Original Article

Classroom Assessment in Higher Education

Higher Education for the Future
10(1) 11–30, 2023

DOI: 10.1177/23476311221143231
journals.sagepub.com/home/hef
SAGE

N. J. Rao[1] and Shilpi Banerjee[2]

Abstract

Classroom assessment is the process of documenting the knowledge, skills, attitudes and beliefs of learners. It provides essential feedback to both instructors and students to improve their teaching methods for guiding and motivating students to be actively involved in their learning. Assessment drives learning. Formative assessments enable the instructor to guide the students to learn well. Summative assessments enable the measurement of levels of attainment of course outcomes and act as feedback to course design and curriculum improvement. This article presents the underlying principles of assessment through a discussion of assessment approaches and their purposes, types of assessment items, quality of assessment and summative assessment plans. Quality assessment instruments can be developed through an understanding of the quality attributes of assessment items, the process of designing assessment instruments, designing a variety of assessment items, and devising plans to evaluate them through rubrics. An approach is presented for creating a summative assessment plan that can also lead to the attainment of outcomes as per the requirements of programme accreditation.

Keywords

Classroom assessment, assessment, assessment items, assessment plan, outcome based education, accreditation

Introduction

Graduate and post-graduate programmes are designed and conducted to facilitate students to acquire a set of programme outcomes identified by the accrediting agency of the country or the university and programme-specific outcomes

[1] Indian Institute of Science, Bangalore, Karnataka, India
[2] Azim Premji University, Sarjapura, Bangalore, Karnataka, India

Corresponding author:
N. J. Rao, Indian Institute of Science, 4th Main, 2nd Cross, NTI 2nd Stage, RMV 2nd Stage, Bangalore, Karnataka 560094, India.
E-mail: njraoiisc@gmail.com

identified by the Department offering the programme. Every course in these programmes is to be described, in the framework of outcome-based education, in terms of course outcomes that also foster the attainment of a subset of these programme outcomes and programme-specific outcomes. Formative assessments enable the instructor to guide the students to learn well. Summative assessments enable the measurement of levels of attainment of course outcomes and act as feedback to course design and curriculum improvement. For the assessments to be effective learning and teaching tools, instructors need to plan assessments (assessment instruments and test items) and conduct the assessment with appropriate tools and techniques. The assessment results can be used for improving teaching and learning.

This article presents the principles of good-quality classroom assessment and a method of designing such assessments. It presents the underlying principles of assessment through a discussion on assessment approaches and their purposes, types of assessment items, quality of assessment and summative assessment plans.

Educational Assessment

Assessment (Merriam-Webster Dictionary) is 'the act of judgment about something'. It also has an equally popular meaning as 'an amount that a person is officially required to pay'. While these two meanings are very distinct from each other, they come from the same source. 'Something' of assessment can relate to education, some aspects of persons, groups, organizations, societies, governments and the environment. Educational assessment, according to the National Institute on Learning Outcomes Assessment (NILOA) concerning the learners, is the systematic process of documenting and using empirical data on the knowledge, skills, attitudes and beliefs of students. Banta and Palomba (2014) define educational assessment as the process of providing credible evidence of resources, implementation actions and outcomes undertaken to improve the effectiveness of instruction, programmes and services in higher education. In education, the term assessment refers to the wide variety of methods or tools that educators use to evaluate, measure and document the academic readiness, learning progress, skill acquisition, or educational needs of students (Nami et al., 2014).

Student learning assessment also called 'classroom assessment', is the process of documenting the knowledge, skills, attitudes and beliefs of learners. Classroom instruction and student learning are influenced by the 'instructional situation' which consists, besides others, the Programme and the Institution. Assessment in the context of higher education, therefore, has come to encompass the entire process of programme and institutional effectiveness. In India, such assessments of some professional programmes are done by the National Board of Accreditation (NBA), and of an Institution is done by the National Assessment and Accreditation Council (NAAC) through their accreditation processes.

Classroom Assessment is a mechanism for providing instructors with essential feedback for improving their teaching methods for guiding and motivating students to be actively involved in their learning. The assessment gives us necessary information about what our students are learning and about the extent to which we are meeting our teaching goals (Primer, 2002). But the real power of assessment comes in using it to give feedback to students. Improving the quality of learning in a course involves not just determining to what extent students have mastered course content at the end of the course; also consists in determining to what extent students are mastering content throughout the course. Thus, in addition to providing the instructors with valuable information about our students' learning, assessment should assist the students in diagnosing their learning. That is, assessment should help students 'become more effective, self-assessing, and self-directed learners' (Angelo & Cross, 2012).

There is considerable evidence showing that assessment drives student learning (Wormald et al., 2009). More than anything else, our assessment tools tell students what we consider to be important. While assessment is one of the last things that university teachers consider when planning and developing courses, it is often the first thing that most students think about (Baud, 1995). Students will learn what we guide them to learn through our assessments (Primer, 2002). In other words, assessment is deciding what we want students to learn and making sure they learn it. Assessment, in many respects, is the glue that links the components of a course—its outcomes, instructional methods and skills development. Any changes in the structure and content of a course require coordinated changes in assessment.

We explore important dimensions of '**classroom assessment**' in the context of higher education programmes as offered in India. Figure 1 presents the important aspects of Classroom Assessment in Higher Education in the form of a concept map.

Types and Purposes of Assessment

The most important classification of assessment is Formative and Summative. This is to consider different objectives for assessment practices (Freeman & Lewis, 2016). Figure 2 presents the details of these two types of assessment, their purposes and their nature. In the context of classroom assessment, assessments are conducted to aid learning, improve teaching, certify students and diagnose prior knowledge.

Formative assessment is defined as all those activities undertaken by instructors or students, which provide information to be used as feedback to modify the teaching and learning activities in which they are engaged (Black & William, 2004). These assessments are integrated with the teaching-learning process and are conducted throughout the teaching. Formative assessment when conducted by instructors is referred to as the **assessment for learning**. Such assessment **improves teaching** by providing the instructors with valuable information on students' learning and helping instructors to review the

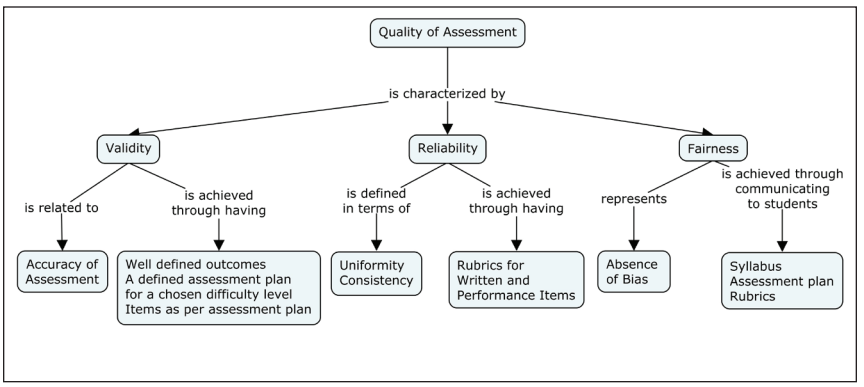

Figure 1 . Overview of Classroom Assessment in Higher Education Programmes.

Source: The authors.

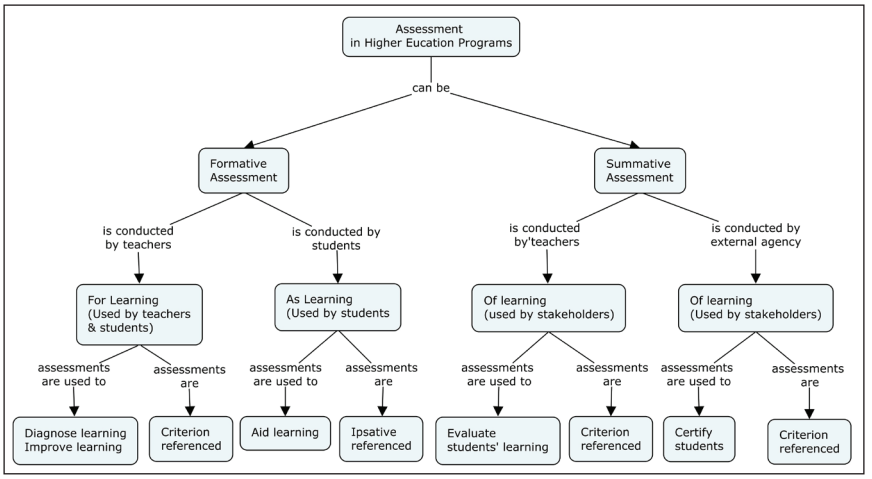

Figure 2. Types and Purposes of Assessment.

Source: The authors.

effectiveness of all instructional activities. Another form of formative assessment is **diagnostic assessment**. It provides instructors with information about students prior knowledge and misconceptions before beginning an instructional activity (Nitko, 1996). It is primarily used to diagnose student difficulties and to guide instruction and curriculum planning. These assessments are designed/referenced to a criterion. In the context of higher education, course outcomes are a **criterion**.

Students play an active role in formative assessment along with the teacher. The assessment which is conducted by the student himself/herself is called **assessment as learning** or self-assessment. It is a process in formative assessment during which

students reflect on the quality of their work (self-monitoring), judge the degree to which it reflects explicitly stated learning outcomes (self-judgement) and revise accordingly (self-revision) (Gipps, 1999). Metacognition is central to self-assessment; where the student consciously takes control of specific cognitive skills, such as checking, understanding, predicting outcomes, planning activities, managing time and switching to different learning activities. Students feel more motivated and confident to learn when they have taken charge of learning with the necessary scaffolding provided by the teacher. Therefore, such assessment **aids learning** by helping students identify their strengths and areas that need to be improved, track their progress and plan activities which will help them in attaining course outcomes. In this assessment, a student is trying to assess his/her progress against his/her own previous achievements. Such referencing is stated as **ipsative** which is self-comparison.

Summative assessment is intended to measure the attainment of course outcomes and report those outcomes to students, parents and administrators. In an educational setting, it occurs after a class, course, semester, or academic year. In the context of a course, summative assessments are typically used to **assign students a course grade** against a set of course outcomes (**criterion**). It is also referred to in a learning context as an **assessment of learning**. The summative assessment is also used for **certification** when it is used for indicating that a student has reached a particular standard. This may be in the form of simple pass or fail or 'competent' or 'not yet competent.' Assessment in these and similar circumstances certifies that a particular level of performance has been attained for a defined set of course outcomes (**criterion**). Summative assessment also helps in the **selection**, for example when choosing students for a next-level course or employment. Assessment in this context is used for prediction, for instance, which students will be able to benefit from further study or how the individuals might perform in employment. This is seen mostly in examinations such as Joint Entrance Examination (JEE), Common Entrance Test (CET), Graduate Aptitude Test in Engineering (GATE), Common Admission Test (CAT) and in tests associated with recruitment. Selection can help the learner choose his/her career. Selection has historically been linked to the ranking of students. Assessments in this form have been a means of positioning students in order of merit or achievement, therefore, they are called **norm-referenced**.

Nature of assessment for higher education courses: 20%–70% of the total marks for a course are allotted to Continuous Internal Evaluation (CIE) in India. Internal summative evaluation is carried out in the form of quizzes, assignments, presentations and tests. CIE results are used as feedback by both instructors and students to improve their teaching and learning respectively. The examination conducted at the end of a course is termed as Semester End Examination (SEE). 80%–60% of the total marks for a course are allotted to SEE. Some Private Universities in India use ratios of 60:40 or 70:30 for CIE: SEE. Students learn better if course outcomes, instruction and assessment are in alignment with one another. Both CIE and SEE assessments play important roles as CIE helps to find out whether instructional activities helped students to achieve those course outcomes and SEE aims to measure whether students have attained these course outcomes.

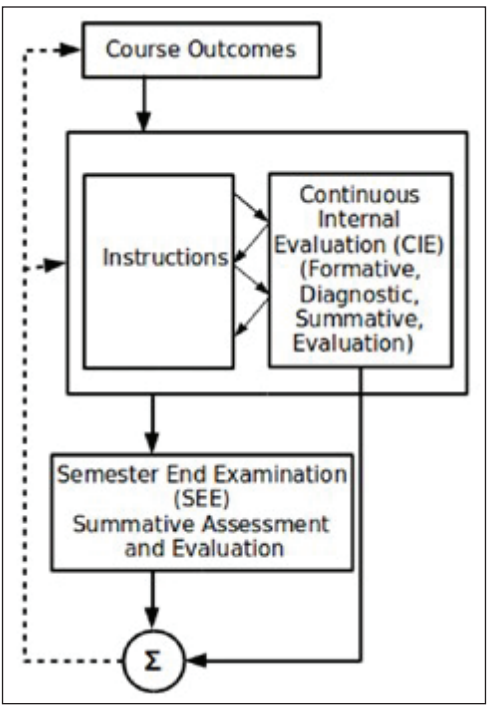

Figure 3. Assessments in Conducting a Course.

Source: The authors.

The role of assessments at various stages in teaching a course is presented in Figure 3.

It is possible to ensure good learning by using many types of assessments. The types of assessments used in programmes will depend on the policies of the Institution/University.

Test (Assessment) Items

Formative and Summative assessments are conducted using either written test items or performance test items. Types of test (assessment) items, their evaluation and classification are presented in Figure 4.

Written test items can be broadly classified as **Selection-type** test items and **Supply-type** test items. They are given in quizzes, class tests, assignments and final examinations, or in the form of reports students are required to generate. The test items can be selection-type items or supply-type items. For each specific course outcome, many test items can be prepared for assessment before instruction, during instruction and at the end of instruction.

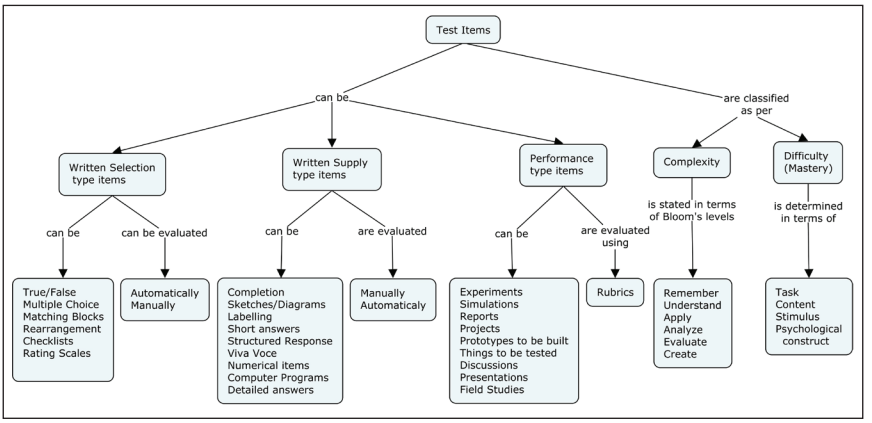

Figure 4. Classification of Test Items and Their Evaluation.

Source: The authors.

Selection type item is a form of questioning which has a single correct answer. Selection-type items include true/false solutions, multiple-choice, multiple-response and matching items. Such items are well suited to the increasingly popular computerized or online assessment format. These items are objective and have unique answers. Tests that include only selection-type items make it possible to compare students' performances directly. For example, it is possible to compare the performance of students within one section to students in a different section, or students under one instructor with those under another instructor. Typically, selection-type items are used to assess lower cognitive levels, as questions asked at lower cognitive levels usually have unique answers (definitions, lists, formulae, etc.). Evaluation of students' performances in selection-type test items can be automated. Examples of selection-type test items are True/False, Multiple choice (with single or multiple answers), Matching Blocks, Rearrangement, Checklists and Rating Scales.

Supply type item is a form of questioning which may have more than one right answer or more than one way of expressing the correct answer. Supply-type items include extended-response questions, essays, hypothesizing, creating plans, solving well-defined problems and ill-defined problems, etc. Well-defined problems have unique answers, there can be subjective variations in arriving at the final answer.

Supply-type test items are necessary when there is a need to test the ability of the student to use higher cognitive levels in the course. The supply-type items may be required where a selection-type cannot be devised. The supply-type items require the students to organize their knowledge. It also demands the ability to express ideas not required for selection-type items. These type of items are valuable in measuring the students' generalized understanding of a subject. On the other hand, a supply-type item may evaluate the student's abilities to innovate (cognitive level: Create) rather than their specific knowledge of the subject matter.

Examples of supply-type test items are Completion, Sketch, Labelling, Short Answers, Structured Response, Viva Voce, Numerical Questions, Detailed Answers and Programming assignments.

The main disadvantage of supply-type test items is that they cannot be evaluated with uniformity, which presents difficulties in automating the evaluation process. The same test item evaluated by different instructors (raters) would probably be given different scores. Even the same test item graded by the same instructor on consecutive days might be assigned different scores. Therefore, standardization of evaluation is not always possible. Still, another disadvantage of a supply-type test is the time required by the student to complete it and the time needed for the instructor to grade it. Such challenges can be overcome by using holistic and comprehensive rubrics which will be introduced in the subsequent section.

Examples of **performance test items** include Experiments, Simulations, Projects, Prototypes to be built, Testing of Things/objects/equipment/physical samples, Discussions, Presentations and Field studies. They will involve working with devices, performing, making presentations, discussions and report generation. These activity items can be very definitive or open-ended. The evaluation of the performance of students in activity-based test items is likely to be qualitative. For every type of performing test item, it becomes necessary to define rubrics.

Written supply-type items and performance-type items are evaluated using rubrics to ensure consistency in evaluation. A rubric is a scoring tool that lays out the specific expectation for an assignment. It divides an assignment into its component parts (dimensions) and provides a detailed description of what constitutes acceptable or unacceptable levels of performance for each of these parts. Rubrics can be used for grading a large variety of assignments and tasks: research papers, book critiques, discussion participation, online discussions, laboratory reports, conduct of experiments, portfolios, group work, projects, term papers, oral presentations, problem-solving and more (Stevens & Levi, 2005). It allows standardized evaluation according to specified criteria, making grading simpler and more transparent.

There are four parts to any rubric, presented in Table 1, which are Task Description, Dimension and Weightage, Scale and Scoring Criteria.

When developing a rubric, the instructor(s) can start with three **Scale Levels** and three **Dimensions.** As the rubric evolves, the number of Scale Levels may increase up to five, and the number of Dimensions may increase up to 10. However, increasing the Scale Levels and Dimensions make it harder to mark (score) the performance of the student.

The **task** description is generally framed by the instructor and involves a 'performance' by the student. The tasks related to the course are given in the syllabus of the course and communicated to the students at the beginning of the semester. The task should have a 'descriptive title' and a 'task description'.

The **dimensions** of a rubric represent the parts or aspects of the tasks completely. They also clarify for students how their tasks can be broken down into more important components. Dimensions should represent the type of component skills

Table 1. Elements and Structure of a Rubric.

Task: Task Description			
	Scale Level 1	Scale Level 2	Scale Level 3
Dimension 1	Scoring criteria		
Dimension 2			
Dimension 3			

Source: Stevens and Levi (2005).

students must combine in a successful assignment, such as depth of understanding of the content, content organization, assumptions made, use of appropriate language, analysis of the conclusions in an essay, gestures during a presentation, design in a product development project, and so on. Appropriate weight is given to each of these dimensions in the assignment. Dimension should not include any description of the quality of the performance, and it is indicated by the 'scale'. Some examples of Dimensions are Knowledge/Understanding, Communication, Use of visual aids, Presentation skills, Context, Evidence, Analysis, Group presentation skills, Group organization and Operating with declared timelines.

Sometimes the **weightage** given to all dimensions is not the same. In such cases, the weights need to be indicated clearly. For example, Dimension one (30%), Dimension two (30%), Dimension three (20%) and Dimension four (20%). Total marks obtained by the student should be computed taking these weights into consideration.

The **scale** describes how well or poorly any given dimension of the task is performed. One may use words such as 'Excellent', 'Competent', and 'Needs Work'. Along with these words marks or mark ranges are also used. Two of the commonly used labels and associated marks or mark ranges are

- Excellent (10 or 10–8); Competent (7 or 7–5); Needs work (4 or 4–0).
- Sophisticated (10 or 10–9); Competent (8 or 8–6); Partly competent (5 or 5–3); Not yet competent (2 or 2–0).

The labels, marks and mark ranges may be chosen by the instructor through peer discussion.

We require some **criteria to score** the performance of each student concerning each one of the identified dimensions. These will also be convenient when there are multiple evaluators of the same performance of the student. If these criteria are written with adequate detail, they will lead to providing timely and detailed feedback to the student and the teacher, encouraging critical thinking, facilitating communication with others, helping teachers to refine their instruction methods, creating a level playing field, saving considerable time to the teacher in marking the subjective assessment, and reducing the time to give feedback.

Classification of Assessment Items

All higher education programmes in India are expected to attain programme outcomes (POs) (identified by the NBA in professional programmes, and to be identified by the University/institution in case of general programmes) and programme-specific outcomes (PSOs) (identified by the Departments offering the programmes). Courses constitute nearly 90% of credits of all programmes. Hence, POs and PSOs can mainly be attained through courses. Courses are defined in terms of course outcomes (COs) which are statements that describe what students should be able to do at the end of a course. As per Outcome Based Education, students learn well when

- they are clear about what they should be able to do at the end of a course/ programme,
- the **assessment** is in alignment with what they are expected to do, and
- the instructional activities are designed and conducted to facilitate them to acquire and demonstrate what they are supposed to do.

Teaching and learning in education programmes involve cognitive, affective and psychomotor domains to varying degrees of emphasis. At present most of the courses are designed and conducted to address only the cognitive domain outcomes. The cognitive domain is characterized by two dimensions: cognitive processes (levels) and knowledge categories (Anderson & Krathwohl, 2001). The CO statements, which are predominantly cognitive, have four components:

- **Action** represents a cognitive/affective/psychomotor activity the learner should perform. An action is indicated by an action verb, occasionally two, representing the concerned cognitive process(s).
- one or more **Knowledge** elements (from the four categories in the case of Basic Sciences, Engineering Science and Humanities and Social Science courses, and eight categories in the case of Engineering courses).
- **Conditions(s)** under which the cognitive activity needs to be performed.
- **Criteria** that the cognitive activity needs to meet.
- **Conditions** and **Criteria** are optional elements of a CO statement.

Test items for formative and summative assessments need to be created as per the selected instruction methods chosen and to ensure the attainment of course outcomes. Two characteristics of test items are important in this context. These are **Complexity** and **Difficulty**.

Complexity of assessment items is associated with the cognitive level of the assessment items. For example, a test item at Evaluate level of Revised Bloom's taxonomy is more complex than a test item at Understand level. As per Revised Bloom's Taxonomy the cognitive processes and their sub-processes, arranged hierarchically, are

- Create (Generate, Plan (Design) and Produce)
- Evaluate (Critique and Check)
- Analyze (Attribute, Organize and Differentiate)
- Apply (Implement and Execute)

- Understand (Explain, Compare, Infer, Summarize, Classify, Exemplify and Interpret)
- Remember (Recall and Recognize)

Action verbs associated with these cognitive processes and subprocesses can facilitate the teacher to create test items at the right level of complexity.

Difficulty of assessment items is a measure of mastery learning of the course outcomes (Banerjee et al., 2015). Many variables are involved in determining the difficulty level of an assessment item. These variables, which are not completely independent of one another, are categorized into learner-dependent, learner-independent, subject-nonspecific, subject-specific and score based. However, the instructor can initially create and arrange the assessment items at three difficulty levels: Low (one), Moderate (two) and High (three).

Difficulty corresponding to an item resides in the associated task, content, stimulus and psychological construct (Osterlind & Friedman, 1999). **Task** difficulty refers to the difficulty that the students face when they generate their responses. This can be influenced by variables such as number of unknowns, number of conditions, numerical complexity and so on. **Content** difficulty is related to the elements of knowledge including facts, concepts and procedures. Some examples of content difficulty are the number of facts, number of concepts, number of procedures, number of prerequisite course outcomes from the same course, etc. **Stimulus** difficulty is related to the way the item is presented to the students which include words, phrases and information which is packed along with the item (Cheng, 2006). This includes item presentation, number of hints, independence of unknowns, usage of technical notations, number of inferences, number of resources and number of assumptions. variables. **Psychological** construct deals with learner-dependent variables, such as task motivation, anxiety, openness to experience, willingness to communicate (Gan, 2011), working memory, task switching, aptitude, self-efficacy, openness and implicitness. Thus, an assessment item is perceived differently by different students about the difficulty level.

Sample of Items of different difficulty levels

- **Low:** Determine the time period of a simple pendulum of length 1 m on the surface of the earth.
- **Moderate:** Determine the time period of a simple pendulum of length 1 m when it is placed in a lift which is moving upwards with an acceleration 2 ms^{-2}.
- **High:** Determine the time period of a simple pendulum of length 1 m with its bob dipped in a non-viscous medium of density one-tenth of bob and placed in a lift which is moving upwards with an acceleration 2 ms^{-2}.

It should be noted that test items at higher cognitive levels are not necessarily of higher difficulty levels. It is possible to design assessment items of lower difficulty levels at higher cognitive levels. Addressing the relevant cognitive levels is essential in designing assessment items. The difficulty levels of items can be reduced, but the complexity levels of items should never be reduced, which is a disservice to student learning.

Quality of Assessment

According to the Standards for Educational and Psychological Testing by the American Educational Research Association (AERA, 2018), **validity, reliability, and fairness** are critical attributes of the quality of any assessment. Figure 5 shows the meaning of these quality attributes and how they are achieved for summative assessment (Banerjee, 2021).

AERA defines **validity** as the degree to which evidence and theory support the interpretation of test scores for proposed uses of tests. Test scores from classroom assessments are used to make inferences about a student's mastery of the course outcomes and thereby measure the attainment of course outcomes. Therefore, to ensure validity, assessments should be aligned to both content domains and cognitive skills as indicated in course outcomes. Validity of an assessment instrument used for the summative purpose is ensured through the following three steps: (a) the **Course outcomes** (CO) for a course are defined. Assessment is then designed in alignment with the course outcomes for an identified cohort of students, (b) the **assessment plan** is designed to represent the weight distribution across course outcomes and cognitive levels. The assessment instruments are designed according to the defined plan, and (c) each defined course outcome is mapped to one of the cognitive levels of Revised Bloom's taxonomy. Most of the **assessment items** in the test should belong to the same cognitive level of the CO selected from the same to which the course outcome is mapped.

AERA defines **reliability** in terms of consistency with regards to the replication of the testing procedures. Reliability in the classroom context is understood in terms of the properties of consistency and uniformity. Classroom assessments comprise a series of informal and formal assessments including observation, projects, portfolios, debates and presentations. Consistency is ensured when there is uniformity in students' performance across multiple methods of assessment. Also, sharing clearly defined assessment expectations in the form of rubrics ensures consistency in students' performance and uniformity in teachers'

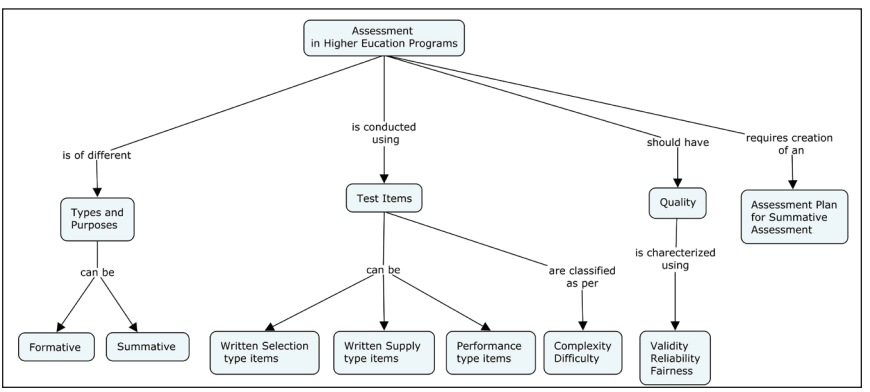

Figure 5. Quality of Assessment.

Source: The authors.

evaluation. Assessment should be designed in a way such that it should result in similar evaluation scores irrespective of what time it is evaluated or who has evaluated it. The items used for well-designed classroom assessment should have detailed **scoring rubrics** which are used for evaluating students' responses.

AERA defines **fairness** as responsiveness to individual characteristics and testing contexts so that testing scores yield valid interpretations for intended uses for different groups of test takers. For assessments to be fair, students should be aware of course outcomes, assessment plans and scoring rubrics.

Summative Assessment Plan

The way students approach assessment determines the way they think about learning. Therefore, the quality of summative assessment determines the quality of learning. Given the centrality of assessment in determining the quality of learning, it is necessary to understand what constitutes quality assessment. Design of quality assessment that is in alignment with Course Outcomes requires-

- Well-stated **Course Outcomes** in AKCC format (A—Action, K—Knowledge elements, C—Conditions and C—Criteria).
- An **assessment plan** compatible with stated Course outcomes.
- Design of **structures** of all CIE and SEE assessment instruments.
- Creation of some **sample test items** for each course outcome that are in complete alignment with it, and sample solutions for them, reflecting what the instructor considers as an appropriate answer and how the selected programme outcomes are addressed. Such sample solutions help students to follow good practices in writing their solutions to test items.
- Creation of an **item bank** with items at appropriate cognitive levels.

Course Outcomes: A course is described in terms of its 'Course outcomes' (COs). Every course outcome addresses a subset of Programme outcomes (POs) and Programme specific outcomes (PSOs). Programme outcomes are what every student graduating from the institution needs to attain. In the case of professional programmes, the POs are identified by the National Board of Accreditation, and in the case of general programmes, they are identified by the institution/university. The POs are discipline/programme non-specific. Programme Specific Outcomes are what the students graduating from a specific programme need to attain. These are identified by the Department offering programme through its Board of Studies. The number of class sessions required, which is indicative of scope, to address a CO is also marked. By tagging the COs of a course with POs, PSOs, Cognitive level, Knowledge categories and the number of classroom sessions we can communicate clearly to the students and other stakeholders about the intent of the course. It will also greatly help in planning and conducting instruction and designing good assessments. The first step in ensuring the quality of assessment is to write Course Outcomes and tag them appropriately. These COs need to be

Table 2. COs of the Course Analog Circuits and Systems.

CO	Course Outcome	POs/PSOs	CL	KC	Class (Hrs)	Lab (Hrs)
CO1	Understand analogue signal processing functions in present-day electronic products.	PO1, PSO1	U	F, C, C & S, PC	3	0
CO2	Design simple signal processing networks using linear and non-linear, passive and active, one-port and two-port electrical networks.	PO1, PSO1	Ap	C, P	5	0
CO3	Understand the behaviour and signal processing applications of passive and active electronic devices.	PO1, PSO1	U	C, P	5	4
CO4	Determine the performance of signal processing and conversion operations with devices that have parameters sensitive to temperature, voltage and time using negative and positive feedback.	PO1, PO4, PO5, PSO1	U	C, FDP	6	4
CO5	Design circuits that perform analogue linear signal processing functions, including amplification, summing, differentiation and integration and non-linear signal processing functions, including log and anti-log amplification, current sensing, rectification and dc voltage regulation using passive and active devices.	PO3, PO4, PO5, PSO1	Ap	F, C, P, C & S	9	8
CO6	Design passive and active biquad analogue lowpass, highpass, bandpass and bandstop filters in the base-band region using Butterworth, Chebyshev, Inverse Chebyshev and Elliptic approximations as per given specifications.	PO3, PO4, PO5, PSO1	Ap	C, P, C & S	8	8

(Table 2 continued)

(Table 2 continued)

CO	Course Outcome	POs/PSOs	CL	KC	Class (Hrs)	Lab (Hrs)
CO7	Design amplitude and frequency stable, and tunable sinusoidal and non-sinusoidal signal generators, crystal oscillators and modulated signal generators.	PO3, PO4, PO5, PSO1	Ap	C, P, C & S	3	2
CO8	Determine the performance of Frequency locked loops and Phase locked loops.	PO1, PO3, PO5, PSO1	Ap	C, P, C & S	4	4
CO9	Understand the history and trends in analogue electronic circuits and systems.	PO1, PSO1	U	F, C	2	0
Total hours of instruction					45	30

Source: The authors.

Notes: CL: Cognitive level; R: Remember; U: Understand; Ap: Apply; An: Analyze; E: Evaluate; C: Create; KC: Knowledge categories; F: Factual; C: Conceptual; P: Procedural; M: Metacognitive; FDP: Fundamental design principles; C&S: Criteria and specifications; PC: Practical constraints; DI: Design instrumentalities.

communicated to students as part of the Syllabus of the course. These tags of COs also greatly help in computing their attainment. A sample description of an engineering course with the COs tagged is shown in Table 2.

Assessment plan: The assessment plan refers to the weightage (marks) given to each CO in assessment instruments of CIE and SEE by the instructor and approved by the Department. We suggest that the weightage assigned to COs be proportional to the number of associated classroom/laboratory sessions. For example, consider the sample course Analog Circuits and Systems presented earlier. The weightage of COs can be worked out as follows:

CO1: (3/75) = 4%; CO2: (5/75) = 6.7%; CO3: (9/75) = 12% and so on.

The values may be rounded and adjusted as required to make the total 100. The allocation of marks for the course presented in Table 2 is shown in Table 3.

The number of assessment instruments under CIE is as per the policy of the Institute. We assume that both CIE and SEE address all COs according to the chosen weightages (classroom sessions). No attempt should be made to divide marks into fractions. Each teacher may work out these numbers differently, keeping the distribution of marks based on the time spent and the scope of and mastering the content. The choice of assessment instruments and weightage given to them can also differ from one course to the other based on the nature of the course and the policy of the Department/Institute.

Table 3. Allocation of Marks to COs of a Course for CIE and SEE.

CO	Class + Lab Hrs	Summative Assessment CIE Marks (rounded)	Summative Assessment SEE Marks (rounded)
CO1	3	4	4
CO2	5	7	7
CO3	5 + 4 = 9	12	12
CO4	6 + 4 = 10	13	13
CO5	9 + 8 = 17	23	23
CO6	8 + 8 = 16	21	21
CO7	3 + 2 = 5	7	7
CO8	4 + 4 = 8	11	11
CO9	2	2	2
Total	75	100	100

Source: The authors.

Table 4. Assessment Plan for CIE of the Sample Course.

		CIE			
CO	CIE Marks	A1	Test 1	Test 2	Lab
CO1	4	0	4	0	0
CO2	7	0	8	0	0
CO3	12	0	8	0	4
CO4	13	0	10	0	4
CO5	23	6	0	8	8
CO6	21	4	0	8	8
CO7	7	0	0	6	2
CO8	11	0	0	8	4
CO9	2	0	0	0	0
Total	100	10	30	30	30

Source: The authors.

One example of an assessment plan for CIE in the case of our sample course, which integrated the laboratory, is shown in Table 4.

An assessment plan for SEE can be done similarly. If all COs need to be addressed to their respective strengths (number of classroom sessions), then the marks distribution among all COs will be the same as that of CIE. As there is sequential dependence among the COs, the allocated marks for each CO can be altered by the instructor and have it cleared by the Departmental Committee. Test papers may have test items carrying different marks. It is possible to have, in SEE assessment instrument, test items carrying one mark to items carrying 20 marks or even more depending on the course. Further, there may be restrictions on the maximum number of sub-items that one main item can have. Many Institutions

Table 5. Assessment Plan of SEE for the Theory Part of the Course as Adjusted by the Teacher.

CO	SEE Marks	SEE Marks as Adjusted by the Teacher
CO1	4	2
CO2	7	6
CO3	12	12
CO4	13	14
CO5	23	24
CO6	21	22
CO7	7	8
CO8	11	11
CO9	2	1
Total	100	100

Source: The authors.

divide test papers into two parts, making one part compulsory and giving some choice in the second part. When there is a choice in any part, it must be an internal choice. In other words, the choice is always between items that belong to the same CO, the same cognitive level and the same difficulty level. SEE assessment instrument must be designed to account for all these issues. The assessment plan of SEE for the sample course is presented in Table 5.

Ensuring alignment in assessment design: Complete alignment of assessment with Course Outcomes requires that all assessments items related to a CO should be in the same cell as that of CO of the taxonomy table as shown in Table 6(a) or 6(b).

The courses in Sciences, Mathematics, Humanities, Social Sciences and Engineering Sciences can use Table 6(b) for planning assignments.

Many teachers consider it is not inappropriate to have some test items that belong to cognitive levels lower than that of the CO. A survey of many assessment instruments across several Institutions indicates instances of absence of test items that belong to the same cell as that of the CO.

Item banks: There is a vast variation among teachers in their opinions, ability and effort in designing assessments. Some are serious and put earnest effort into designing their assessment instruments. Some teachers, on the other extreme, may treat examinations as necessary evils. Consequently, they do not put adequate effort into designing assessments. Many times, teachers may have to operate under many informal and formal directives from the management. Sometimes, the effort needed to respond to the test items is wrongly estimated. The test items are sometimes recycled from the test papers of previous years. Many times, language issues lead to misunderstandings among the students. The sampling of the content may be non-uniform, leading to the contents of courses not being addressed. Many contingencies related to creating assessment instruments do not give adequate time to teachers to develop good and appropriate test items. Teachers new to teaching and teachers who are assigned new courses on short notice, require time to understand assessment processes related to creating assessment

Table 6(a). Alignment of Assessment Items (AI) with Course Outcome (CO) of An Engineering Course in Revised Bloom-Vincenti Taxonomy Table.

	Knowledge Categories							
Cognitive Processes	F	C	P	M	FDP	C & S	PC	DI
Remember								
Understand								
Apply			CO, AI					
Analyze								
Evaluate								
Create								

Table 6(b). Alignment of Assessment Items (AI) with Course Outcome (CO) of Humanities/Science/Commerce Course in Revised Bloom Taxonomy Table.

	Knowledge Categories			
Cognitive Processes	F	C	P	M
Remember				
Understand			CO, AI	
Apply				
Analyze				
Evaluate				
Create				

Source: Rao (2020).

instruments. It is difficult to assure the quality of assessment during this learning period. All these factors bring in a significant variation in the quality of assessment. One solution to ensure the quality of assessment is to create Item (Question) Banks (Banerjee et al., 2016).

The primary objections to the use of Item Banks are (a) The students are likely to have access to the Item Bank. The standard answers will get created and become available to students. In such a case, all items will then become Remember level items irrespective of the cognitive level implied by the related action verbs; (b) If the number of items in the Item Bank is small, students have less interest in attending classes and actively participate in learning. Even good students are likely to lose motivation to learn; (c) Even if one starts with large Item Banks, institutional dynamics are likely to lead to a reduction in its size; (d) The teachers are likely to instruct only to give answers to the items in the Item Bank.

Many of these objections can be overcome by making (a) the Item Bank large, with greater than 1,000 items, (b) the Item Bank dynamic through archiving 10% of items and adding 10% of items every year, (c) a significant percentage of items use numerics, making it possible to create infinite variations of the same item, thus preventing the conversion of an item into Remember type, (d) the Item Bank compatible with the quality of available students and (e) the processes related to assessment transparent to all the stakeholders to overcome, at least partly, the

institutional dynamics. Besides these, every Institution/Department can design and implement its checks and balances to ensure the quality and appropriateness of assessment.

Summary

Educational Assessment includes the collection and interpretation of data or evidence which provides information about students' learning, teaching, classroom practices, institutional processes and so on. Such information is important to improve student learning, instruction and institutional practices. There are several purposes of assessment and each of them plays an equally important role to support quality education. This article presents the purposes of assessment, an understanding of which will help teachers to design assessment items and instruments and use the results of assessment appropriately.

The quality of the assessment instrument determines the usefulness of evidence or inferences about levels of students' learning. Quality assessment instruments can be developed through an understanding of the quality attributes of assessment items, the process of designing assessment instruments (writing course outcomes, developing an assessment plan, creating assessment instrument structures and designing item banks), and designing a variety of assessment items and devising plans to evaluate them through rubrics.

Declaration of Conflicting Interests

The authors declared no potential conflicts of interest with respect to the research, authorship and/or publication of this article.

Funding

The authors received no financial support for the research, authorship and/or publication of this article.

ORCID iD

N. J. Rao https://orcid.org/0000-0002-1167-724X

References

American Educational Research Association (AERA). (2018). *Standards for educational and psychological testing*. American Educational Research Association.

Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.

Angelo, T. A., & Cross, K. P. (2012). *Classroom assessment techniques*. Jossey Bass Wiley.

Banerjee, S. (2021). Ensuring quality in classroom assessments. *Teacher*, *15*(2), 7–9.

Banerjee, S., Rao, N. J., & Ramanathan, C. (2015, October). *Rubrics for assessment item difficulty in engineering courses*. In 2015 IEEE Frontiers in Education Conference (FIE) (pp. 1–8). IEEE.

Banerjee, S., Rao, N. J., & Ramanathan, C. (2016, December). *Designing item banks in alignment with course outcomes for engineering courses*. In 2016 IEEE Eighth International Conference on Technology for Education (T4E) (pp. 152–155). IEEE.

Banta, T. W., & Palomba, C. A. (2014). *Assessment essentials: Planning, implementing, and improving assessment in higher education*. John Wiley.

Baud, D. (1995). *Enhanced learning through self-assessment*. Routledge Falmer, Taylor, and Francis.

Black, P., & William, D. (2004). The formative purpose: Assessment must first promote learning. *Yearbook of the National Society for the Study of Education*, *103*(2), 20–50.

Cheng, L. S. (2006). *On varying the difficulty of test items*. In Annual Conference of the International Association for Educational Assessment.

Freeman, R., & Lewis, R. (2016). *Planning and implementing assessment*. Routledge.

Gan. (2011). Second language task difficulty: Reflections on the current psycholinguistic models. *Theory and Practice in Language Studies*, *1*(8), 921–927.

Gipps, C. (1999). Chapter 10: Socio-cultural aspects of assessment. *Review of Research in Education*, *24*(1), 355–392.

Nami, Y., Marsooli, H., & Ashouri, M. (2014). Hidden curriculum effects on university students' achievement. *Procedia-Social and Behavioral Sciences*, *114*, 798–801.

Nitko, A. J. (1996). *Educational assessment of students*. Prentice-Hall.

Osterlind, S. J., & Friedman, S. J. (1999). Constructing test items: Multiple-choice, constructed response, performance, and other formats. *Journal of Educational Measurement, 36*(3), 267–270.

Primer. (2002). *An introduction to assessment: The basics*. http://archive.wceruw.org/cl1/flag/start/primer2.htm

Rao, N. J. (2020). Outcome-based education: An outline. *Higher Education for the Future*, *7*(1), 5–21.

Stevens, D. D., & Levi, A. J. (2005). *Introduction to rubrics: An assessment tool to save*. Stylus Publishing.

Wormald, B. W., Schoeman, S., Somasunderam, A., & Penn, M. (2009). Assessment drives learning: An unavoidable truth. *Anatomical Sciences Education*, *2*(5), 199–204.

Authors' Bio-Sketch

N. J. Rao was the Chairman of CEDT (Centre for Electronics Design and Technology, IISc between 1981 and 1996, and Chairman, Department of Management Studies between 1998 and 2006 and superannuated as a Professor at CEDT in July 2006. He was a member of several committees associated with NBA, and a member of the Core Committee that defined the new Accreditation processes of NAAC. His research areas included Control Systems and System Dynamics. His present research interests include higher education, pedagogy and education technologies.

Shilpi Banerjee is working at Azim Premji University since 2017. She has a PhD in Educational Assessment from International Institute of Information Technology, Bangalore. She brings the experience in teaching in various engineering colleges across the country and applied research experience in the area of educational assessment. Her current research interests include the development of feasible quality assessment prototypes for formative purposes, assessment design and statistical evaluation of large-scale assessment data. She is also involved in designing and offering courses in various aspects of assessment to teacher educators, education functionaries, practitioners and MA Education students. She is part of various technical committees set up by state and national boards to strengthen the design of board examinations and classroom assessments.