

ASSESSMENT IN HIGHER EDUCATION

JUNAID SARFRAZ KHAN

*Department of Examinations, University of Health Sciences, Lahore – Pakistan***ABSTRACT**

Assessment drives learning and influences the quality of learning by directing the approach students take towards learning and by aligning the outcomes of the educational program with teaching methodologies and the educational environment. Assessment needs to be recognized as a multidimensional entity and not a singular activity or concept that transcends across three domains cognitive, affective and psychomotor. Any assessment requires activation of and access to different cognitive, affective and psychomotor skills at multiple levels and their applications through a fusion in a multidimensional collusion of stored memories, learned knowledge and behaviour and acquired skills. Another dimension that requires consideration here is the context in which assessment takes place. Context of assessment can be defined in terms of the environment in which assessment takes place, its background, stakes as well as the stakeholders involved. New formats and mediums are being used in all areas of education both as a learning / teaching strategy as well as for assessment. Computerized, computer – aided or online teaching and learning have paved the way for computer – assisted – assessment techniques. Whether assessment is formative or summative, influences its design, approach and outcomes. To the administrator, the results of the assessment, either formative or summative, provide data that will help establish current policies or bring changes to them. To the program developers, the same results establish the worth of the program or otherwise. To the trainees, the scores or feedback help in understanding their deficiencies in relation to the clearly predefined goals and objectives of the educational program. The public places great emphasis on the nature of assessment and the outcomes related to it since it is the public that is going to use the product of the medical education programs and confidence in the product will be related to their acceptability of the assessment and its outcomes. This paper identifies different formats of assessment and their contextual relevance.

Keywords: Assessment, Formative assessment, summative assessment, Bloom's Taxonomy, Context.

BACKGROUND

Assessment drives learning and influences the quality of learning by directing the approach students take towards learning and by aligning the outcomes of the educational program with teaching methodologies and the educational environment. In any educational program students learn that, which shall be assessed rather than what is required. Assessment, therefore, requires

a strategic planning whereby teaching and learning is driven through it to achieve the desired goals in a competency – oriented, outcomes – based educational program.

Assessment needs to be recognized as a multidimensional entity and not a singular activity or concept. Table 1 presents the taxonomy of educational learning and assessment divided into three domains cognitive,

Table 1: Bloom's taxonomy of educational objectives.

<i>Cognitive Domain</i>			
<i>Category</i>	<i>Behavior Description</i>	<i>Examples</i>	<i>Keywords</i>
<i>Knowledge</i>	Recall data or information.	Multiple-choice test, recount facts or statistics, recall a process, etc.	Arrange, define, describe, label, list, recognize, relate, reproduce, select, state.
<i>Comprehension</i>	Ability to grasp the meaning of material,	Explain or interpret a given scenario or statement, suggest treatment,	Explain, reiterate, classify, give examples, illustrate, translate, review, report, discuss.
<i>Application</i>	Ability to use learned material in new and concrete situations,	Put a theory into practical effect, demonstrate, solve a problem.	Use, apply, discover, manage, execute, solve, produce, implement, construct, change, Prepare.

<i>Category</i>	<i>Behavior Description</i>	<i>Examples</i>	<i>Keywords</i>
<i>Analysis</i>	Interpret organizational principles, structure, construction.	Identify constituent parts and functions of process or de-construct a methodology.	Analyze, break down, catalogue, compare, quantify, test, examine, experiment, relate, graph, diagram, plot.
<i>Synthesis</i>	Ability to put parts together to form a new whole.	Develop plans or procedures, integrate methods, resources, ideas.	Develop, plan, build, create, design, revise, formulate, propose, establish, assemble.
<i>Evaluation</i>	Ability to judge the value of material for a given purpose.	Select the most effective solution. Hire the most qualified candidate.	Review, justify, assess, present a case for, defend, report on, investigate, direct, appraise, argue.
<i>Affective Domain</i>			
<i>Receiving</i>	Awareness, willingness to hear, selected attention.	Listen to teacher, take interest in learning, participate passively.	Asks, chooses, describes, follows, gives, holds, identifies, locates, points to, selects, replies, uses.
<i>Responding</i>	React and participate actively.	Participates in class discussions. Questions new ideals, concepts, models, etc.	Answers, assists, aids, complies, discusses, greets, helps, performs, presents, reads, recites.
<i>Valuing</i>	Attach values and express personal opinions.	Decide worth and relevance of ideas, experiences.	Argue, challenge, debate, refute, confront, justify, persuade.
<i>Organization</i>	Reconcile internal conflicts; develop value system.	Qualify and quantify personal views, state personal position.	Build, develop, formulate, defend, modify, relate, prioritize, reconcile, contrast, arrange.
<i>Internalize or characterize values</i>	Adopt belief system and philosophy.	Shows self-reliance when working independently.	Act, display, influence, solve, practice, proposes, qualifies, questions.
<i>Psychomotor Domain</i>			
<i>Perception</i>	The ability to use sensory cues to guide motor activity.	Detects non-verbal communication cues.	Recognize, distinguish, notice, touch, hear, feel, etc.
<i>Set</i>	Readiness to act.	Mental, physical or emotional preparation before experience.	Arrange, prepare, get set, states, volunteers.
<i>Guided response</i>	The early stages in learning a complex skill that includes imitation and trial and error.	Imitate or follow instruction, trial and error.	Imitate, copy, follow, try.
<i>Mechanism</i>	Basic proficiency.	Competently respond to stimulus for action.	Make, perform, shape, complete.
<i>Complex Overt Response</i>	Skillful/expert proficiency.	Execute a complex process with expertise.	Coordinate, fix, demonstrate.
<i>Adaptation</i>	Skills are well developed and the individual can modify movement patterns to fit special requirements.	Alter response to reliably meet varying challenges.	Adapts, alters, changes, rearranges, reorganizes, revises, varies.
<i>Origination</i>	Creating new movement patterns to fit a particular situation or specific problem.	Develop and execute new integrated responses and activities.	Design, formulate, modify, re-design, trouble – shoot.

Table 2: Requirements and concepts behind assessment.²

<i>Learning Domain</i>	<i>Activities</i>	<i>Delivery Considerations</i>	<i>Assessment</i>
<i>Cognitive</i>	Self-check quizzes Case studies Drill and practice Short answer essay Project or problem – based activities.	Web-enhanced materials supplementing classroom lectures; Hybrid course with cognitive content on the web; Multimedia simulations of challenging and key concepts.	Project based for higher cognitive skills Multiple choice or short essay questions Case Studies.
<i>Affective</i>	Goal setting Self – reflective writing in a journal Practice tutorials designed for student success.	Face-to-face meetings Motivational videos Streaming audio explanations and encouragement Interactive video, web casts, conference calls.	Self-assessment using check-list Pre / post attitude survey related to course content Retention/success in course.
<i>Psychomotor</i>	Practice of desired skill with feedback Arranging sequences of an activity in correct order.	Face-to-face demonstrations Demonstration videos Pictures with audio and text explanations Interactive video demonstrations.	Performance of skill matches set standard as observed by an instructor or designee.

affective and psychomotor presented by Bloom.¹ behind assessment of the three domains and their sub-levels.

The pitfall to avoid here is not to consider each layer within each domain separately but in a multi-layered, multidimensional manner. Any assessment requires activation of and access to different cognitive, affective and psychomotor skills at multiple levels and their applications through a fusion in a multidimensional collusion of stored memories, learned knowledge and behaviour and acquired skills. Repeated activation and application of learned knowledge, behaviour and skills reinforces the same and improves it through the value of the experiences gained through its application.

Another dimension that requires consideration here is the context in which assessment takes place. Context moulds assessment, learning through assessment and the outcomes of the assessment. Context of assessment can be defined in terms of the environment in which assessment takes place, its background, stakes as well as the stakeholders involved.

New formats and mediums are being used in all areas of education both as a learning / teaching strategy as well as for assessment. Computerized, computer – aided or online teaching and learning have paved the way for computer – assisted – assessment techniques. These have evolved from the very basic, resembling pen – and – paper tests, to use of increasingly greater adaptive technology and newer formats requiring multimedia and constructed responses to finalize the programs of both learning and assessment, embedding virtual reality and simulations. This brings all three i.e. the learning experiences, the assessment context and the learning through assessment as close to reality as possible.

Table 2² presents the requirements and concepts

Whether assessment is formative or summative influences its design, approach and outcomes. Formative assessment is defined as “a range of formal and informal assessment procedures employed by teachers during the learning process in order to modify teaching and learning activities to improve student attainment”.³ Whereas summative assessment (or summative evaluation) refers to the assessment of the learning and summarizes the development of learners at a particular time.⁴ Since the purpose of formative assessment is to provide feedback on assessment, stakeholders approach formative assessment differently than summative assessment in which the stakes are higher. What needs to be recognized is the power and potential of formative assessment in aligning educational strategies to achieve the outcomes of the program and to make summative assessment a success.

If formative assessment is to be rooted within the educational environment, depicting the outcomes of individual components of the program as close to reality as possible, it will influence tremendously in driving learning in the right direction especially at the right time. Formative assessment is not formative on account of it being assessment but by virtue of the feedback that is generated out of it and presented to the students, rather all stakeholders. The analogy to consider here is that of multiple test – drives, pit – checks and fine tuning by the entire team of a Formula 1 course before the final outcomes, the final drive which is competitive and is to involve the driver and the car above without the rest of the team and whose stakes are so high that failure could represent considerable losses for the entire team (the stakeholders).

Assessment can therefore be classified on the basis

of its functionality. Airasian and Madaus⁵ classified assessment into the following categories:

1. *Placement assessment:*

Examples of placement assessment are entrance tests like Medical Colleges Admission Test (MC-AT), assessment of students at the beginning of the course to place them into groups based on their background knowledge, skills and behaviour level or during the course of the program to again assign them into groups that may require different facilitatory or instructional approaches. To arrive at these decisions, a host of different tests and inquiries can be used including simulated real-time performance measures, pen – and – paper tests, self – reports, peer – reports, direct observation of behaviour, records of past achievements and experiences and outcomes of counseling sessions.

2. *Formative assessment:*

This assessment typically does not count towards assigning course grades. Its purpose is to provide feedback to the stakeholders on the alignment of the strategic goals of the program and the progress towards those goals by the stakeholders. Therefore, it requires 360° feedback to be fully effective. As discussed previously, for formative assessment to be effective, it needs to be set in as much of a real – time, objective and competency – oriented setting, as possible, assessing the program strategic goals as realistically as can be. Only then can it guide the instruction and learning in the direction where it shall culminate in achievement of the program goals, fully assessed through summative assessment at the end of the program.

3. *Diagnostic assessment:*

Whereas formative assessment aims to provide feedback and correction of deficiencies in instruction and learning in an educational program, the purpose of diagnostic assessment is to provide the stakeholders with a diagnosis of the problems or obstacles hindering the progress of instructions or learning in the right direction and at the right time so that adequate remedial actions can be taken by the stakeholders concerned to achieve the strategic goals of the program. Diagnostic assessment is, therefore, a specialized assessment requiring specific tools like psychoanalysis, direct observations etc.

4. *Summative assessment:*

Summative assessment is the final evidence of achievement of cognitive and psychomotor gains and changes in behaviour that were intended in the educational program. It is used for assigning course grades and for certifying competency in the outcome – oriented competency – driven higher

education program. Summative assessment is important in providing the feedback to all stakeholders that the outcomes have been achieved. In medical education, summative assessment certifies that the product of the medical education program is safe to progress to the next stage of competency development and, to finally become an independently functioning health professional. This certification is important for the public trust in the program and its products. Summative assessment therefore, must be an assessment of the competencies of the product as close to the real environment as possible for that assessment to be sufficiently valid and reliable to foster feelings of trust in the product.

Assessment can also be classified by virtue of interpretation of assessment procedures. That is, assessment can be norm – referenced or criterion-referenced.

5. *Norm – referenced assessment:*

This can be defined as a measure of performance: cognitive, psychomotor or behavioural skills separately or in a combination, interpreted in terms of an individual's standing in a known group relative to others within the group.

6. *Criterion – referenced assessment:*

This can be defined as the measure of performance cognitive, psychomotor or behavioural skills against predefined criteria, reference or measure. As an example, if the objective of an educational program was to train a typist to type 40 words per minute, a certain referenced test shall measure the competence of the student against the yardstick or objective of 40 words typed per minute. Therefore, criterion – referenced assessment is also called objective referenced assessment.

Standard – based assessments in medical education fall within this category as well. They may typically involve the use of checklists where performance of the candidates are measured against set criteria; pass and fail are not dependent on the relative standing of an individual student within the cohort but by achieving minimum safe standards. Most of the tests in medical education at present, however, are a mix of the two varieties, that is, they measure the student competence against fixed predefined criteria and objectives but also report on the relative standing of individuals within the cohort.

Our final distinction between the two, categories is that whereas criterion – referenced tests are typically designed to measure the degree of competency or mastery achieved against predefined objectives, norm – referenced tests tell us of the

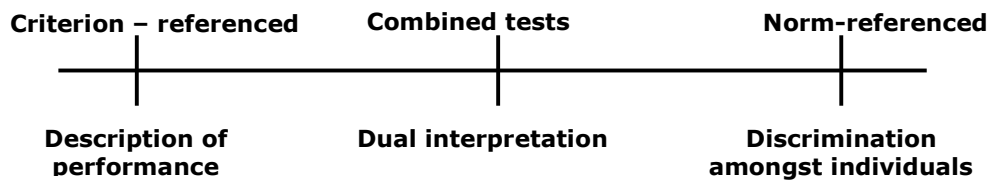


Table 3: Comparison of NRT and CRT.

Common characteristics	<i>NRT and CRT</i>	
	Both require specification of the achievement domain to be measured.	
	Both require a relevant and representative sample of test items.	
	Both use the same types of test items.	
	Both use the same rules for item writing (except for item difficulty).	
	Both are judged by the same qualities of goodness (validity and reliability).	
Both are useful in educational assessment.		
Differences	<i>NRT</i>	<i>CRT</i>
	Typically covers a large domain of learning skills, with just a few items measuring each specific task.	Typically focuses on a delimited domain of learning tasks, with a relatively large number of items measuring each specific task.
	Emphasizes discrimination among individuals in terms of relative level of learning.	Emphasizes description of what learning tasks individuals can and cannot perform.
	Favors items of average difficulty and typically omits very easy and very hard items.	Matches item difficulty to learning tasks, without altering item difficulty or omitting easy or hard items.
	Interpretation requires a clearly defined group.	Interpretation requires a clearly defined and delimited achievement domain.

relative standing of each individual within the group. Of note here is the arbitrary distinction between the two based on the relative standing, either within the group, or against a criterion. As already stated, it is perhaps more common these days, to focus on both, with each test providing a description of competency achieved and to the level that it has been achieved within the group and thereby, the information how the group as a whole has achieved those objectives.

This represents a continuum as shown below⁶:

Comparison of norm – referenced tests (NRTs) and criterion – referenced tests (CRTs) is provided in table 3.

Cronbach⁷ further classified assessment into two broad categories:

a) *Measures of maximum performance:*

These measures or tests assess the performance of individuals when they are maximally motivated to perform and achieve the highest.

b) *Measures of typical performance:*

These tests are designed to determine the normal

or typical in routine performance of the individuals.

In medical education, examples of the two performances can be derived from practice. Typical routine performance of practitioners is seen in the day to day, run – of – the – mill activities of health professionals, in activities that they consider routine, like working in the Out Patient Department (OPD) diagnosing a set of routine diseases. Maximum performance is observed when individuals are challenged by encounters that are other than normal or routine, when they have to perform at the best of their abilities to arrive at the desired outcomes. This may be a rare or challenging diagnosis, a particularly complicated surgical procedure, etc.

Of importance in this distribution is the position a test has in the continuum from the routine to maximum. This will largely depend on the context in which the test is applied, the objectives of the test and the outcomes that are being measured. Secondly the objectives of the program shall also determine how to shift the routine to the maximum in the day to day activities of the practitioner

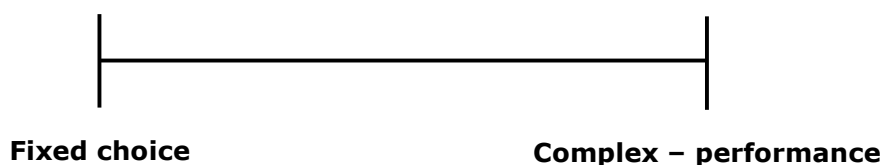


Table 4: Summary of various categorizations of assessments in Higher Education.

<i>Basis for Classification</i>	<i>Type of Assessment</i>	<i>Function of the Assessment</i>	<i>Illustrative Instruments</i>
<i>Nature of assessment</i>	Maximum performance	Determines what individuals can do when performing at their best	Aptitude tests, achievement tests
	Typical performance	Determines what individuals will do under natural conditions	Attitude, interest, and personality inventories; observational techniques; peer appraisal
	Fixed-choice test.	Efficient measurement of knowledge and skills, indirect indicator.	Standardized multiple-choice test.
<i>Form of assessment.</i>	Complex – performance assessment placement.	Measurement of performance in contexts and on problems valued in their own right Determines prerequisite skills, degree of mastery of course goals, and / or best mode of learning.	Hands – on laboratory experiment, projects, essays, oral presentations Readiness tests, aptitude tests, pretests on course objectives, self – report inventories, observational techniques.
<i>Use in classroom instruction.</i>	Formative.	Determines learning progress, provides feedback to reinforce learning, and corrects learning errors.	Teacher – made tests, custom – made tests from textbook publishers, observational techniques.
	Diagnostic.	Determines causes (intellectual, physical, emotional, environmental) of persistent learning difficulties.	Published diagnostic tests, teacher-made diagnostic tests, observational techniques.
	Summative.	Determines end-of-course achievement for assigning grades or certifying mastery of objectives.	Teacher-made survey tests, performance rating scales, product scales.
<i>Method of interpreting results</i>	Criterion referenced.	Describes student performance according to a specified domain of clearly defined learning tasks (e.g., adds single – digit whole numbers).	Teacher – make tests, custom – made tests from test publishers, observational techniques.
	Norm referenced.	Describes student performance according to relative position in some known group (e.g., ranks 10 th in a classroom group of 30).	Standardized aptitude and achievement test, teacher – made survey tests, interest inventories, adjustment inventories.

in our case. This shift is paramount towards the road to competence.

Another distinction that is applied to the methods of assessment is based on the continuum of fixed choice tests and complex – performance assessment.

At the far end of the continuum are the various

formats of the Multiple Choice Questions also known as the objective selected – response – test items including the Extended Matching and the True / False varieties. These tests are highly efficient because students can respond to a large number of questions relatively quickly, thereby covering a large area of the curriculum over a short period of time with high validity, reli-

ability, efficiency and feasibility. Since objectivity and comprehensiveness are more important to the test results than the use of machines. Both certainly improve efficiency.

Major problems associated with the fixed choice tests are firstly, the emphasis on low – levels of knowledge at the expense of problem – solving and conceptual skills. Secondly, according to Resnick and Resnick⁸, such tests drive instruction towards accumulation of facts and procedures rather than construction of knowledge through understanding and analysis.

The last few decades have seen a paradigm shift in higher education in general towards standard – setting, quality – control and quality assurance, outcome – based and competency – oriented assessment. This paradigm shift has been reflected in assessment through construction of multidimensional, multilevel and complex performance assessment techniques including written essays, Objective Structured Clinical Examinations (OSCE), creative exercises that require analysis, comprehension and conjugation of various cognitive, psychomotor, and affective elements.

Falling between these extremes are tests that require short answers, like the short – essay – questions or the structured – answer – questions. Interestingly none of the examples provided here against the categories of the continuum can be depicted as stereotypes. A long essay question on account of the way it has been constructed can very well fall short of assessing higher order cognitive process and a short essay question, when constructed with care, can extract through application and creativity in its design the same in its response. The same can be said of the fixed – choice selected – response test items which when provided with a multidimensional problem solving scenario may require higher order thinking to elicit a response.

Complex – performance assessments can be built into authentic assessments in vitro like OSCE, or the more authentic real – time in vivo work – place as-

ssments.

One of the drawbacks of performance – based complex assessment models is the subjectivity they bring into the assessment process. Assessment of performance at levels of competence requires scoring by competent and qualified assessors. Training these assessors in applying objectivity and a criterion-referencing system in assessment can obviate a number of these concerns.

Table 4⁶ provides a summary of various categorization of assessments used in higher education with examples of test instruments applied. Of particular note is the multiple faceted nature of assessment and the multiple uses of the instruments depending on how they are constructed and the context that they are used in.

REFERENCES

1. Bloom BS. *Taxonomy of Educational Objectives, Handbook 1; The cognitive domain*. New York: McKay; 1956.
2. Vinson C. *Learning Domains and Delivery of Instruction* [internet]. 2011[cited 2012 Feb 13]. Available from: <http://pixel.fhda.edu/id/learning_domain.html>
3. Ministry of Education, New Zealand. *Assessment: policy to practice*. Wellington, New Zealand: Learning Media; 1994.
4. Wikipedia – the free encyclopedia, 2011. *Summative assessment* [internet]. 2011 [cited 2012 Feb 13]. Available from: <http://en.wikipedia.org/wiki/Summative_assessment>.
5. Airasian PW, Madaus GJ. Functional types of student evaluation. *Meas. and Eval Guid.* 1972; 4: 221-233.
6. Linn RL, Miller MD. *Measurement and Assessment in Teaching*, 9th edn. New Jersey: Davis KM; 2005.
7. Cronbach LJ. *Essentials of psychological testing*, 5th edn. New York: Harper and Row; 1990.
8. Resnick LB, Resnick DP. *Assessing the thinking curriculum: New tools for Educational reform*. Gifford, BR and O'Connor, MC (eds.). *Changing assessment: Alternative views of aptitude, achievement and instruction*: 37-75, Boston: Kluwer Academic Publishers; 1992.