training evaluation, a comprehensive, quiz-type knowledge test is administered to assess the initial level of knowledge. This test is prepared with more knowledge-type questions rather than understanding and skill-oriented questions.

During the concurrent evaluation session, performance-oriented tests are given. At this stage a mix of objective and subjective tests is employed. This consists of an unannounced surprise performance test comprising more in-depth understanding and descriptive type of questions, judgment of performance of trainees in practical laboratory sessions (judges' rating) and judging the comprehension and communication skills through a pre-decided term paper prepared by the trainees and also presented by each trainee in a seminar. Concurrent evaluation is oriented towards reflecting upon the performance, soft skills and measuring the knowledge as well as skills (behavioural component) likely to reflect in the sphere of activity of the trainees. Here, knowledge and skill constructions are assessed at three levels – personal, professional and disciplinary[31].

During the post-training session, the knowledge test, similar to the one given during the pre-training session is repeated with the purpose of measuring the knowledge gain. At this stage a questionnaire is also administered[32] to measure the perceptions, reactions and attitude of the trainees about organizational effectiveness of the training including logistic support. This questionnaire consists of pre-tested fixed alternative Likerts rating-scale questions and some open-ended questions. The dimensions in the scale rated are satisfaction in terms of educational experience, attention, methodology, interest, training materials, instruction in terms of knowledge and quality and practical sessions. The quantitative data on test scores and rating scores is analysed using appropriate statistical tools like central tendency measures, skewness coefficient, coefficient of variation, indices and paired *t*-test and multiple correlation coefficients. The open responses are subject to subjective interpretation.

The above methodology is being adopted for evaluating all the training programmes of the institute. The results of the evaluation of one such course offered in the disciplines of biochemistry, biotechnology and microbiology are presented here as an illustration. The programme was organized in May–June 2008, in which 24 trainees from various academic institutions in India participated. The sample was homogenous with respect to age and the academic background as standardized during the selection procedure.

The mean pre-training knowledge test score was as low as 8.67 out of a possible score of 25 and the range of scores was wide, i.e. 3.25–17.75. The coefficient of skewness, +1.25 indicated that the scores of the sample were distributed more above the reported mean. The coefficient of variation was as high as 39.11, which indicated a higher variation among the samples on knowledge domain before commencement of the training.

The mean knowledge score in the post-training knowledge test increased to 17.45 after completion of the training and the range narrowed down to 12.5–22. The coefficient of variation (CV) decreased to 15.08, which indicated that the sample was more homogenous as far as knowledge scores are concerned. However, the coefficient of skewness was negative at –0.02, which indicated that the sample distribution was below the reported mean. The *t* value of 1.35 indicated that there was significant difference in mean knowledge scores prior to and after training at 0.01 level of significance.

The mean of the objective-cum-descriptive performance test was 11.78, out of a maximum possible score of 20, with a wide range, i.e. 4.3–18.3. The CV was acceptable at 29.31. Here also, the negative coefficient of skewness of –0.33, indicated a distribution of trainees below the reported mean. The mean rating score of practical skill session judgment was as high as 3.45, with a maximum possible score of 5 and with a low CV of 22.3. Coefficient of skewness was positive at 0.71. The mean on training assignment score was as high as 6.5 out of the maximum possible 10, with a low CV of 12.33, which indicated uniform performance by the sample. Here also the coefficient of skewness was –1.66. The mean score for seminar presentation was 11.85, when the maximum possible score was 20, with a narrower range of 8.8–16.6 and lower CV of 17.89. Here, coefficient skewness was positive at 0.66.

The efficacy of training was measured using a five-point rating scale on four dimensions, namely relevancy to needs, content of training, quality of instruction and practical and overall effectiveness. The mean rating was highest on the 'content' dimension (4.54) and the dimensions of 'relevancy to needs' and 'quality of instruction' followed next with a mean rating score of 4.11.

The training organizational effectiveness index on dimensions of logistics was as high as 0.85. This indicated a high score on qualitative aspects of the training organization. The open responses in efficacy and organizational aspects were also listed and analysed. The most frequent response was to increase time allotted to hands-on training and improving laboratory support and provision of on-campus accommodation for the trainees. With regard

Table 1. Statistical analysis

Statistics	Pre-test (max 25)	Performance test (max 20)	Practical (max 5)	Assignment (max 10)	Seminar (max 20)	Post-test (max 25)
Mean	8.67	11.78	3.45	6.5	11.85	17.45
Range	3.25–17.75	4.3–18.3			8.8–16.6	12.5–22
CV	39.11	29.31	22.3	12.33	17.89	15.08
Coefficient of skewness	+1.25	–0.33	+0.71	–1.66	0.66.	–0.02

Table 2. Correlation analysis

	Pre-evaluation	Performance test	Post-evaluation
Pre-evaluation	1	0.686**	0.528*
Performance test	0.686**	1	0.539*
Post-evaluation	0.528*	0.539*	1

*Correlation is significant at the 0.05 level.
**Correlation is significant at the 0.01 level.

to the satisfaction in terms of expectations, the most frequent response was 'gain in skills' of using various biochemical, biotechnological and microbiological techniques (67%), followed by 'gain and exposure of knowledge' in life sciences (33%). The summary of statistical analysis is given in Table 1.

Multiple correlation analysis was done with the scores obtained in three phases of evaluation pre-test, performance test and post-test. There was a positive and significant correlation between these sets of scores. This is a clear evidence for the validity of the tests employed and the results obtained. The data are given in Table 2.

In this communication, we have analysed the contemporary theories of training evaluation – the need for training evaluation in an ever-changing organizational and development context; the purposes of training evaluation and the evolution of theoretical models of training evaluation, including the widely used Kirkpatrick's model. Training evaluation is important from an institutional, social and economic perspective. The facilitator role and investments on HRD and training of personnel in scientific and reserach organizations are on the rise. We have also dealt with a case study of an academic training programme evaluation conducted at our Institute. The method follows a combination of formative and summative evaluation techniques and is an approximation of Kirkpatrick's model. Multiple methods are employed to determine the knowledge gain, performance of skills and organizational effectiveness of training. The method measures three dimensions, namely reactions (perceptions as to how well the training worked), learning (how well the training worked to transfer the knowledge and skills) and performance (the degree to which learners can apply what they have learned in their spheres of activity). The results indicated that there was significant gain in knowledge due to training, improved practical skills in using various biochemical, biotechnological and microbiological techniques and that the training was highly effective in terms of organizational logistics. One limitation of the study maybe that it does not attempt to interpret the results in terms of impact or cost–benefit of the training, as stated in the fourth-level of Kirkpatrick's model. The training programme evaluated is of purely academic nature, and as an exception since the vision of the course is to provide learning opportunities, the evaluation up to the performance level itself can provide some soft returns in the context of a research organization. The methodology is also continuous and participatory in nature, as it involves the learners and resource persons from the very beginning of the evaluation process and provides equal importance for learner–instructor rapport and open feedback from participants.

1. Mishra, R. K., Management development and training – the Indian experience. Paper presented in IASIA Conference on Public Administration: Challenges of Inequality and Exclusion, Miami, USA, 2003.
2. Lynton, R. and Pareek, U., *Training for Organizational Transformation – For Policy Makers and Change Managers*, Sage Publications, New Delhi, 2000, pp. 15–18.
3. Riech, A. H., Why I teach. *Chron. Higher Educ.*, 1983, **9**, 27–31.
4. Bramely, P. and Newby, A. C., The evaluation of training part I: Clarifying the concept. *J. Eur. Ind. Train.*, 1984, **8**, 10–16.
5. Figari, G., Evaluation of quality aspects of vocational training programmes. *Eur. Center Dev. Vocat. Train.*, 1994, 18.
6. Carnevale, A. P. and Schulz, E. R., Return on investment: accounting for training. *Train. Dev. J.*, 1990, **44**, S1–S32.
7. Holcomb, J., *Make Training Worth Every Penny*, Wharton, Del Mar, CA, 1993.
8. McMahon, F. A. and Carter, E. M. A., *The Great Training Robbery*, Falmer, New York, 1990, pp. 37–39.
9. Hoyle, A. R., Evaluation of training – A review of literature. *Public Adm. Dev.*, 2006, **4**, 275–282.
10. Sefried, E., Evaluation of quality aspects of vocational training programmes. *Eur. Center Dev. Vocat. Train.*, 1998, p. 24.
11. Fiore, K. E. and Rose, S. D., Practical considerations and alternative research methods for evaluating HR programmes. *J. Bus. Psychol.*, 1999, **14**, 44.
12. Massey, O. T., *Evaluating Human Resource Development Programs: A Practical Guide for Public Agencies*, Allyn & Bacon, Massachusetts, 1996.
13. Dixon, N. M., New roots to evaluation. *J. Train. Dev.*, 1996, **50**, 82–86.
14. Gordon, J., Measuring the 'goodness' of training. *J. Train. Dev.*, 1991, **28**, 19–25.
15. Phillips, J. J., A rational approach to evaluating training programs including calculating ROI. *J. Lend. Credit Risk Manage.*, 1997, **79**, 43–50.

16. Phillips, J. J., *Handbook of Training Evaluation and Measurement Methods*, Gulf, Houston, TX, 1991.
17. Kirkpatrick, D. L., Techniques for evaluating training programs. *J. Am. Soc. Train. Dev.*, 1959, **13**, 11–12.
18. Anita, P. B., Becky, F. A. and Mavin, M., *Supervisor–Team Training: Issues in Evaluation*, University of Berkeley, USA, 2006.
19. Fullard, F., A model to evaluate effectiveness of enterprise training programmes, *Int. Enterpreneurship Manage. J.*, 2006, **3**, 263–276.
20. Phillips, J., ROI: The search for best practices. *Train. Dev.*, 1996, **50**, 42–46.
21. Stoel, D., The evaluation heavy weight match. *Train. Dev.*, 2004, **58**, 46–48.
22. Alliger, G. and Janak, E., Kirkpatrick's levels of training criteria: thirty years later. *Pers. Psychol.*, 1989, **42**, 331–342.
23. Alliger, G., Tannenbaum, S. Bennett, W., Traver, H. and Shotland, A., A meta analysis of the relations among training criteria. *Pers. Psychol.*, 1997, **50**, 341–358.
24. Bates, R., A critical analysis of evaluation practice: the Kirkpatrick model and the principle of beneficence. *Eval. Program Plann.*, 2004, **27**, 341–347.
25. Holton, E. F. The flawed four-level evaluation model. *Human Resour. Dev. Q.*, 1996, **7**, 5–21.
26. Swanson, R. A. and Holton, E. F., *Foundations of Human Resource Development*, Berrett-Koehler Publishers, San Francisco, CA, 2001.
27. Newby, T., *Training Evaluation Handbook*, Pfeiffer, San Diego, CA, 1993.
28. Pulley, M., Navigating the evaluation rapids. *Train. Dev.*, 1994, **48**, 19–24.
29. Clark, D., Instructional system development-Evaluation phase. http://www.nwlink.com/~Donclark/hrd/sat6.html#introevaluate, 2007.
30. Johnson *et al.*, Law Enforcement Training in Southeast Asia: A Theory Driven Evaluation. *Police Pract. Res.*, 2006, **7**, 195–215.
31. Wiessner *et al.*, Constructing knowledge in leadership training programmes. *Community Coll. Rev.*, 2007, **35**, 88–112.
32. Report, Directorate of Publications and Information on Agriculture, ICAR, New Delhi, 1997.

Received 11 August 2008; revised accepted 26 November 2008

Soil erosion limits for Lakshadweep Archipelago

Debashis Mandal* and K. P. Tripathi

Central Soil and Water Conservation Research and Training Institute, 218, Kaulagarh Road, Dehradun 248 195, India

Soil loss tolerance limits (*T* value) define the soil loss amounts that are tolerable to maintain, continuously and economically, the sustainability of soil productivity. Within these limits, soil erosion and soil formation processes are in equilibrium. The Lakshadweep Islands is prone to soil erosion and about 20 running kilometre seashore line is being subjected to severe erosion. The unique land and soils of the Lakshadweep Coral Islands require careful management to protect the fragile ecosystem. Soils of ten inhabited islands of Lakshadweep were studied in detail to assign *T* values, for suggesting a conservation plan. The *T* value for the whole Archipelago varied between 7.5 and 12.5 t ha^{-1} yr^{-1}. The spatial delineation of soils with respect to *T* value can facilitate the management of these valuable resources and prevent their degradation.

Keywords: Conservation plan, soil erosion, soil loss tolerance, soil sustainability.

SOIL is an essential natural resource, which is available in limited quantities. Soil functions are mainly in crop production and as a filtering agent indispensable for the maintenance of water quality. In tropical agro-ecosystems, soil erosion is the main land-degradation process, especially if land use is intense[1]. Soil erosion can reduce crop productivity, due either to physical degradation or nutrient depletion[2]. Soil erosion is also an environmental hazard. In this case, the impacts are called off-farm, while silting and pollution of water resources are the major consequences[3]. Erosion limits have to be defined in order to keep these impacts at acceptable levels.

Soil loss tolerance is the maximum rate of annual soil erosion that may occur and still permit a high level of crop productivity to be obtained economically and indefinitely[4]. The *T* value is also sometimes called 'permissible soil loss'. Within these limits, soil erosion and soil formation processes are in equilibrium. Soil loss tolerance depends on the soil type. In very deep and homogenous soils, the effects of erosion will be less pronounced than in shallow soils encountered on highlands of semiarid zones or highly weathered soils whose nutrient storage and availability depend largely on the organic matter of the surface layer[5]. Determination of soil tolerance is intended to compare the expected soil loss with the soil loss tolerance. If soil loss is less than or equal to the soil loss tolerance, soil loss can be still permitted. The maximum soil loss tolerance for tropical regions[5] is 25 t ha^{-1} yr^{-1}. A commonly used soil loss tolerance rate is 5–12 t ha^{-1} yr^{-1} for shallow to deep soils[6,7]. However, the current used rates for tolerable soil loss are far too high for fragile tropical soils with low levels of fertility[5,7]. It has also been indicated that tolerance values for tropical soils have not yet been formulated at the international level[5]. Established annual soil loss tolerance limits[7–9] vary between 0.2 and 11 t ha^{-1} yr^{-1}.

It is important to mention here that soil formation is a positive feedback process, i.e. the product of the process accelerates the production of the product. Therefore soils have to be kept in place to make more of them. The estimated rate of soil loss from farmland has a disastrous consequence for food production. Further, each harvesting removes the plant nutrients from the soil. In a self-

*For correspondence. (e-mail: demichael@rediffmail.com)