

1. DEFINITION OF EVALUATION

Evaluation is defined as a systematic process by which the worth or value of something is judged. Evaluation in Physical education can be defined as a continuous process based upon criteria which cooperatively developed and concerned with measurement of :

- The performance of learners.
- The effectiveness of teachers
- The quality of the programme

Evaluation could be of :

- Students
- Teachers
- Programmes and courses in relation to educational objectives.

2. PURPOSES OF EVALUATION

- (a) Incentive to Learn (motivation)
- (b) Incentive to perform (motivation)
- (c) Feedback to student/trainee
- (d) Modification of Learning/training Activities

- (e) Selection of students/trainee
- (f) Success or failure
- (g) Feedback to teacher/coach
- (h) University/College/Department/School public relations
- (i) Protection of society (certification of competence)

3. COMPONENTS OF EVALUATION

Five basic components of evaluation are as under :

1. Audience,
2. Purpose,
3. Questions
4. Scope, and
5. Resources

To determine these components, one must ask the following questions :

- For what audience is the evaluation being conducted?
- For what purpose is the evaluation being conducted?

- What questions will be asked in the evaluation?
- What is the scope of the evaluation?
- What resources are available to conduct the evaluation?

Measurement and Evaluation examples in Our Daily Routine are :

- alarm clock
- petrol pump gauge
- speedometer
- meeting new people
- opinion of the class
- current trends
- public health initiatives
- promotion of physical activity/physical education
- problems with obesity or NCDs (non communicable diseases)
- evaluation standards for school/college programmes

Evaluation

Definition 1 : The process of making judgments about the results of measurement in terms of the purpose of the measurement.

Definition 2 : The process of obtaining information (data) and using it to form judgments, which in turn are used in decision making.

Evaluation is use of measurement in making decisions which involves interpretation of a score. Hence, place value judgment on the measurement.

Terminology review can be made from the following such as Which is a measurement?, Which is a test? or Which is an evaluation?

- skinfold caliper
- 55 mm
- 'You are in the healthy fitness zone'

'Skinfold caliper' is a test; '55 mm' is a measurement whereas 'You're in the healthy fitness zone' is an evaluation.

Therefore, evaluation involves measurement (collecting data), judging the value of the data by comparing data to some standard and making decisions based on these data.

4. KEY FOCUS OF EVALUATION

Evaluations may be focused on progress in student learning (student evaluation), the effectiveness of Institution/school programmes (programme evaluation), and the effectiveness of the curriculum (curriculum evaluation). Teachers also reflect on the effectiveness of their instruction (teacher self-evaluation).

Evaluation of Students : Assessment techniques are selected or designed to gather data/information related to how well students are achieving the learning objectives of the curriculum. The specific assessment technique used at any given time depends on several factors, including the type of learning outcomes namely movement skill development, fitness development, conceptual understanding, personal and social skills, attitudes and values; the subject area content, the instructional strategies used, the student's level of development and the specific purposes of the evaluation. Various tests and measurement techniques are used to evaluate students / trainees/players on different parameters.

Evaluation of Programme : Evaluation of a programme is a systematic process of gathering and analyzing data/information about some aspect of a programme (may be college/school/training/coaching) in order to make a decision or to communicate to others involved in the decision-making process.

Evaluation of a programme can be conducted at two levels: relatively informally at the team level/classroom level or more formally at the laboratory/ground/classroom, institute/school or state/national levels.

At the team level/classroom level, evaluation of a programme is used to determine whether the programme being presented to the trainees/sportspersons/students is meeting their

needs and the objectives prescribed by the planner or administration. Evaluation of a programme is not necessarily conducted at the end of the programme but is an ongoing process. For example, if particular lessons/training session appear to be poorly received by students/trainees or if they do not seem to demonstrate the proposed learnings from a unit of training/study/coaching, the problem must be investigated and changes should be made.

By evaluating their programmes at the laboratory/ground/classroom, instructors/coaches/trainers/teachers become thoughtful practitioners. The data/information gathered through programme evaluation will assist instructors/coaches/trainers/teachers in programme planning and in making decisions for improvement.

Most evaluations of programmes at the team level/classroom level are relatively informal, but they should be done systematically. Such evaluations should include identification of the area of concern, collection and analysis of data/information and judgment or decision-making.

Evaluation of Curriculum : At the institution or state level, Evaluation of Curriculum involves making judgments about the effectiveness of authorized curricula in order to find out if they are being effectively implemented and whether they are meeting the needs of students.

Evaluation of Curriculum involves gathering data/information in order to determine how well the curriculum is performing. The principal reasons for curriculum evaluation are to make improvements such as changes to the curriculum document to provide resources or inservice to teachers.

It is projected that evaluation of curriculum be a shared, collaborative effort involving all of the main education partners in the institution/state/nation. Physical education teachers contribute in the development of instruments, validation, field testing, scoring and data interpretation.

In the assessment phase, data/information is gathered from students, teachers and administrators. The data/information obtained from educators indicates the degree to which the curriculum is being implemented, as well as the strengths and weaknesses of the curriculum. The information from students indicate how well they are achieving the proposed learning outcomes and may provide indications about their attitudes toward the curriculum.

All curricula must be included within the scope of curriculum evaluation. Evaluations should be conducted during the implementation phase for new curricula and regularly on a rotating basis thereafter.

Self-Evaluation of Teacher

An important aspect of good teaching practice is that of self-evaluation of teacher. Teachers can use various techniques to assess their strengths and weaknesses.

Terms Clarification: To enhance understanding of the evaluation process, it is useful to distinguish between the terms *assessment* and *evaluation*.

Assessment is a preliminary phase in the evaluation process. In this phase, various techniques are used to gather information about student/trainee progress.

Evaluation is the weighing of assessment information against some standard, such as the foundational objectives of the physical education curriculum, in order to make a judgment or decision. This may then lead to other decisions and action by the teacher/trainer, student/trainee or parent and administration or management.

5. TYPES OF EVALUATION

There are three main types of student evaluation: *formative*, *summative* and *diagnostic*. Assessment techniques are used to gather data/information for each type of evaluation.

1. **Formative evaluation** is a continuous classroom process that keeps students and

educators informed of students' progress toward programme learning objectives

The main purpose of formative evaluation is to improve instruction and student learning. It provides teachers with valuable information upon which instructional modifications may be made. This type of evaluation helps teachers understand the degree to which students are learning the course material and the extent to which their skills, fitness levels, knowledge, understanding and attitudes are developing. Students are given direction for future learning and are encouraged to take responsibility for their own progress.

2. Summative evaluation occurs most often at the end of a unit of study. Its primary purpose is to determine what has been learned over a period of time, to summarize student progress and to report to students, parents and educators on progress relative to curriculum objectives.

Seldom are evaluations strictly formative or strictly summative. For example, summative evaluation can be used formatively to assist teachers in making decisions about changes to instructional strategies or other aspects of students' learning programmes. Similarly, formative evaluation may be used to assist teachers in making summative judgments about student progress. It is important that teachers/trainers/coaches make clear to students/trainees the purpose of assessments and whether they may later be used summatively.

3. Diagnostic evaluation usually occurs at the beginning of the school year or before a unit of instruction. It identifies students who lack prerequisite skills, fitness, knowledge or understanding so that remedial help can be arranged. It also serves to identify gifted learners to ensure they are being sufficiently challenged. Diagnostic testing also identifies student interests. Diagnostic evaluation provides information essential to teachers in designing appropriate programmes for all students.

Teachers/trainers conduct all three types of evaluation during the course of the institution/school year or course period

6. GENERAL PRINCIPLES OF EVALUATION

the following five general guiding principles provide a framework to assist teachers in planning for student evaluation

(i) **Evaluation is a planned and continuous activity**: Evaluation is an essential part of the teaching-learning process. It should be a planned, continuous activity that is closely linked to both curriculum and instruction

(ii) **Evaluation is directed by proposed learning outcomes**: Evaluation should be guided by the intended learning outcomes of the curriculum and a variety of assessment strategies should be used

(iii) **Evaluation process is communicated in advance**: Evaluation plans should be communicated in advance. Students should have opportunities for input regarding the evaluation process. The weighting of criteria and in some instances the establishment of criteria can be negotiated among teacher and students.

(iv) **Evaluation must be fair and impartial**. It should be sensitive to family, classroom, school and community situations. It should be free of bias. Students should be given opportunities to demonstrate the extent of their skills, fitness levels, knowledge, understanding and attitudes.

(v) **Evaluation is for positive feedback**: Evaluation should help students. It should provide positive feedback and encourage students to participate actively in their own learning.

Phases of the Evaluation Process

Although evaluation is not strictly sequential, it can be viewed as a cyclical process made up of four phases namely preparation, assessment, evaluation and reflection. This process involves the teacher as a decision maker throughout all these four phases.

1. Preparation phase: During the preparation phase, decisions are made that identify what is to be evaluated, the type of evaluation to be used (i.e. formative, summative or diagnostic), the criteria against which student learning outcomes will be judged and the most appropriate assessment techniques with which to gather information on student progress. The teacher's decisions in this phase form the basis for the remaining phases

2. Assessment phase: During the assessment phase, the teacher identifies information-gathering strategies, constructs or selects instruments, administers them to the students, and collects and organizes the information on student learning progress. The teacher continues to make decisions in this phase. The identification and elimination of bias (e.g., gender and culture bias) from the assessment techniques and instruments, and determining where, when and how assessments will be conducted are examples of important considerations for the teacher.

3. Evaluation phase: During the evaluation phase, the teacher interprets the assessment information and makes judgments about student progress. Based on the judgments or evaluation, teachers make decisions about student learning programs and report on progress to students, parents and appropriate school personnel.

4. Reflection phase: The reflection phase allows the teacher to consider the extent to which the previous phases in the evaluation process have been successful. Specifically, the teacher evaluates the utility and appropriateness of the assessment techniques used. Such reflection assists the teacher in making decisions concerning improvements or modifications to subsequent teaching and evaluation.

7. DATA/INFORMATION GATHERING AND RECORD KEEPING

Reporting to students and parents must be based upon real evidence. In order to build up

a record of growth, teachers will have to rely to a great extent upon accurate observation and adequate record keeping.

Evaluation in physical education should focus on both process and content. Many of the objectives in the physical education curriculum are best assessed through ongoing observation

B. ASSESSMENT TECHNIQUES

Methods of Organization

These methods refer to the broad organizational structure of the assessment techniques a teacher/trainer may decide to use.

- Assessment Stations
- Individual Assessments
- Peer and Group Assessments
- Learning-Contracts
- Self-Assessment
- Portfolios

Data Recording Methods: These methods may be used within the structure of any of the methods of organization and with any of the ongoing student/trainee activities, as well as tests and quizzes.

- Anecdotal Records
- Observation Checklists
- Rating Scales

Ongoing Activities of Students: These techniques are those a teacher/trainer would use throughout the course of a regular institution/school working day when students/trainees are engaged in their usual learning activities.

- Performance Assessment
- Written Assignments
- Presentations
- Discussions
- Journal Writing
- Problem Solving
- Projects
- Homework
- Conferences/seminars
- Interviews

Assessment techniques

Assessment techniques

Standards for Evaluation

Norm-referenced Standard : Norms developed based on a large number of individuals. Norm-referenced Standards are used to judge performance in relation to the performances of others of a defined group. Percentile ranks are a common method for developing norms. Percentile rank is percentage that can be expected to score below a given value. If an individual's score reaches the 35th percentile, then that person scored better than 35% of the group.

Criterion-referenced Standard : This is used to determine if an individual has attained a specified level. A predetermined standard that shows whether an individual has achieved a desired level of performance. Hence, the performance of an individual is compared against a predetermined standard, not against the performance of others.

Steps Involved in Making an Evaluation

1. Define the objective or the purpose of the test.
2. Measure the performance or administer the test.
3. Find or develop a standard.
4. Compare a person's performance on the test to a standard.
5. Make the evaluation then discuss and distribute the results in the most appropriate manner.

Example : If a group of people take the same test on two different days, the scores obtained should be approximately the same.

A reliable test will yield data that are stable, repeatable, and precise. Reliability of a test refers to the dependability of test scores. The following factors contributing to Reliability :

- (i) Testing environment should be favourable
- (ii) Testing environment should be well-organized
- (iii) Administrator should be competent to run a favourable testing environment
- (iv) Range of talent
- (v) Motivation of the performers
- (vi) Good day vs bad day
- (vii) Learning, forgetting, and fatigue
- (viii) Length of the test
- (ix) Test difficulty
- (x) Ability of the test to discriminate
- (xi) Number of performers
- (xii) Nature of measurement device
- (xiii) Selection of scoring unit
- (xiv) Precision
- (xv) Errors in measurement
- (xvi) Number of trials
- (xvii) Recording errors
- (xviii) Classroom management
- (xix) Warm-up opportunity

Validity : The American Psychological Association (APA) reported that validity is the most important characteristic to a test or measuring instrument. The validity of each test can only be evaluated in terms of a particular purpose and for a particular group.

Example : A strength test that is valid for college-aged students is not necessarily valid for sedentary adults.

11. EVALUATION TECHNIQUES IN PHYSICAL EDUCATION

Some examples are as under :

Age, Height, Weight, and Body Types

- Most early measurements were in this area
- W. H. Sheldon developed somatotypes:
 - Ectomorphy (lean)
 - Mesomorphy (muscular hardness)
 - Endomorphic (heavy softness)

Power Measurement

- Vertical Jump Test (Dudley Sargent)
 - Correlates with 50 yd dash, high jump, standing broad jump, 8 lb shot put

Agility Measurement

- Shuttle run, Burpee Test

General Motor Ability Measurement

- Brace Motor Ability Test
- Iowa Brace Test

Balance Measurement

Kinesthetic Measure

- Difficult to measure; has a great deal to do with physical performance

Flexibility Measurement

- Sit and Reach Test is an example

Rhythm and Dance Measurement

- No test which is easy for teachers to use with classes

Speed and Reaction Time Measurement

- 40 yd dash; 50 yd dash; 100 meters; Stick test (reaction)

Strength Measurement

- Dynamometers; weights (one-rep max);

Muscular Endurance Measure

- Pull-ups, push-ups, chin-ups, repetitive weight lifting (curls, bench press, etc.)

Cardiovascular Endurance

- Harvard Step Test (field measure); Cooper Twelve-minute run-walk;

Physical Fitness Measurement

Military fitness tests; Kraus-Weber Test, AAHPERD Youth Fitness test etc.

Sports Skill Measurement

Posture Measurement

Social Qualities Measurement

Knowledge Measurement

Psychological Parameters Measurement

Physiological Parameters Measurement

Biomechanical Analysis Measurement

Purposes of knowledge tests

- (a) Assigning a grade or summative evaluation
- (b) Measuring progress or formative feedback
- (c) Providing feedback to students or programme participants as to their status and what the class or programme knowledge expectations are ?
- (d) Motivating students or programme participants to learn the material tested
- (e) Assessing teaching or instructional effectiveness

Construction and Administration of Psychomotor (Skill) Tests

12. TYPES OF PSYCHOMOTOR TESTS

Motor ability : The innate and acquired ability of an individual to perform motor skills of a general nature, exclusive of highly specialized sports or gymnastic skills.

Measurement of motor ability no longer popular because :

1. Existence of general motor ability questioned; abilities specific to performance task.
2. Construct validity of motor ability test batteries has never been established.
3. Lack of consensus on what the components of motor ability are.

Motor capacity : The individual's potential ability to perform motor skills.

Motor educability : The individual's ability to learn new motor skills. These tests are no longer popular.

Prefer to measure physical performance components of agility, balance, cardio