MONOGRAPH 4

Evaluation in science and technology education

4.1 Overview

Evaluation is an integral part of the teaching–learning process. Through systematic evaluation procedures science and technology teacher educators can measure the progress of pre-service and in-service teacher trainees in attaining the objectives of teaching programmes. Evaluation also enables teacher educators to measure the effectiveness of their own teaching strategies and methods. Similarly, trainee teachers need to develop skills and competence in evaluation techniques which they can then apply in their own teaching in schools. Evaluation results can also play a part in an accountability system for ensuring that institutions, teacher educators and teachers deliver good quality science and technology education.

It is important that the objectives of a lesson or course should be clear, so that science and technology teacher educators can select appropriate learning experiences, teaching methods and information-gathering techniques in order to achieve the objectives.

4.2 Objectives

This monograph aims at helping science and technology teacher educators to:

- acquire the necessary knowledge and skills to evaluate trainees;
- acquire the skills of evaluating their own teaching;
- become aware of the problems associated with evaluation;
- become aware of the need for and methods of record keeping;
- develop the necessary understanding and skills of evaluation amongst trainee teachers.

4.3 The purpose of evaluation

Evaluation is carried out for the following purposes:

- to measure progress;
- for accreditation and certification;
- to determine whether objectives of a teaching programme have been achieved;
- to rank learners;

- to identify gaps in the knowledge and skills of learners;
- to provide information on progress to parents, teachers, teacher educators and decision makers.

The teacher educator should emphasise to the teacher trainees that the purpose of evaluation is not only to let the learner and parents know about the progress of the learner. It is also of vital importance to teachers because it informs them about the success of their own teaching. Information from evaluation can be used by teachers to plan their lessons. This should be demonstrated by the teacher educator.

4.4 Types of evaluation

The purpose of evaluation determines its type. Informal evaluation techniques may be used to collect information needed for decision-making. Evaluation may be carried out to collect a variety of information about teacher trainees, such as academic performance, performance in the classroom, ability to handle classroom material, presentation and communication skills and ability to plan and execute practical work. Similarly, teacher educators may evaluate trainees with regard to their knowledge and skills in planning and administering evaluation in school.

It is the use of the information collected which determines the nature of evaluation. If the information is used to diagnose the cause of a particular problem faced by a trainee, the evaluation is called **diagnostic**. If information gathered is used to monitor teachers' progress during pre-service or in-service programmes and to take corrective measures whenever required, evaluation is called **formative**. This type of evaluation is carried out periodically and is therefore also called **continuous** evaluation. Evaluation which is carried out at the end of a course or session is called **summative** evaluation.

4.4.1 Formative evaluation

Formative evaluation is carried out to provide the teacher, as well as the learner, with feedback that can be used in taking corrective measures at different stages in the teaching–learning process. Objectives of formative evaluation are:

- to determine the progress being made by the learners and the difficulties they may be experiencing;
- to guide the learning process;
- to decide which students need further help and the nature of activities and strategies that might help them;
- to pace the learning process and make the learner more alert to his or her performance.

Formative evaluation is conducted mainly to improve learning. This does not mean that tests used for formative evaluation cannot be graded. However, the grading is done to find strengths and weaknesses in the teaching–learning process. Formative evaluation, therefore, provides both teachers and learners with feedback on the achievement of the objectives. Teacher educators should model how this can be done in their own teaching.

4.4.2 Continuous evaluation

As the term implies, continuous evaluation is carried out periodically during the teaching-learning process. It has two main purposes:

- for formative evaluation;
- to supplement summative evaluation.

This means that the final grade a learner receives in a course is not based on a single final examination, but takes into account the results of the continuous evaluation.

Characteristics of continuous evaluation

- **Systematic.** It is systematic in that measurements to be made of the teacher's performance at different time intervals have to be planned.

- **Comprehensive.** It is comprehensive in that it makes use of many different techniques to collect information over the range of abilities. Information about cognitive, affective and psychomotor skills may be collected through tests, projects, assignments, observations, interviews and questionnaires. This holistic evaluation provides valuable information on the teachers.

- **Guidance-oriented.** Continuous evaluation is oriented towards providing guidance in that the information obtained is used to guide the teacher trainees' future professional development.

It is important that teacher educators demonstrate these characteristics of continuous evaluation in the evaluation of the teacher trainees and then enable the teacher trainees to use the strategies with their students during teaching practice. This will help the trainees to develop appropriate skills.

Problems associated with continuous evaluation

Teacher educators should be aware of the problems outlined below, discuss them with pre-service trainees, offer suggestions and allay fears.

- **Workload of the teacher.** There can be no denying that continuous evaluation makes demands on the teacher. Whether this work is above and beyond that which the teacher can reasonably be expected to undertake as part of teaching depends on particular situations. As was stressed above, evaluation and feedback are essential in teaching. Where continuous evaluation is used to provide feedback, it is reasonable to expect teachers to integrate this into their normal workload.

  The problem with continuous evaluation arises, however, when the outcomes need to be recorded for use by other teachers and outside bodies. This usually has to be done in a standard format that involves the teachers in additional clerical work and in carrying out evaluations at stipulated periods for record keeping rather than at times most appropriate to the teaching situation.

- **Teacher–student relationship.** Some teachers are concerned that formally evaluating students will interfere with teacher–student relationships and with the role of the teacher in providing feedback and counselling. The perception is that students will be afraid to give ideas and opinions because they may be penalised during evaluation.

  Teacher educators must realise that evaluation and feedback are one and the same and that evaluation, of any form, needs to be an integral part of teaching. When students understand this and are able to establish and accept the need for evaluation and feedback, the teacher–student relationship can be enhanced. It appears that the fear of teacher trainees is associated with an inability to link evaluation with teaching.

4.4.3 Summative evaluation

Evaluation which is carried out at the end of a course with the sole purpose of grading, promotion or selection is termed summative evaluation. Results of summative evaluation are also used for accountability purposes of individual teacher educators or teachers or of their institutions. Summative evaluation also informs policy makers about the standards of education and the efficacy of the curriculum.

Summative evaluation is normally carried out using paper and pencil tests and/or practical tests administered towards the end of a term or academic year. Such tests may be conducted by the institution itself or by an external examining agency.

In some countries, teacher trainees' success or failure at the end of a teacher education programme is mainly determined by an external examination body. It has been advocated that teacher educators should have some input in determining the final grades of their teacher trainees.

For a valid evaluation of the student teacher, the results of continuous evaluation should be used to supplement summative evaluation.

4.5 Methodology of evaluation

Harlen (1983) has suggested that evaluation begins by asking a few questions. She suggests the following sequence of questions:

- What is the purpose of evaluation?
- What information is required to serve the purpose?
- What methods are appropriate to gathering this information?
- What form of record is suitable for accumulating the information?
- What interpretation and use are to be made of the information?
- Does this use serve the initial purpose?

The first four questions must be asked before the evaluation takes place and the last two after the information has been collected through informal means such as observation or formal means such as tests. Thus the purpose of evaluation and the information required for making judgements determine the tools and techniques to be used to collect the information.

Teacher educators should ask themselves these questions when planning evaluation of trainees. After the evaluation exercise has been completed the teacher educators should discuss the basis of decisions on each question with the trainees. This will help the trainees to see evaluation in action and identify appropriate method(s) for collecting a particular kind of information, and the most suitable method of recording it. The trainees may then be given appropriate assignments to practise these skills. They may be asked to decide the kind of information to be collected, the tools and techniques to be used to collect the information and the form of record to be kept. A common mistake which occurs is to set a task or question designed to evaluate a particular objective or criterion which does not provide opportunity for the learner to demonstrate the achievement of that objective. This must be avoided.

In order to give practice in the interpretation and use of information collected, trainees may either be given some hypothetical information to interpret or asked to collect information themselves and interpret and use it. It is important for teacher educators and the trainees to know that the purpose of evaluation and the type of information collected could inform vital decisions. These decisions may be regarding the achievement of each learner or the class as a whole, difficulties being faced by individual learners and the class, the efficacy of the teaching–learning strategy utilised, and remedial activities that may be necessary.

Nature of information required for evaluation

The nature of information to be collected for different purposes depends upon the objectives being evaluated. In most countries the objectives relate to the cognitive, psychomotor (including processes) and affective domains, but the emphasis placed on each varies, as the following examples illustrate. In some countries there is heavy emphasis on content while in others the emphasis is on the development of psychomotor skills. The development of scientific attitude, interest and values is emphasised by all to varying degrees.

In **Namibia**, the syllabus gives the aims of the science course as:

- to acquire sufficient understanding and knowledge to become citizens of a confident world;
- to take or develop an informed interest in matters of scientific importance;
- to recognise the usefulness and limitations of scientific processes and appreciate their applicability in other disciplines and in everyday life;

In the **Kenya** Syllabus for Primary Schools Science, the objectives for standards 1–3 emphasise the development of attitudes, skills and acquiring information about the immediate surroundings through first-hand observation.

In **Nigeria** the objectives of science teaching are the acquisition of basic scientific skills: observing, manipulating, classifying, communicating, inferring, hypothesising, interpreting data and formulating models.

In **Uganda** the objectives of science teaching are to develop in children skills of observation, experimentation and evaluation.

Thus, the objectives of science and technology education about which information will have to be collected are in the areas of:

- knowledge and understanding of scientific facts, principles, methods and materials;
- process skills such as observation, classification, communicating, inferring, hypothesising, experimenting, designing and problem solving;
- attitudes, interests, curiosity, critical thinking, perseverance, openness.

4.6 Information-gathering techniques

A variety of techniques can be employed to collect information about learners' progress. These techniques may be classified as:

Formal

- written tests;
- practical tasks carried out by individuals or groups;
- questionnaires;
- scheduled interviews;
- rating scales;
- check lists.

Informal

- observation during normal activities;
- written records of work produced including drawings;
- listening to oral explanations.

4.6.1 Evaluation of problem solving

To solve a given problem, a learner has to use certain ideas and skills. From the point of view of evaluation, it is vital that these are carefully identified and documented, so that the extent to which they have been used successfully can be evaluated. For the purposes of grading and assessing progress, a series of statements in terms of what learners will or will not be able to do at each stage should is needed. Each stage can be assigned a mark. The panel shows how this approach is being followed by the INSTANT project in Namibia.

Possible investigative skills to be assessed and developmental stages

SKILL 1: Using and organising techniques, apparatus and materials

Stage 1: Can follow written, oral or diagrammatic instruction to carry out one operation; needs help to do more; uses familiar apparatus adequately, but needs showing how to use unfamiliar apparatus; rather thoughtless over safety points.

Stage 2: Can follow written, oral or diagrammatic instruction to carry out an operation involving a series of steps and uses familiar apparatus adequately and safety; needs demonstration of how to use unfamiliar apparatus.

Stage 3: Can follow written, oral or diagrammatic instruction to carry out an operation involving a series of steps and is able to modify the instructions to improve the operation of the equipment; uses familiar apparatus adequately and safely; can make a fair attempt at using unfamiliar apparatus.

SKILL 2: Observing, measuring and recording

Stage 1: Can follow detailed instructions to make observations; can make simple measurements using a simple measuring device; records results in an appropriate way when shown how to do so.

Stage 2: Makes measurements, given a brief outline of how to do it; can use some more complex measuring devices such as those having a scale where 1 division is equal to 0.1 of a unit or 2 units; reads most devices but may not do so with complete accuracy; records results in an appropriate way given an outline format.

Stage 3: Makes relevant observations that are as accurate as possible; can read any scale correctly; records results correctly without being given a format.

SKILL 3: Handling experimental observations and data

Stage 1: Can process results adequately given detailed instructions on how to do it; can draw one obvious conclusion from the results.

Stage 2: Can process results adequately given outline instructions on how to do it; can recognise results that might have experimental errors; can draw conclusions from the data.

Stage 3: Can process results adequately without help, recognises experimental errors and knows how to deal with them; can identify possible reasons for the error; can draw conclusions; and also makes general deductions from data.

Skill 4: Planning and investigating

Stage 1: Can suggest a simple experiment to investigate a practical problem, although this may not work; can attempt modifying the experiment if it does not work.

Stage 2: Can list a series of steps to carry out an investigation; can modify the steps that do not work well; can recognise the need to control variables though may not be very clear about how to do it systematically.

Stage 3: Can list a series of logical steps to carry out an investigation that is likely to work in practice; modifies the steps that do not work well; recognises the need to control variables in a systematic way.

4.6.2 Evaluation through observation during normal activities

Observation (as a technique to collect information) provides qualitative information about the objectives being evaluated. Through observation, skills, attitudes and acquisition of concepts can be evaluated over an extended period of time. Since the learners are being evaluated through observation of their normal activities, they are not under stress; also, there is no loss of

teaching–learning time. All the learners need not be evaluated at the same time through one activity. To aid observation, actual classroom activities could be video-recorded. These recordings could also be used by trainees to develop and practise observation skills.

Observation is a very powerful information-gathering technique since it can be used to collect information about a wide range of behaviours spread over several objectives. However, the teacher educators and the teachers need to recognise that observation goes beyond simply looking at what students are doing. For the technique to be used successfully, the activities to be observed and the particular stages within them, must be clearly identified. The emphasis is not on observing whether or not the learner is doing the right things but on what the learner can or cannot do.

In addition, criteria have to be established, against which observations can be measured and recorded. The criteria are generally derived from the learning objectives and the stage of development of the trainee. This aids proper interpretation of the observations and in establishing the learner's progress, information which is very useful for the teacher educators and teachers in planning future activities for the learners.

One disadvantage of collecting evaluation information through observation is the element of subjectivity. The observer has to be skilled in the technique: an unskilled observer may miss certain important factors that should be noted.

4.6.3 Using written tests for gathering information

Written tests may contain a variety of questions such as:

- short-answer;
- essay-type;
- multiple-choice;
- fill in the blanks.

Written tests are used mainly for collecting information about students' cognitive abilities and some drawing skills. They cannot be used for testing a wide range of psychomotor skills. They can, however, be used for both formative and summative evaluation. Most written tests include essay-type and multiple-choice questions, and some characteristics of these are listed below.

Essay-type questions

These questions require more extended written answers and the ability to integrate and express ideas involving:

- description, explanation and prediction of processes and structures;
- description of instruments, apparatus, etc.;

- factual knowledge;
- presentation of theoretical knowledge;
- interpretation of experimental and numerical data;
- discussion of results of experiments and solutions of problems.

Advantages

- They reduce guesswork in answers.
- They reduce the possibility of cheating.
- They provide freedom of response.

Disadvantages

- In marking essays, the teacher tends to carry impressions from one paper to another.
- Essays cannot be used effectively with learners with low levels of language development.

Evaluation through essays at primary school level should be limited to short paragraphs only.

Multiple-choice type questions

Multiple-choice questions can be useful in covering a range of curricular objectives. A multiple-choice question usually consists of a question or statement, called the 'stem', followed by four or five choices of answers or 'options'. The stem may also be a partial statement that is completed by one of the choices. Only one of the options is the correct answer to the question or ending to the statement. The other options are called 'distracters'.

Other forms of multiple-choice questions include:

- **True/false questions.** These are not particularly useful since there is a 50% chance of guessing the right answer.

- **Fill-in-the-blank questions.** Here possible responses are supplied to the learner, only one of which is correct.

- **Matching questions.** These are presented in the form of two lists of statements. For each statement in one list there is one correct statement in the other list and the learners have to match them. The primary cognitive skill tested by matching exercises is recall.

Teacher education programmes should include practice in the construction and review of different types of questions and written tests. The review should include analysis of each item of the test.

4.6.4 Evaluation of individuals in group work

Most evaluation in science and technology education is carried out by teacher educators, but student teachers can be good evaluators of each other especially when they work as a group on an activity. An example of how this works is drawn from South Africa.

An innovative teacher in a school, at St Stithians, South Africa, came up with an interesting approach to evaluating group work

An overall point (percentage) is awarded for the activity or assignment of the group as a whole. The members of the group then decide how the marks should be distributed amongst themselves, according to the contribution each made to the work. The total marks for individuals must, in the end, equal the mark the teacher gave to the group. If they feel they all contributed equally, they can each earn the same percentage. According to the teacher, this strategy works very well. Within a group, each individual knows very well who did the work or who was just sitting around!

4.7 Record keeping

There are many different ways to keep records of learners' achievements and progress. The nature of the record depends to a large extent upon the technique used to collect information. For example, the record of an observation tends to be descriptive, giving information on the absence or presence of the trait being observed. Similarly, the record of a written test tends to be numerical. The type of record to be maintained also depends upon the purpose or use of the record. In general, a numerical record is not useful if one wants to know what learners can or cannot do. But if the purpose is grading the learners then a numerical record is more useful than a descriptive record.

During formative evaluation, most of the information gathered is interpreted and used immediately for taking corrective measures in the learning activities of individuals and for planning future activities. A record of the stages of development reached by the learner can be very useful in deciding long-term activities.

A common practice is to prepare a list of activities and tick them off as individuals complete them. This kind of record is not very helpful since it does not tell the reader what the learner can or cannot do. To improve the usefulness of such a list, a column can be added where the skills observed/achieved could be written. Alternatively, a list of possible skills to be assessed could be drawn up, and for each learner the particular skills observed/achieved ticked off.

Activity no.	Skill 1	Skill 2	Skill 3	Skill 4	etc.
Name 1					
Name 2 etc.					

4.8 Profiles

A profile is a method of building up a picture of attainment that allows separate grades, scores, marks and other measures to be recorded for each student. The picture can be broadened by including factors such as attitudes, health data, punctuality, absenteeism and self-esteem. These additional attributes can be judged on, say, a 3- or 5-point scale.

Profiles can be of two types:

- An extension of the conventional evaluation data, in the sense that student performance is reported on a larger number of dimensions.

- Information presented as a record of the student's abilities, skills and attitudes, which are explicitly stated and which describe the student's characteristics and attributes on a criterion-referenced basis.

In the first type of profile, the data obtained from examinations is put forward as a kind of report. The data are presented in greater detail in the sense that greater numbers of subdivisions are included. But the data themselves are still either the percentage scores, or the percentages related to a 3- or 5-point scale, which may be numerical or labelled, for example, above average, average, below average. When the middle grade is labelled as average (average for the student population), the profile is said to be based on norm-referenced evaluation.

4.8.1 Developing criterion-related profiles

If adequate information is given about the learners, the profile system mentioned above can be meaningful. But in the absence of such data (which is often the case), more meaning is attached to the profile by including the specific criteria against which measurements are taken. There is

however a tendency for the profiles to be based on criteria and be related to criterion-referenced evaluation.

In a criterion-referenced profile, the numbers represent a 5-point range, with 5 being a high proficiency and 1 a low proficiency rating. The numbers could be easily substituted by grades or use could be made of other systems. For example the ability to use a scientific instrument could be graded as:

5 = can use a scientific instrument accurately and safely;

4 = can use a scientific instrument, with guidance, and is aware of safety aspects;

3 = can use a scientific instrument, with guidance, but is not aware of safety aspects;

2 = cannot use a scientific instrument, even after guidance, but is aware of need for safe practice;

1 = cannot use a scientific instrument, even after guidance, and does not consider safety aspects.

4.8.2 Extending the profile

Profiles can extend beyond characteristics that relate to cognitive ability or practical skills. The example below shows part of a profile covering aspects of attitude. Again, a criterion-referenced format is being followed.

Relationship with fellow students

4	A leader/dominant personality
3	Accepted member of student groups
2	Likes to join with other students, a follower
1	Independent, quite isolated; tends to be on his/her own

Ability to work with others

4	Works well as the leader of a team
3	Works well as a member of a team
2	Prefers to work on his/her own
1	Does not fit in well as a member of a team

Punctuality in the last year of school

3	Excellent
2	Some lateness
1	Poor

Discipline

4	Self-discipline – able to relate to a normal adult/child, teacher/student situation
3	Accepts a specified pattern of behaviour and rules
2	Accepts an imposed pattern of behaviour and rules where there is a degree of supervision
1	Does not always accept a pattern of behaviour required by the student group

4.8.3 Problems with profiling

Three are three main issues associated with the profiles as a method of record keeping:

- the time it takes to develop and maintain the profile;
- the type of profile to be designed and the qualities to be described in them;
- the technical aspects of the educational measurement on which the profile is to be based, including validity and reliability.

The type of profile developed depends on the needs of the end-user.

Profiles are usually developed to show what a student can achieve. It is a record of achievement and attitudes. The technical aspects of the profile are to do with the development of levels that show gradations in academic achievement or attitudinal development. The acceptability or validity of the profile depends on how meaningful the statements being made are to the end-user.