

DEFINITION OF ASSESSMENT

'A general term for the processes of ascertaining whether training is efficient in achieving prescribed objectives. It covers both validation and evaluation'.

Training is only as good as the results it achieves. These results can be considered from a variety of viewpoints. Trainers have a tendency to view training through the process itself, and are interested in achieving objectives concerned with learning; the trainees' immediate manager will tend to look at what people can do as a result of the training they receive; and senior management may wish to know whether investment in training was worthwhile. All these viewpoints are relevant and important. Trainers are involved in training, because they enjoy helping people to learn; departmental managers see training to improve performance. Therefore, the view taken of the extent to which 'good training' is being done depends on who is making the assessment - and for what reason. Assessment is carried out by a variety of people, each with a particular point of view. Assessment may be imperfect in that it is often based on subjective opinion rather than objective measurement. However, an even greater danger is that no assessment is done by anybody. The result of this is likely to be that:

- Trainees do not know whether they have achieved the desired outcomes.
- Trainers lack feedback to enable them to improve the quality of training.
- Money may be wasted on inadequate training because no one is aware of its poor quality.
- Management may not associate improved performance with training.
- Spending may be reduced on training because no one knows of the benefits it produces.

Training design should take account of these points to ensure trainees know their achievements and trainers are given feedback to enable them to improve the quality of training. They, along with their client or management can see a link between training and performance.

The techniques of assessment can be used to provide the necessary information. Each of these terms is important when designing training.

FORMATIVE ASSESSMENT

Depending on the training methods being used, there can be many instances where trainees can be given feedback - information about their performance, which will help them to improve. This type of assessment is built into the learning event and becomes an integral part of the learning process: it may not even be recognised as 'assessment'. It is, however, an essential feature of good training as it provides feedback to both the trainee and the trainer. Excellent examples of formative assessment can be found in computer-based training and in group exercises and role-plays.

Computer-based Training and interactive video programmes, both based on the principles of 'programmed learning', include 'remedial loops'. As a result of making an error, the trainee is presented with additional learning to overcome the difficulty. Typically, this is along the lines "Let's consider the situation again from another point of view....", or "Sorry, you've got this one wrong. Let me explain in more detail...". Interactive video can take this a stage further by showing the trainee the consequences of a certain decision and using the situation to provide additional or remedial learning.

This also applies to experiential learning events, based on group exercises or role-plays, where assessment is simply part of the learning experience. This leads to the trainer asking such questions as "How did you feel when Roy said.....?" or "What are the consequences of Roy's actions.....?"

Formative assessment should be non-threatening and designed to help the trainee to learn. There should be no sense of taking a 'test', or the results of assessment being given to the trainee's boss. Trainees should be assured that the results of the test will not be divulged to anyone other than the trainer to evaluate the

progress of learning, give the trainee confidence and, where necessary, adjust the learning process. The concept of Andragogy relies on the sensitive use of formative assessment, which should be 'designed in' to the learning event.

SUMMATIVE ASSESSMENT

The Driving Test is an example of summative assessment. The learning event is over and the trainees' driver has to demonstrate competence. The test will produce a 'pass' or a 'fail' result based solely on performance, not on learning. The use of summative assessment presents considerable design problems. For example:

- It may be difficult to establish suitable standards of performance.
- It is often difficult to devise measures to assess effectively that standards have been achieved.
- The cost of assessment can be high.
- Trainees may react against the prospect of being 'tested'.
- What happens to people who fail the test?

Summative assessment is used at schools and colleges, often resulting in something like 40% of students failing their exams. This arises from deliberate design of the test to produce these results in order, for example, to decide which minority of candidates is to be allowed to proceed to more advanced studies. Training cannot afford failure on anything like this scale. You are employed to design training that will result in improved performance. You are paid to design successful training, based on the need to develop competence to the level required by the client. Failure to improve the trainees' performance to an acceptable standard means that you have failed to meet the client's needs. These are considerations, which should permeate the whole of your design activities. Summative assessment is an important consideration in the design process and may often make a useful or even vital contribution to training. The reason we can travel safely by air is that rigorous standards and tests are used to ensure pilots and ground staff are competent.

NORMATIVE AND CRITERION-REFERENCED ASSESSMENT

Consider the final of the 100 metres in the Olympic Games. Eight athletes are lined up ready to race. They are the finest sprinters in the world. Each trained to be a 'master performer'. In about 10 seconds we have the results of the race, with one 'winner' and seven 'losers'. This is a familiar scene in sport, but it is hardly appropriate in training for work. People are trained, not to compete against each other but to help them improve their performance, as individuals and as members of teams. The reason for training is to enable them to make a better contribution to their organisation. The result of training should be that all are 'winners': training objectives have been achieved, performance improved and the problem solved.

As a designer you may be faced with an environment that is rewarding to 'winners'. Such 'winners' could be sales people being competitively paid to win, or participants on training programmes, which produce the 'Trainee of the Year'. When designing assessment, particularly summative assessment, you are faced with two options. One is to replicate an Olympic final to find out which trainee is the 'best', although you should be sure that you have good reason for needing to know who it is. This is called 'Normative Assessment'. The other is to strive towards a situation in which everybody is a winner, in that everybody achieves the standard, or criterion, for success. This is called 'Criterion-Referenced Assessment'. Discussions with your client should enable you to decide which of these options to use, although for most situations the latter is preferred.

MEASUREMENT CRITERIA

When you are designing and using evaluation measures, it is important to assess the extent to which they conform to three criteria:

1. Technical Accuracy. The measures must accurately and consistently test those characteristics, which they are intended to test.
2. Administrative Convenience. The measures must not place an unacceptable administrative burden upon the organizers or users.
3. Political Acceptability. The measures must not conflict with any professional, functional or status-related expectations of the users.

Ignoring these factors can create resistance to evaluation. This can affect the level of cooperation from the parties involved, which is vital to the success of evaluation, and may lead to its initial rejection or ultimate failure. We will now look at these three criteria in more detail.

Technical Accuracy

Evaluation is measurement. It follows that evaluation can only be as accurate as the measuring instrument used. To obtain accurate measurements we need to use tests and other instruments that are as objective as possible. They should be as free from biased judgement as we can make them. They should be both reliable and valid.

Reliability is concerned with the consistency of the results produced. Therefore, for example, a test administered by two different assessors to the same trainee should produce the same result. In other words, the same test given to the same trainee on two separate occasions (with no additional learning occurring in the intervening period) should produce the same result.

Validity is concerned with whether the test measures what it is intended to measure. For example, if a trainee were given a written examination on the principles of operating a particular piece of equipment in order to assess that persons operating skill, the test would be invalid. It could perhaps be highly reliable but it would be incapable of assessing the trainee's practical ability to operate the machine, as distinct from theoretical knowledge about what should be done. A practical test would be required.

Evaluation, which is both reliable and valid, should be capable of producing consistent results and should measure what it was intended to measure. However, evaluation can be too open to subjective interpretation to be reliable: it could also be objective, but inappropriate. A common mistake is to use a reliable measuring instrument and to assume, because it is objective, that it is also valid. This often happens when written tests are used to measure a practical skill. Skill is the ability to do something and cannot be measured by a test, which will indicate only knowledge about the subject.

Given the choice between objectivity and validity, it is usually better to choose validity. You might, for example, have to rely on a line manager's report on a trainee's improved attitude to customers after a training session, rather than conduct a detailed and scientific observation. You should, however, be aware of the dangers of losing objectivity. An evaluation, which is completely subjective, will not be reliable.

Administrative Convenience

Sometimes, course evaluation questionnaires ask such questions as 'Did the course you have attended produce practical ideas which you can apply in your work?' Ample space being provided for comments.

Although ample space is provided for comments, the questions are closed questions which encourage only 'Yes' or 'No' answers. Revision of the questionnaire by substituting open questions requiring comment that is more detailed would provide much more information. It would also reveal the need for the replies to be discussed with the trainees and their managers. To do so can be very time-consuming and expensive. Given scarce resources, most trainers would cut down on evaluation rather than other aspects of training.

To cope with the realities of the situation, therefore, you have to achieve a balance. You should aim to produce the fairest system of assessment with the highest levels of accuracy that can be achieved with the resources you have been allocated.

Political Acceptability

The third criterion is that evaluation measures should be politically acceptable. You should not implement measures, which create feelings of annoyance and irritation or even rejection on the part of those affected. Measures, which are too frequent and numerous, take too long to complete or ask questions which are seen as too intrusive or not relevant are likely to be unacceptable.

In any case, many people approach tests and questionnaires with suspicion and are not always going to be reassured by your explanations of the purpose behind them. They are likely to form their own view of how much assessment is acceptable, rather than relying on your judgement on the matter.

The people most affected are, of course, the trainees themselves. However, you may extend the assessment to organizational departments by asking the trainees' colleagues, their supervisors, their managers and even their clients about the effects of training. You will then have to persuade those concerned of the fairness, importance and relevance of the assessment you are making. You may need to overcome suspicion of an ulterior motive behind the questions you are asking. In short, you have to balance the requirements of accuracy against what is administratively convenient and acceptable to those affected. Your task is to keep pushing at the boundaries and to use your position to achieve the highest level of accuracy that is achievable in the context.

METHODS OF OBTAINING INFORMATION

You have considered the scope of an evaluation and the importance of linking evaluation with objectives. You have also considered the type of objectives you should be seeking to measure in your evaluation. We now move on to a discussion of the practical methods of obtaining and recording the information you need for evaluation purposes. Such data-gathering tools or instruments are wide-ranging in nature and include the categories described below.

Questionnaires

Questionnaires are a versatile and inexpensive evaluation instruments, and can provide anonymous, standard and measurable replies. However, return rates can be low, questions misunderstood and information falsified to impress.

Tests

Many types of tests can be used in evaluation including oral, objective, performance and essay-type tests. They are reliable and easy to administer and process. However, there are difficulties in designing tests appropriate to objectives and test results must be carefully interpreted to provide reliable conclusions.

Simulations

Simulation is often essential to prevent trainees making dangerous or costly mistakes. They can provide hypothetical problem situations in a safe environment where conditions can be readily controlled and measured. The simulation must be relevant to the work environment and followed up by on-the-job evaluation. Simulations can be time-consuming and difficult to learn from in non-technical situations.

Individual Interviews

Individual trainees are interviewed to assess reactions to or opinions about training, and/or to analyse progress in implementing an agreed action plan. These interviews can be either structured to ask pre-set questions or unstructured to solicit more general information about the training. Interviews may be supplementary to other forms of evaluation and are a flexible, revealing and user-friendly ways of obtaining information. However, there are problems relating to the need for trained interviewers, the cost and time of one-to-one interviews and the possibility that interviewees may feel threatened by a one-to-one situation and may not give their true responses to the training.

Group Interviews

A group of trainees can be interviewed at any stage before, during or after training to identify training needs, progress and outcomes. As with individual interviews, trained interviewers are needed and there is a danger that the more vociferous members of the group could have a disproportionate say. The benefits are that it is a flexible and revealing exercise, and it is less expensive in terms of time and money than carrying out individual interviews. Both group and individual interviews are suitable for evaluating at the levels of reactions, job performance and organisation.

Observations

Trainees can be observed during the performance of a task by trainers, supervisors or peers. To minimize the observer's effect on the trainee, video can be used. The advantages of observation are that actual performance can be assessed in a flexible and, as far as the trainee is concerned, relatively stress-free way. The disadvantages are that the method can only be used for observable performances and it can be unreliable, especially if the observer influences trainees to act in an unnatural way. Observation also requires trained observers, and so can be costly and time-consuming.

Activity Sampling

Activity sampling is a form of observation that takes random samples of the trainee's activities and calculates what proportion of time is spent on various activities over the period of sampling. It is often used in conjunction with other methods to verify the interpretation of the random observations. The method is useful where the time spent on various activities is important because it can provide quantitative measures. However, it is expensive and time-consuming and does not take into account the difficulties of the various activities undertaken.

Critical Incidents

Here supervisors or others familiar with the job are asked to describe the appropriate actions that should be taken in both typical situations and times of crisis. The information is then classified, analysed and

interpreted to identify the areas on which most attention should focus during training. This approach identifies the most critical aspects of job performance, particularly for complex jobs. Like some of the other methods, it is expensive and time-consuming. It is most effective where the trainees have positions with the same, or similar, job content.

Performance Review

Performance review usually involves prolonged or continuous assessment of job performance, generally through discussions between employee and management. The aim is to identify areas that have been improved or may need improvement through training. This has the advantage of involving those with the most interest in the outcomes of training - the trainee and the line manager - but as these individuals may not have specialist knowledge of evaluation techniques, the resulting information may not be reliable.

Recording Information

Every organisation has some system for recording information about the performance of individuals, departments or the organisation as a whole in terms of output, quality of products and services, costs and time. These records exist not principally to evaluate training but as part of the control and management processes of the organisation.

It may be that information arising from the evaluation of training can be stored as part of this overall performance record system. Alternatively, the information may be stored separately, as long as it is accessible to those who need to use it.

Several methods of storing information exist, including: on a manual card index; in appraisal forms kept on the personnel files of individual members of staff; on a computer database. However the information is stored, it should be presented in a form that is intelligible to managers who are not training specialists.

Selecting Methods

Generally, when designing or selecting a specific measuring instrument, the following issues should be considered:

- What data do I need?
- How objective and unbiased are the data likely to be?
- How will I use the data to measure outcomes against objectives?
- How will I analyse the data?
- Who will use the information? Trainers? Trainees? Supervisors? Senior managers?
- Does a suitable instrument already exist?
- Can I adapt an instrument used for a similar purpose?
- Is the instrument going to be cost-effective?

You have probably found that there is no ideal instrument. Your final selection may well be a compromise based on the competing considerations of the various instruments available.

Evaluation at Level 4 'Results Achieved'

You may have noticed that no methods of collecting training information are mentioned for being applicable to the ultimate value level of evaluation. This is because:

- What is measured at the ultimate value level is the combined effect of all management action to achieve the organisation's purposes
- It would be artificial, and could give a false impression of the nature of training, to attempt to isolate its effects at this level
- Often the ultimate value effects are so remote from training as to be virtually immeasurable

- Any attempt to evaluate at this level would probably have to be based on evidence from performance records kept for purposes other than the evaluation training.

Cost-Benefit Analysis

Its relevance to evaluation is considered briefly here. Before embarking on cost-benefit analysis, you need to decide what your purpose is. The purpose might be to:

- Prove that training already completed has had beneficial effects;
- Predict the possible outcome if planned training is implemented;
- Improve general decision-making on training and related matters.

If you want to predict whether training will pay, the basis for the decision should be agreed beforehand. This involves trying to determine whether the return expected from training investment is high enough to match the criteria laid down by management. An analysis of this kind also allows a stricter basis for the definition of training objectives and gives senior management a clearer view of training effectiveness. Criteria that can be used to determine if investment is worthwhile include:

Payback period - how long will it take to pay back the original investment?

Rate of return - what percentage interest is the organisation getting from training investment?
(Rate of Return = Benefits divided by Investment Costs)

Discounted cash flow - bearing in mind the changing value of money, what return would I want in the future for payments I make now?

For example, sales assistants in a large retail organisation consistently made errors when entering sales into cash registers. An average of 32 defective slips per day was returned, totalling 8,000 in a 50-week year. In staff time, corrections cost the company Rs1, 320,000 a year.

A training programme was arranged to improve performance. It consisted of running two sessions of three hours each in the training room for small groups of assistants. The total cost of training including overheads, trainers' preparation time, course evaluation time and the time that the staff were away from work was Rs105, 000. In the first year, after programme completion, a 50 per cent saving in the cost of defective slips was achieved, amounting to RS 660,000 per year. This meant a net saving of RS 555,000.

In this example, the cost benefits are obvious and easy to quantify. This is not always the case. Costing training is relatively easy and straightforward. Serious difficulties arise, however, when we try to quantify the results of training in financial terms. Remember that we must measure results because we cannot take it for granted that there will be financial benefits from training.

The effects of training on the output of jobs such as personal assistant, personnel officer, supervisor or manager are not so directly measurable. We need to translate the 'output' for such jobs into some kind of tangible result - e.g. customer complaints - so that we can assign a money value to it and thus more easily assess any benefits from training. There will remain, however, some intangible outputs such as improved morale, job satisfaction and enhanced reputation, which it would be difficult to quantify.

If your training and development policy sets out the criteria for what it expects from training (as a financial policy does for any other investment), you have to get agreement on whether training can be judged against the same investment criteria as any other business investment. However, you will need to convince senior management that all costs and benefits should be considered, including non-financial benefits. Good decisions are not always based on financial considerations alone, especially where values relating to social benefits or human relations are as important as they are in training.

Analysis of Less Quantifiable Benefits

Some benefits cannot be measured by a cost-benefit analysis. For example, it is unlikely that such an analysis could prove that supervisory or management training is cost-effective. This is because the benefits of training of this type are often, for example, increased morale, better teamwork or greater loyalty. These benefits do not lend themselves to financial quantification.

Certain courses are mandatory by law, such as food hygiene and health and safety. The main purpose of the legislation is to benefit employees and/or the community at large, rather than the organisation. However, it can be argued that the organisation also obtains benefits from such training.

The benefits to be considered may not always be benefits to the organisation. The ultimate value of training may be seen by some for being - at least in part - the benefits it produces for the individual trainee, the community, the industry sector or the country as a whole. Thus, some would attach considerable value to such benefits to the individual as:

- Less stressful work
- More job satisfaction
- Opportunity for increased earnings
- Developing a satisfying career

Where such outcomes are explicitly agreed by senior management as relevant to the organisation, they can be considered appropriate but not amenable to measurement by cost-benefit analysis. So far as commercial organisations are concerned, financial benefits will generally predominate, as the main purpose for their existence is to produce profits for their shareholders. Nevertheless, many commercial organisations would not wish to exclude social benefits of the kind we have been discussing.

In organisations providing services to the general public - e.g. education establishments, hospitals, local authorities and welfare organisations - human or social benefits will predominate. These might include cleaner streets, improved examination results or shorter waiting lists. However, this does not mean that these organisations can ignore financial benefits, as they must ultimately account for the provision of the service to the public, central government or other source of their funding.

It is worth asking whether your client departments would purchase your training from the marketplace if they were allowed to do so. If they would, it is reasonable to assume that your training is beneficial to the organisation. If they would not, an urgent review of training policies is probably called for.

It is important that management should identify those factors, which should be considered, and the measurement criteria to be applied to them, and those factors, if any, which should be ignored.