

Evaluating training

Fiona Beddoes-Jones

Key learning points

- Evaluation versus validation.
- What the Kirkpatrick model is.
- Applying each level of evaluation to your training in practice.
- Evaluating unconscious learning.
- The importance of goals, time and measurement in evaluation.

Evaluation versus validation

When you evaluate something, you measure it. When you validate something, you are checking that what you are doing is a valid activity – that it's appropriate, suitable, applicable, legitimate and justified. Therefore, with all training interventions, the validation of what you are doing or planning to do must come first. It is completely pointless to evaluate an inappropriate training programme.

What the Kirkpatrick model is

Originally designed in 1959 as part of a PhD dissertation, the Kirkpatrick model was designed

by Donald Kirkpatrick – then Professor of Marketing at the University of Wisconsin – to evaluate the impact of any training intervention. It is based on four hierarchical levels. Although you may not realise it, the training objectives we spend so long writing are all based at one of these four levels:

Level 1: Satisfaction – reactions and impressions

Level 2: Learning – usually conscious learning

Level 3: Behaviour change – application of the learning

Level 4: Results – qualitative, quantitative and financial.

Applying each level of evaluation to your training

Level 1: Satisfaction

This is concerned with the reactions and first impressions of your participants, and is sometimes referred to as the 'happy sheet' stage.

Unfortunately, it often has less to do with the actual value of the learning that takes place and its value to the organisation than it does with the delegates'



satisfaction with the venue and the lunch. This is the lowest level of value. However, it is immediate feedback and, as such, does provide relevant information, particularly if you used focused questions. Be aware, though, that sometimes those training events that participants rave about because they are exciting and novel may not lead to much valuable integrated learning over time. (Some poorly-facilitated, outdoor, competitive team games, thinly disguised as teambuilding spring to mind.) Many trainers ask delegates to rate their responses against a scale of 1 to 5. This makes numerical calculations easier; you will be able to make statements such as 'Ninety-five per cent of participants agreed or strongly agreed that the programme was relevant to their needs at work'.

Tip

Completing a good-quality Training Needs Analysis (TNA) will help you to ensure that you are doing the right thing before you go ahead and waste valuable resources.

Evaluating unconscious learning

Obviously this kind of evaluation is impossible. If you ask a delegate what they have learned and they say that they don't know, then you know that their learning has occurred on an unconscious level. You need to bring whatever it is that they have learned into their conscious awareness. You can do this by asking some good-quality questions such as these:

- So, what was the most valuable/useful/beneficial part of the course / this section, and how can you apply what you have learned to your relationships / job / communication skills?
- If you were to take a guess at what you have learned today that will be useful for you at work, what might that be?

These questions have effectively given their brains both permission to guess and a command to come up with something relevant.

This kind of question is very useful. Often your delegates will come up with something of which they were previously unaware.

The importance of goals, time and measurement in evaluation

Evaluating any training activity is pointless if you don't have either an objective or a base line to measure your results against. Clearly stated aims, objectives, goals, outcomes and outputs are all critically important. After all, how can you participants

perform if they don't know exactly what it is they are meant to be achieving and when? Saying that, in six months' time, you (or, even better, senior management) want each delegate to have found a way of saving their organisation £100K, measured cumulatively over five years, makes it much more likely that worthwhile results will be achieved. Did you notice that this example contains a goal, a time-frame and a quantitative financial measure? The earlier delegates become used to your expectations, the longer their brains have to create the strategies needed to achieve them. State your expectations early, and state them clearly and simply.

You will have noticed that Levels 3 and 4 can only be evaluated over time. Level 3 evaluations often take place after one to three months, whilst Level 4 evaluations can take six months to a year (or even longer) to be fully realised. Remember, though, that if you don't revisit them, these results won't be noticed by anyone. Your training interventions will be forgotten in the mists of time, and all anyone will remember is the venue and the lunch.

References

- 1 D. L. Kirkpatrick, 'Techniques for Evaluating Training Programs', *Training and Development Journal*, vol. 33, no. 6, 1979, pp. 78–92.
 - 2 National Training Awards website: <http://www.nationaltrainingawards.com>
- See also:
- D. L. Kirkpatrick, 'Evaluation', in R. L. Craig (ed.), *Training and Development Handbook*, 4th edition, McGraw-Hill, New York, 1996.
- D. L. Kirkpatrick, *Evaluating Training Programs: The Four Levels*, Pfeiffer Wiley, 1998.

Author

Fiona Beddoes-Jones is the author of the psychometric instruments Thinking Styles®, Think Smart™ and Cognitive Team Roles™. She is a member of the British Psychological Society, the Association of Business Psychologists and the CIPD, and is also a Neuro Linguistics Programming master practitioner. She runs in-house training workshops for trainers who want to develop their training skills.

Telephone: 0044 1476 861010 • E-mail: fiona.bj@cognitivefitness.co.uk • Website: www.cognitivefitness.co.uk

Answers to case study 1

Level 1: Reaction *We don't know precisely. However, we can assume, from the other levels, and because they applied what they learned in practice, that they would have been satisfied or more than satisfied with their training (which in this case was their personal feedback on their Thinking Styles profile).*

Level 2: Learning *They learned how they themselves think, how the other team members think and process information, and the implications this has for debate within meetings.*

Level 3: Behaviour change *Clearly they have all applied their learning. This is evidenced by the shorter meetings and by their feeling that they are better at communicating and relationship building.*

Level 4: Results *The measure of 1.5 hours saved every week by ten senior managers equates to a saving of £25K per annum. Not bad for two days' work. It would be interesting to re-evaluate the meetings after one year, to see if the benefits had been maintained after the inevitably changing group dynamics of leavers and joiners.*

Tip

When designing your Level 1 questions, it is a good idea to include some questions such as these:

- What outcomes did you look for from attending the training course?
- What were your objectives in attending the course?
- Did you discuss your attendance on the course, and your objectives, with your line manager?

There are three key benefits in this approach: Firstly, it will encourage your participants to have clear objectives for attendance and to discuss these with their manager. Secondly, if their response is that they did not achieve their objectives, you will be able to see very quickly if they had any in the first place. Thirdly, when you are evaluating longer-term benefits to the individual and the organisation over time (Levels 3 and 4), you will be able to refer back to their responses and evaluate against them easily, because you have a written record.

Level 2: Learning

This is usually conscious learning. You will often be able to measure participants' learning against course objectives in the form of statements such as: 'By the end of the training, participants will have learned X, Y and Z'. Kirkpatrick defines learning as '[the] principles, facts and techniques that were understood and absorbed by the participants' (p. 82). Evaluation at this level sometimes takes the form of a quiz or test. It is relatively immediate feedback, however; it does not measure whether your participants remember the learning or whether they have integrated it into their daily working lives.

Tip

Depending on the content and purpose of your training, a control group can be used at this level. Direct comparisons can be made between what the training group have learned and can achieve in reality, and what has been learned and achieved by a group that has not been trained.

Level 3: Behaviour change

This is about application of the learning, and is sometimes referred to as the transfer of learning. Many training interventions are planned and delivered with no thought of measuring the learning or knowledge transfer later on.

Changes in attitude can only be evaluated by observing behaviour changes. Verbal accounts cannot be relied upon; some people will say that they believe one thing and then go on to do something that proves the opposite. Soft skills are ideal to evaluate at this level; communication skills learned in one context are quite easily transferred to other areas at work. Because skills like presentation skills and the use of computer software packages take time to develop, I would always include the evaluation of skills development as a Level 3 activity. Some trainers include it under Level 2.

Tip

Learning logs, diaries, narrative accounts and observer accounts can all be used to evaluate behaviour changes at Level 3.

Level 4: Results

This is qualitative, quantitative and financial. Evaluations at this level are the Holy Grail for trainers and are notoriously difficult (but not impossible) to measure. Level 4 results are not limited to the return



on investment (ROI) of your training. They can also include any results which contribute to the effective running of the organisation, and things which are considered beneficial to the business. Customer goodwill, for example, can be measured via a survey and assigned a financial value. After all, it is always included on the balance sheet in a merger, acquisition or management buyout. Be aware that the complexity of multi-variables can make direct financial correlations difficult. You will need to make (and state) your assumptions clearly, and get them agreed by the relevant people, before you make any Level 4 claims.

Tip

If you, or your organisation, want to win one of the UK's four prestigious National Training Awards,² you will need to apply both the training cycle and Kirkpatrick's model, particularly at Level 4. The judges are very keen on entries in which the bottom-line financial benefits have been measured. Reading details of previous winners is a worthwhile activity in relation to evaluating training.

Case study 1

In 2000, the Woodford site of BAe Regional Aircraft were able to cut the length of time they needed for their weekly production meetings, from three hours down to one-and-a-half hours. The ten senior managers who attend the meeting are delighted with the progress they have made since being profiled with the psychometric, Thinking Styles®. Their internal consultant, Irene Foxley, said at the time:

They are much more focused now. They describe themselves as 'sharper' and, because they now understand how each of the others thinks and processes information, they waste much less time in argument and debate. They all feel that their communication skills and their relationships have improved. Instead of halving the meeting time, they have decided to build more into it – a decision they are all more than happy with. Early estimates on the bottom-line benefits to the organisation, calculated against the value of the senior managers' time, are in the region of £25,000 per annum.

Activity

Read Case study 1 again. Which of Kirkpatrick's levels are indicated here and how? (Answers at the end of the article.)

“ It is completely pointless to evaluate an inappropriate training programme ”

Apply each of these levels of evaluation to your training

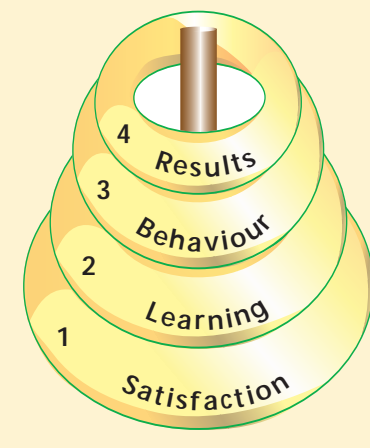


Fig. 1: Levels of Evaluation



Case study 2

Kaizen is the Japanese word for 'continuous improvement'. Measuring continuous improvement is a Level 4 activity. On a first-line supervisors' management programme I ran in a fruit juice factory, some time ago, about the productivity and efficiency improvements the supervisors could make for themselves, we identified that every 20 minutes, on each production line, a carton of fruit juice was opened and tested. This juice was emptied into a clean plastic dustbin which, once full, was emptied down the drain. I asked one of the supervisors if there was

anything wrong with the juice in the bins. There wasn't; it had passed the tests and it was perfect juice. I asked why the juice couldn't be tipped back into the beginning of the process rather than being thrown away. 'What a good idea,' they said. With ten production lines running 24-hour shifts, that one action saved the organisation £500,000 a year. Over five years the cumulative benefit to the organisation, of this one small implemented idea, was in the region of £2.5 million. How I wish I had been paid on a commission basis linked to savings!