


EDUCATIONAL
TRAINING
MANUAL
No. 20



EDUCATION STAFF TRAINING DEVELOPMENT PROGRAMME

PROJECT EVALUATION
AND
EVALUATION OF TRAINING



UNITED NATIONS
Economic Commission for Africa
January 1992

**EDUCATIONAL TRAINING
MANUAL NO. 20**

**UNITED NATIONS
ECONOMIC COMMISSION FOR AFRICA
Public Administration, Human Resources
and Social Development Division**

EDUCATION STAFF TRAINING DEVELOPMENT PROGRAMME

PROJECT EVALUATION AND EVALUATION OF TRAINING

DEFINING THE PURPOSE OF THE EVALUATION

Project evaluation involves the ex-post analysis of the functioning, outcomes, and costs of a project. If the evaluation is done as part of midterm monitoring, it will focus on ways of improving the project (for example, by redesigning its mode, curricula, or management) or perhaps on decisions about its continuation, expansion, or replication. If it comes at the end of a project, it will be mainly concerned with the project's success or failure and with drawing lessons applicable to subsequent projects.

The nature of the project is important to evaluation. Are we looking at a training system as a whole, a training institution, or a course? If at a course, is it intended for precareer training or for upgrading? Is it short or long? Off-the-job or on-the-job? The objectives of the project also affect our approach. Are the objectives defined in terms of efficiency, in one sense or another, or of equity? Finally, we need to know from whose viewpoint the analysis is being conducted.

THE NATURE OF THE PROJECT

Evaluating a single precareer course that is entirely off-the-job is a straightforward type of project analysis. Here it is relatively easy to isolate costs and benefits. In the case of a course intended to upgrade skills, workers' performance before and after training may be measured. Evaluating on-the-job training is more difficult because of questions of costing that require special treatment.

Analysis of an entire training institution (unless it offers only one course) is inevitably more complex than evaluating single courses. Many institutions offer courses in a bewildering variety of subject matter, length, skill level (initial or upgrading), format (on-or off-the-job, full-time), and purpose (vocational or nonvocational). In such cases analysis of the demand-supply situation has to cover a wide range of occupations, and outcome-cost analysis, whether concerned with external or with internal efficiency, has to be disaggregated to the level of the individual course. The same is true of analysis of a whole system, unless it consists of only one mode. Thus, evaluation of, say, an on-the-job apprenticeship system can be conducted at an aggregate level, but evaluation of a system that includes several routes has to be disaggregated.

THE OBJECTIVES OF THE PROJECT

The criteria used by a project analyst obviously depend on the objectives of the project, but identifying these can be difficult. Projects often have multiple, sometimes conflicting, objectives. Most projects aim at being efficient, in the sense of achieving the highest possible outcome-cost ratio, but efficiency has several dimensions.

Efficiency can be defined in terms of academic performance-which may not be fully measurable by test or examination scores, including as it does increases in skills,

changes in attitudes, cognitive development, and acquisition of knowledge. We need to know who is supposed to learn what, under what conditions, and by what date.

Efficiency also can be defined in terms of skill on the job after graduation. Measurement of a project's impact in this respect will not be easy in any case, but it will be impossible unless we know what the project was trying to achieve.

Another dimension of efficiency is productivity and income in employment, which may or may not vary directly with the degree of skill acquired by the trainee.

In projects that are not oriented toward the labour market efficiency may be particularly difficult to define. Gains in academic achievement or acquisition of skill may be among the results, but the project may also have important, not easily measured, effects on the morale or well-being of the participants. (Adult literacy classes are an example).

Finally, a project's efficiency may be defined partly in terms of its indirect or spill-over effects, such as a change in the role of women or the wider impact on the labour market of changes in the supply of skills.

For some projects success is measured in terms of their impact on equity. Their objective is to improve the relative position of a given underprivileged group with respect to any or all of the effects discussed above.

Whatever the objectives of a project, the analyst needs to have them clearly set out before he can select his criteria for evaluation. Where there are multiple objectives, and particularly if they conflict, it may be necessary for him to assign weights to them to arrive at an overall assessment.

THE POINT OF VIEW

The criteria used in the analysis also depend on the viewpoint that is adopted. Are we looking at the project from the point of view of the nation, or of society as a whole (the most usual practice)? From the narrower budgetary point of view of the government or the treasury? Or from the perspective of the aid agency involved in financing the project, the local community, or the firm that sponsors a trainee or an individual student? Taking any of these viewpoints can be useful for different purposes, and each implies a variation in method.

THE INITIAL CHECKLIST

In order to be clear about the purpose of a project analysis, it is useful to fill out an initial checklist (table 1.1). In the following chapters we will assume, for purposes of exposition, that we have checked the following items: A.2, B.1.b, B.2.a, B.3.a, B.4.a, C.1.a-c, D.1. In other words, we assume that we are conducting an evaluation of a training institution that offers long-term precareer, entirely-off-the-job courses; that the project's objectives include efficiency in all except the non-labour-market and spill-over senses; and that we are appraising it from the point of view of society as a whole.

Table 1. Initial Checklist: Purpose of Project

Instructions: Check appropriate box.

- A. Stage of evaluation
 - 1. Midterm
 - 2. Final
 - 3. Other (regular or periodic)
- B. Description of project
 - 1. Scope
 - a. Whole system
 - b. Multicourse training institution
 - c. Single course training institution
 - d. Single course
 - 2. Length of training
 - a. Long (more than three months)
 - b. Short (up to three months)
 - 3. Focus
 - a. Precareer
 - b. Upgrading
 - 4. Mode
 - a. Entirely off-the-job
 - b. Entirely on-the-job
 - c. Mixed

- C. Project objectives
 - 1. Efficiency
 - a. Academic performance
 - b. Skill on the job
 - c. Productivity in employment
 - d. Non-labour-market
 - e. Spill-over
 - 2. Equity
- D. Point of view
 - 1. Nation or society
 - 2. Treasury
 - 3. Aid agency
 - 4. Local community
 - 5. Firm
 - 6. Individual student

INTERRELATIONS WITH INDUSTRY

- . Training and employment
- . Formal links and services
- . Industrial links of staff
- . Industry-like environment

THE QUANTITATIVE EVALUATION

Quantitative indicators of efficiency of operations include student flow rates and performance, staff load, provision and utilization of facilities and resources, and breakdowns of training costs.

STUDENT FLOW RATES

The most useful measures of student flow are

- . Admission rate (the proportion of applicants admitted to the ceruse)
- . Dropout rate (the number of students who leave during the ceruse without taking final tests or examinations, as a proportion of students enroled at the beginning of the course)

- Repetition rate (the number of students who repeat a stage of training as a proportion of the students enrolled in that stage in the previous year)

- Pass rate (the number of students completing the course successfully as a proportion of the students enrolled in the final year or stage of the course).

STUDENT PERFORMANCE (EFFICIENCY INDEXES)

An overall indicator of internal efficiency in terms of student performance is found by dividing the number of graduates by the number of students entering at the beginning of the course to yield a percentage. If the amount of repetition is significant, however, it is more useful to employ a measure that indicates how much additional time over the planned time is required to produce graduates. Examples are

- Average time required to produce a graduate (total student-years spent on training (including time spent by dropouts) divided by number of graduates produced; this can then be compared with planned time

- Output-input ratio (the number of graduates, multiplied by planned course length in years, as a proportion of the total number of student-years spent in training).

STAFFING

Indicators include

- Student-teacher ratio (by course or for the institution)

- Average class size (preferably separately for classroom work and laboratory or workshop activities).

EVALUATING THE EFFICIENCY OF OPERATIONS

This section deals with the efficiency of operations (the internal efficiency) of a technical school or vocational training centre. It examines the main aspects of internal operations, including the content and relevance of the courses, the teaching methods, the quality of the staff, the adequacy and utilization of space, the appropriateness and use of equipment, the effectiveness of management, and the interrelations with industry.

Evaluation of the efficiency of operations (accompanied by an understanding of the underlying factors) is a powerful management tool, both for the director or principal of the school and for the ministry of agency responsible for supervising the institution. It is particularly important when a system or institution is being expanded or when measurements or external efficiency indicate deficiencies in the system that call for improvements in operations. Evaluation also provides a basis for comparing the performance of different institutions.

Assessing efficiency, and in particular identifying inefficiencies and their causes, require subjective judgment as well as quantitative analysis. The diagnosis must be carried out by an experienced technical educator, and it requires the full involvement and cooperation of the senior staff of the institution concerned.

The process of evaluation starts with identification of the key factors that are commonly recognized as being of primary importance in determining the efficiency of a system. In this chapter we identify those key factors and describe how they are assessed through the use of checklists, questionnaires, and other means.

A training programme can be judged only after observing the teaching process, including practical laboratory and workshop activities, to assess the method and quality of instruction. Examination results provide data for the quantitative study of internal efficiency, but the quality of the examination methods to the training must also be assessed. The source and quality of students and trainees and the processes by which they are selected and counselled at entry and assisted to find appropriate employment after training ends should also be evaluated.

An important factor in determining efficiency is the quantity and quality of teaching staff. The evaluation team will need to look at the procedures for selection, appointment, assignment, and promotion of staff, as well as their salaries and other conditions of service. The quality of staff is assessed and deficiencies and training needs are identified by studying personnel records that show qualifications and experience and by observing teachers' performance, the quality of the teaching programmes and materials, and the students' work. The utilization and performance of support personnel, particularly laboratory and workshop employees and professional staff such as librarians, must also be evaluated.

The quality and effectiveness of training can be affected significantly by the adequacy and utilization of physical resources (buildings, equipment, and materials). Overprovision and under-utilization are as inefficient as under-provision, with its consequent overcrowding and inadequate facilities. Here, as for the other key factors, the evaluation team must weigh carefully the quantitative and qualitative assessments. Facilities can be measured in unit areas of teaching space or unit costs of equipment; the actual utilization of space or equipment can be compared with the maximum theoretical use to yield utilization factors. For each quantitative measure there are broad values or

norms that are accepted internationally as good practice. However, it must be kept in mind that an apparently adequate quantitative figure may conceal inefficient or ineffective procedures. For example, a reasonable average level of utilization (say, 75 percent for workshops) may represent extreme overcrowding for part of the time and zero use for the remainder. Workshops and equipment may be in use most of the time, but the training exercises and activities may be of poor quality. The expenditure on equipment may be reasonable, but the equipment may be inappropriate for the objectives of the training programme.

The effectiveness of the school's management is evaluated by examining the organizational structure and the management style and effectiveness of the senior staff. The latter can be judged only qualitatively, but that judgement will be illuminated by the evaluation of other key factors; that is, good or poor ratings in such areas as utilization of facilities are part of the supporting data for judging management effectiveness.

The interrelationship of the training institution with industry is probably the most important single indicator of its efficiency and effectiveness. This interrelationship can be measured in terms of employment of trainees, formal links between the institution and industry, the staff's industrial experience and connections, and the extent to which the institution engages in production or other practical activities and creates an environment similar to that of industry. (These measures overlap with some used in determining external efficiency.)

Normally, a training centre or school is part of a vocational training or technical education system and is responsible to a government organization such as the ministry of labour, manpower, or education. The system may impose constraints or conditions that have a major influence on the operations of the institution. For example, in an extremely centralized system in which curricula, courses, and training material are developed centrally, teachers are recruited and appointed centrally, and all materials are procured and distributed from a central office, there may be little scope for the school or centre to control its own efficiency. It is therefore necessary in any evaluation to look at relevant aspects of the national system.

We have mentioned costs only incidentally, since they are normally included under external efficiency. However, patterns and levels of expenditure on training activities-not just raw numbers-should be taken into account in evaluating operations. For example, a comparison of cost-effectiveness of two similar training institutions could give misleading indications of effectiveness if one centre achieved lower costs over the short term by, for example, failing to replace equipment or restock consumable materials, or neglecting maintenance. The checklists and worksheets include questions designed to identify satisfactory or unsatisfactory budgetary practices.

THE QUALITATIVE EVALUATION

Experience has shown that certain key factors, summarized in the outline below, have a primary influence on the overall efficiency of an institution. The checklist uses these factors as a systematic basis for guiding the evaluation team in making value judgements on the quality of the institution's operations. Those judgments must be made by experienced technical educators and must take into account the environment and objectives of the project, as described earlier. The goal is to build up a profile that shows the state of health of the institution with respect to each key factor so that deficiencies can be identified and the scope and nature of remedial action can be determined.

Level, Content, Quality, and Relevance of the Training Programme

- . The format and content of the curricula and syllabuses
- . The implementation of the courses: the teaching process, methods, materials used, and training activities
- . The examination scheme; content and conduct of examinations.

STUDENTS

- . Selection methods, entry qualifications, and sponsorship
- . Counselling, guidance, placement, and follow-up
- . Student-staff relationships.

STAFFING AND STAFF DEVELOPMENT

- . Staffing policy, salaries, and other conditions of service
- . Selection and qualifications of staff
- . Size and quality of staff; turnover
- . Size, quality, and salaries of support staff
- . Staff development plans; training (pre-service and in-service)

PHYSICAL RESOURCES

- . Range, areas, and layout of accommodations
- . Facilities, services, and maintenance
- . Range, relevance, and adequacy of equipment
- . Equipment use, maintenance, and repair
- . Replacement and updating of equipment
- . Use, replacement, and storage of consumable materials

ORGANIZATION AND MANAGEMENT

- . Institutional development plan and objectives
- . Organizational structure and responsibilities
- . Management information system: availability and use
- . Management style and effectiveness.
- . Average teacher workload (normally expressed as teaching hours or contact hours per week).

FACILITY SCHEDULES AND UTILIZATION OF SPACE

Indicators include

- . Average area of workspace (area of classrooms, or laboratories, or workshops, divided by the normal working capacity)
- . Average areas of support spaces (area of library, communal spaces, living accommodations, and the like divided by number of students using each kind of area)
- . Space utilization (the actual student occupancy of total teaching space as a proportion of the total capacity of the teaching space).

COSTS

the most important cost measure is the cost per student per year or cost per graduate. Other analyses of cost also provide valuable comparative data:

- . Staff salaries as proportion of total cost
- . Cost per student per year for consumable materials
- . Maintenance cost as proportion of capital costs.

THE CENTRAL TRAINING SYSTEM

the section deals with methods for evaluating the main aspects of the central government training system to gain an understanding of the context within which the training institution operates. Such an evaluation can also provide the basis for a study of the technical and vocational education system as a whole.

The key factors in evaluating the central system are:

- . Policies, planning, and development
- . Central and regional control
 - Training programmes and course
 - Staffing
 - Physical resources
 - Finances
 - Management organization

SUMMARY ASSESSMENT AND PROCEDURES

The main purpose of this manual is to provide management with a diagnostic tool for identifying areas that require improvement. However, it can also be used to provide a profile of an institution so that different institutions within a training system can be compared with each other or with a "standard" institution.

The summary assessment (Table 2) is meant to facilitate both uses by telescoping the data gathered through questionnaires and observations into a convenient numeric profile. The summary is organized according to indicators of internal and external efficiency. Since no ranking or weighting of the key factors exists, the summary cannot be made much more compact-it is not possible to arrive at a single measure of the

efficiency and effectiveness of a training institution. Instead, the summary provides a multidimensional picture of an institution and permits comparisons between schools.

It should be noted that the overall assessment is not an arithmetic average of all the items under the key factor. The evaluators should weigh the individual responses and make an intuitive judgment of the overall assessment.

The guidelines in this section will be particularly useful for ministries or agencies which have not yet conducted regular evaluations of their technical schools or vocational training centres. We have tried to achieve an appropriate balance of detail and coverage so that the evaluations can be carried out with reasonable expenditures of time and manpower and still provide a systematic and comprehensive evaluation of the performance of any institution. The guidelines are not rules. We expect that in practice they will be modified to suit local circumstances and purposes. The scope of the evaluation and the depth of detail will be decided with an eye to the cost, the staff available, and the use to be made of the results. In the same way, the procedures for conducting the evaluation will vary according to circumstances.

The technical and vocational education system may include tens or hundreds of institutions. If the number is large, we recommend that as a first stage a sample of about ten, representative in terms of type, size, location, and so on, be chosen for initial evaluation. These initial evaluations serve several purposes. They quickly lead to the identification of any common problems or issues. They also allow the guidelines to be adapted and questionnaires and instruments to be modified to suit the agency's needs. If the agency is attempting evaluation on a significantly more detailed basis than in the past, these initial evaluations also provide a training experience for the staff.

Table 2 Summary Assessment

EFFICIENCY OF OPERATIONS

Qualitative Indicators

Content and quality of courses	1	2	3	4	5
Format and content of curricula and syllabuses	1	2	3	4	5
Quality of curricula and syllabuses	1	2	3	4	5
Review and revision of curricula and syllabuses	1	2	3	4	5
Teaching methods	1	2	3	4	5
Examinations and assessment	1	2	3	4	5

Overall assessment:

Students or trainees	1	2	3	4	5
Entry and selection	1	2	3	4	5
Counselling and career guidance	1	2	3	4	5

Staff-student relationships	1	2	3	4	5
-----------------------------	---	---	---	---	---

Overall assessment:

Staffing and staff development	1	2	3	4	5
Staffing policies	1	2	3	4	5
Selection and qualifications	1	2	3	4	5
Staff development and training	1	2	3	4	5
support staff	1	2	3	4	5

Overall assessment:

Physical resources	1	2	3	4	5
Teaching facilities	1	2	3	4	5
Support facilities	1	2	3	4	5
Utilization of space	1	2	3	4	5
Provision of equipment	1	2	3	4	5
Equipment utilization	1	2	3	4	5
Consumable materials	1	2	3	4	5

Overall assessment:

Organization and management	1	2	3	4	5
Objectives and plans	1	2	3	4	5
Organizational structure	1	2	3	4	5
Information system	1	2	3	4	5
Management style and effectiveness	1	2	3	4	5

Overall assessment:

Interrelations with industry	1	2	3	4	5
Training and employment	1	2	3	4	5
Formal links and services	1	2	3	4	5
Industrial links of staff	1	2	3	4	5
Industrial environment	1	2	3	4	5

Overall assessment:

Average time required to produce a graduate/planned time

Output-input ratio

Cost per student per Year^{1/}
Cost per graduate
Cost per student/cost per student in general high
school

EXTERNAL EFFICIENCY (OUTCOMES AND COSTS)^{2/}

Employers' opinions on employability of graduates
Preferred to other applicants
Same as other applicants.
Less qualified than other applicants
No opinion

Teachers' opinions of graduates
Well prepared for a good job
Only adequately prepared for a job
Not well prepared for a job

Graduates' opinions on their ability to secure
employment in their field of training
With my training it is easy to get a job
With my training it is not easy to get a job
With my training it is very difficult to get a job

Employment rate of graduates one year after graduation
Proportion of those employed working in the same or
similar fields for which they were trained
Rate of return of the programme

The early evaluations should also be designed to help the evaluators arrive at a reasonable consensus of standards through exchange of information and by varying the membership of teams.

If the ministry or agency is establishing regular evaluation as part of its management review process, it is advisable to set up a small unit responsible for the administration of evaluation. However, it is normally better not to create a team of full-time evaluators, but instead to create a panel or pool of senior ministry staff that also

^{1/} For short courses use cost per student per hour.

^{2/} Based on interviews with employers, teachers, and graduates, and on quantitative indicators (where data are available) on employment and earnings of graduates.

includes active principals or directors and, if possible, representatives from industry or technical ministries.

INTERNAL EFFICIENCY

Each institution to be evaluated is given a set of questionnaires to be completed in advance, together with a general note prepared by the agency explaining the purpose of the evaluation. If possible, preliminary briefing meetings are held with the directors of the institutions and their senior staff so that the questionnaires can be distributed and the purpose of the evaluation explained. The briefings should stress the importance of evaluation in planning and its value in identifying needs for resources. The full cooperation of staff should be obtained; subsequently, throughout the conduct of the evaluation, the team members must be on their guard to avoid acting as inspectors or inquisitors. Opportunities should be taken to exchange experiences and provide advice as well as to gather information.

The evaluation of the school or centre takes place one to four weeks after the questionnaires are distributed. The evaluation team should include at least two persons: a technical and vocational educator whose experience and knowledge cover both developing and developed countries, and a person with a background in economics, particularly manpower and labour market economics. The evaluation of specialized training programmes may require the addition of relevant specialists. The team will establish its own work pattern.

DAY 1. An Initial meeting with the principal or director and key senior staff is held. The data, reports, and completed questionnaires that were requested in advance are presented to the team, and problems or missing data are identified (about two hours).

The team then makes a brief tour of the institution to gain familiarity with the main facilities and layout (about one hour).

Next, the team reviews the completed questionnaires and clarifies any apparent anomalies or errors in the answers. It then proceeds, in discussion with the director or principal and the senior staff, to complete questionnaires (about three hours). These questionnaires are to be used as a basis for structured interviews and need not be rigidly followed. To preserve confidentiality, the team may wish to interview the director alone to complete his/her questionnaire, but senior staff should be brought in for the subsequent discussions, both to take advantage of their knowledge and to give them experience in the process of conducting evaluations.

DAY 2. The evaluation team proceeds to the detailed observation of teaching and training activities and use of equipment and facilities. If there is a relatively large number of laboratories or workshops, it is advisable to choose only a sample for detailed

study. During this period in the classrooms, laboratories, and workshops the evaluation team should review critically the content and treatment of the courses or subjects (about four hours).

The evaluation team may conduct its interviews of teaching staff and trainees while carrying out the observations in the classrooms and workshops, if, however, only a relatively small sample of the activities is covered, it would be better to arrange for more broadly based interviews with group of six to eight staff members and a like group of trainees.

At this stage the evaluation team should be able to complete its assessments of the key factors covering the qualitative and quantitative aspects of internal efficiency (a bout three hours).

A third day may be required if a very large institution (2,000 or more training places and a wide range of training programmes) is being evaluated.

The evaluation team concludes its visit to the institution with a brief presentation to the senior staff of the main findings.

It will be clear from the above that the evaluation exercise is significantly more searching and revealing than the normal institutional visit, and only with practice with the team acquire the skills necessary to carry out the interviews and assemble the data. Since the team is required to make value judgments about quality of training, it is vital that the members' experience and background give them adequate status. Even so, they may find it difficult to make judgments about the quality of teaching in subjects that are outside their own specializations. In that case it would be advisable to form slightly larger teams covering a representative range of the technical specializations offered in the institution.

EXTERNAL EFFICIENCY

The evaluation of external efficiency involves the collection of data from employers and graduates of the training programmes. Arrangements for the interviews should be made beforehand, either by the local representatives of the responsible agency or by the director of the institution. It is useful for a staff member of the institution being evaluated (preferably the guidance counsellor or the person in charge of liaison with industry) to be present.

THE SAMPLE. Only a sample of firms should be singled out for interviews, that is, six to either in the city where centres or schools are being evaluated. The firms should be selected on the basis of two criteria: high probability of employing graduates from the institutions being evaluated, and size (large, medium, and small firms should be represented). The information is collected in interviews, initially with the personnel

officer, then with foremen and with graduates of the school or recent trainees. The limited purpose of the evaluation should be made clear from the outset, and firms will usually give full cooperation.

THE QUESTIONNAIRE. As discussed, interviews and questionnaires, although costly and time-consuming, are extremely helpful in assessing the effectiveness of the training from the points of view of the trainee and the employer.

In planning a questionnaire survey, the first issue to be resolved is what groups to interview. At a minimum, graduates, dropouts, and supervisors of the graduates should be interviewed. However, if time and resources allow, teachers, directors of programmes, and other employers could also be included.

The second issue is what areas to cover. In general, the questionnaires should include the characteristics of the respondent, the effects for training on employment and earnings, assessment of the training programme, and assessment of the graduates.

For inquiry through questionnaires to succeed, efficient management and the cooperation of everyone concerned are required. In addition, attention must be given to the following considerations.

Drafting the questions requires great care. It demands a good knowledge of the language and culture of the interviewees, as well as some technical knowledge of the skill or occupation of interest. Whenever possible, questions should not be open-ended, since this invites a wide range of subjective replies and comments that are not comparable. Specific alternative replies to each question should therefore be offered to respondents.

Better-educated respondents may be able to complete a well-made questionnaire by themselves, and budget limits may necessitate this shortcut. but in general, personal interviewing, although costly, yields the most complete and useful responses, since the interviewer is able to clear up misunderstandings and ask follow-up questions. At the very least, personal delivery and collection of the questionnaires, despite the expense, is preferable to expecting people to reply to an impersonal survey by mail.

Under ideal circumstances a significant statistical sample of former trainees should be drawn. When availability of funds determines the size and structure of the sample and the questionnaire or interview procedure, the only option may be to restrict the interviews to graduates at their place of work. This precludes interviewing unemployed trainees.

Training and instructions for interviewers must be careful and detailed to ensure uniform interpretation of the questions and the smooth evolution of the interview.

Finally the answers have to be coded, tabulated, and cross-tabulated so that the evaluator can draw the pertinent conclusions. Sufficient resources, time, and people must be allocated to do this routine but essential chore properly.

Similar answers to the same question, but from different viewpoints, reinforce conclusions. Conversely, answers that differ from one group to another may cast doubt on the validity of a single conclusion, and further evidence may be required. When graduates of the school are being appraised, a supervisor may be asked to rate graduates compared with non-graduates and a graduate may be asked to rate fellow graduates compared with other workers in his group. Similarly the graduate's assessment of the quality of teachers should be supplemented by the opinions of administrators. Because of the variety of institutional set-ups and the different characteristics of particular national and cultural environments, it is impossible to spell out before hand the exact questions to be asked and the manner in which they are posed. But some basic issues are common to all evaluations and must be covered in questionnaires and interviews.

EVALUATION OF TRAINING

Effective training directors will make an effort to evaluate all their training activities. The success of these efforts depends to a large extent on a clear understanding of just what "evaluation" means. This paper will attempt to accomplish two objects: (1) to clarify the meaning of evaluation and (2) to suggest techniques for conducting the evaluation.

These objectives will be related to "in-house" classroom programmes, one of the most common forms of training. Many of the principles and procedures can be applied to all kinds of training activities such as performance review, participation in outside programmes, programmed instruction, and the reading of selected books.

EVALUATION CLARIFIED

Nearly everyone would agree that a definition of evaluation would be " the determination of the effectiveness of a training programme." But this has little meaning until we answer the question: In terms of what? We know that evaluation is needed in order to improve future programmes and to eliminate those programmes which are ineffective. The problem is how to begin.

Evaluation changes from a complicated, elusive generality into clear and achievable goals if we break it down into logical steps. These steps can be defined as follows:

- Step 1: Reaction. How well did the conferees like the programme?
- Step 2: Learning. What principles, facts, and techniques were learned?
- Step 3: Behaviour. What changes in job behaviour resulted from the programme?
- Step 4: Results. What were the tangible results of the programme in terms of reduced cost, improved quality, improved quantity, etc.?

With this clarification of the meaning of evaluation, training directors can now begin to pinpoint their efforts at evaluation. They better realize what they are doing, and they recognize the limited interpretations and conclusions that can be drawn from their findings. As they become more experienced and sophisticated in evaluation design and procedures, they slowly begin to obtain more meaningful results on which future training can be based.

These four steps will now be defined in detail with examples and suggested guideposts. It is important to stress that the described procedures and techniques can be used in almost any organization. It is also important to stress that the results from one organization cannot be used in another organization. Obviously, there are many factors that would influence the results. These variables include the group, the conference leader, and the approach to the subject.

STEP 1: REACTION

Reaction may best be defined as how well the trainees liked a particular training programme. Evaluating in terms of reaction is the same as measuring the feelings of the conferees. It is important to emphasize that it does not include a measurement of any learning that takes place. Because reaction is so easy to measure, many training directors do it.

GUIDE FOR EVALUATING REACTION

1. Determine what you want to find out.
2. Use a written comment sheet covering those items determined in step 1 above.
3. Design the form so that the reactions can be tabulated and quantified.
4. Obtain honest reactions by making the forms anonymous.
5. Allow the conferees to write in additional comments not covered by the questions that were designed to be tabulated and quantified.

STEP 2: LEARNING

It is important to recognize that a favourable reaction to a programme does not assure learning. All of us have attended meetings in which the conference leader or speaker used enthusiasm, showmanship, visual aids, and illustrations to make a presentation well accepted by the group. A careful analysis of the subject content would reveal that the speaker said practically nothing of value-but said it very well.

LEARNING DEFINED

There are several definitions of learning. For the purpose of this chapter, learning is defined as follows: the principles, facts, and skills which were understood and absorbed by the conferees. In other words, it does not include the on-the-job use of these principles, facts, and skills. This application will be discussed later in this paper in the section on behaviour.

GUIDEPOSTS FOR EVALUATING IN TERMS OF LEARNING

Several guideposts should be used in establishing a procedure for measuring the amount of learning that takes place:

1. The learning of each conferee should be measured so that quantitative results can be determined.
2. A before-and-after approach should be used so that any learning can be related to the programme.
3. As far as possible, the learning should be measured on an objective basis.
4. Where possible, a control group (not receiving the training) should be compared with the experimental group which receives the training.
5. Where possible, the evaluation results should be analyzed statistically so that learning can be proved in terms of correlation or level of confidence.

These guideposts indicate that evaluation in terms of learning is much more difficult than evaluation in terms of reaction, as described earlier. A knowledge of statistics, for example, is desirable. In many cases, the training department will have to call on the assistance of a statistician to help plan the evaluation procedures, analyze the data, and interpret the results.

SUGGESTED METHODS

Classroom Performance It is relatively easy to measure the learning that takes place in training programmes that are teaching skills. The following programmes would fall under this category: job instruction training, work simplification, interviewing skills, induction techniques, reading improvement, effective speaking, and effective writing. Classroom activities such as demonstrations, individual performance of the skill being taught, and discussions following a role-playing situation can be used as evaluation techniques. The training director can organize these in such a way that he or she will obtain a fairly objective evaluation of the learning that is taking place.

CONCLUSIONS ABOUT LEARNING

It is easy to see that it is much more difficult to measure learning than it is to measure reaction to a programme. A great deal of work is required in planning the evaluation procedure, in analyzing the data that are obtained, and in interpreting the results. Wherever possible, it is suggested that training directors devise their own methods and techniques. As has been pointed out in this section, it is relatively easy to plan classroom demonstrations and presentations to measure learning where the programme is aimed at the teaching of skills. Where principles and facts are the objectives of the training programme, it is advisable to use a paper-and-pencil test. Where suitable standardized tests can be found, it is easier to use them. In many programmes, however, it is not possible to find a standardized test, and training directors must use their skill and ingenuity in devising their own measuring instruments.

If training directors can prove that their programmes have been effective in terms of learning as well as in terms of reaction, they have objective data to use in selling future programmes and in increasing their status and position in the company

STEP 3: BEHAVIOUR

Robert Katz, professor at Dartmouth, wrote an article in the July-August 1956 issue of the Harvard Business Review. The article was called "Human Relations Skills Can Be Sharpened." In it he stated that those who want to change their job behaviour must meet five basic requirements:

1. They must want to improve.
2. They must recognize their own weaknesses.
3. They must work in a permissive climate.
4. They must have some help from someone who is interested and skilled.
5. They must have an opportunity to try out the new ideas.

It seems that katz has put his finger on the problems that exist in a transition between learning and changes in behaviour on the job.

Evaluation of training programmes in terms of on-the-job behaviour is more difficult than the reaction and learning evaluations described in the two previous sections. A more scientific approach is needed, and many factors must be considered. During the last few years a number of attempts have been made, and more and more effort is being put in this direction.

Several guideposts are to be followed in evaluating training programmes in terms of behavioral changes:

1. A systematic appraisal should be made of on-the-job performance on a before-and-after basis.
2. The appraisal of performance should be made by one or more of the following groups(the more the better):
 - a) The person receiving the training
 - b) The person's superior or superiors
 - c) The person's subordinates
 - d) The person's peers or other people thoroughly familiar with his or her performance
3. A statistical analysis should be made to compare performance before and after and to relate changes to the training programme.
4. The post-training appraisal should be made three months or more after the training so that the trainees have an opportunity to put into practice what they have learned. Subsequent appraisals may add to the validity of the study .
5. A control group (not receiving the training) should be used.

CONCLUSIONS ABOUT BEHAVIOUR

These are some of the best approaches that have been used to measure effectiveness of training programmes in terms of on-the-job behaviour. Only the methods and instruments used in these studies have been mentioned. The results, although interesting, cannot be borrowed by other training directors, but the techniques can.

For those interested in evaluating in terms of behavioral changes, it is strongly suggested that these studies be carefully analyzed. The references in this chapter indicate where the detailed articles can be found.

Once more I would like to emphasize that the future of training directors and their programmes depends to a large extent on their effectiveness. To determine effectiveness, attempts must be made to measure in scientific and statistical terms. Measuring changes in behaviour resulting from training programmes involves a very complicated procedure. But it is worthwhile and necessary if training programmes are going to increase in effectiveness and their benefits made clear to top management.

It is obvious that very few training directors have the background, skill, and time to engage in extensive evaluations. It is therefore frequently necessary to call on industrial psychologists, research people and consultants for advice and help.

STEP 4: RESULTS

The objectives of most training programmes can be stated in terms of results such as reduced turnover, reduced costs, improved efficiency, reduction in grievances, increase in quality and quantity of production, or improved morale, which, it is hoped, will lead to some of the previously stated results. From an evaluation standpoint, it would be best to evaluate training programmes directly in terms of results desired. There are, however, so many complicating factors that it is extremely difficult, if not impossible, to evaluate certain kinds of programmes in terms of results. Therefore, it is recommended that training directors evaluate in terms of reaction, learning, and behaviour.

Certain kinds of training programmes, though, are relatively easy to evaluate in terms of results. For example, in teaching clerical personnel to do a more effective typing job, you can measure the number of words per minute on a before-and-after basis.

CONCLUSIONS ABOUT RESULTS

the evaluation of training programmes in terms of "results" is progressing at a very slow rate. Where the objectives of training programmes are as specific as the reduction of accidents, the reduction of grievances, and the reduction of costs, we find that a number of attempts have been made. In a few of them, the researchers have attempted to segregate factors other than training which might have had an effect. In most cases, the measure on a before-and-after basis has been directly attributed to the training even though other factors might have been influential.

SUMMARY

One purpose of this chapter is to stimulate training people to take a penetrating look at evaluation. Their own future and the future of their programmes depend to a large extent on their ability to evaluate and use evaluation results.

Another objective has been to clarify the meaning of evaluation. By breaking it down into reaction, learning, behaviour, and results, the training person can begin to do

something about it and can gradually progress from a simple subjective reaction sheet to a research design that measures tangible results.