PRINCIPLES OF HIGH QUALITY ASSESSMENT

Allan M. Canonigo

http://love4mathed.com
1. Clarity of learning targets
2. (knowledge, reasoning, skills, products, affects)
3. Appropriateness of Assessment Methods
4. Validity
5. Reliability
6. Fairness
7. Positive Consequences
8. Practicality and Efficiency
9. Ethics
Assessment can be made precise, accurate and dependable only if what are to be achieved are clearly stated and feasible. The learning targets, involving knowledge, reasoning, skills, products and effects, need to be stated in behavioral terms which denote something which can be observed through the behavior of the students.
Cognitive Targets

Benjamin Bloom (1954) proposed a hierarchy of educational objectives at the cognitive level. These are:

- **Knowledge** – acquisition of facts, concepts and theories
- **Comprehension** - understanding, involves cognition or awareness of the interrelationships
- **Application** – transfer of knowledge from one field of study to another or from one concept to another concept in the same discipline
- **Analysis** – breaking down of a concept or idea into its components and explaining the concept as a composition of these concepts
- **Synthesis** – opposite of analysis, entails putting together the components in order to summarize the concept
- **Evaluation and Reasoning** – valuing and judgment or putting the “worth” of a concept or principle.
Skills, Competencies and Abilities Targets

- Skills – specific activities or tasks that a student can proficiently do
- Competencies – cluster of skills
- Abilities – made up of relate competencies categorized as:
  i. Cognitive
  ii. Affective
  iii. Psychomotor

Products, Outputs and Project Targets
- tangible and concrete evidence of a student’s ability
- need to clearly specify the level of workmanship of projects
  i. expert
  ii. skilled
  iii. novice
2. APPROPRIATENESS OF ASSESSMENT METHODS

a. Written-Response Instruments
   - **Objective tests** – appropriate for assessing the various levels of hierarchy of educational objectives
   - **Essays** – can test the students’ grasp of the higher level cognitive skills
   - **Checklists** – list of several characteristics or activities presented to the subjects of a study, where they will analyze and place a mark opposite to the characteristics.
2. APPROPRIATENESS OF ASSESSMENT METHODS

b. Product Rating Scales
- Used to rate products like book reports, maps, charts, diagrams, notebooks, creative endeavors
- Need to be developed to assess various products over the years

c. Performance Tests - Performance checklist
- Consists of a list of behaviors that make up a certain type of performance
- Used to determine whether or not an individual behaves in a certain way when asked to complete a particular task
d. Oral Questioning – appropriate assessment method when the objectives are to:
- Assess the students’ stock knowledge and/or
- Determine the students’ ability to communicate ideas in coherent verbal sentences.

e. Observation and Self Reports
- Useful supplementary methods when used in conjunction with oral questioning and performance tests
3. PROPERTIES OF ASSESSMENT METHODS

- Validity
- Reliability
- Fairness
- Positive Consequences
- Practicality and Efficiency
- Ethics
Something valid is something fair.
A valid test is one that measures what it is supposed to measure.

**Types of Validity**
- **Face**: What do students think of the test?
- **Construct**: Am I testing in the way I taught?
- **Content**: Am I testing what I taught?
- **Criterion-related**: How does this compare with the existing valid test?

Tests can be made more valid by making them more subjective (open items).
Validity – appropriateness, correctness, meaningfulness and usefulness of the specific conclusions that a teacher reaches regarding the teaching-learning situation.

- **Content validity** – content and format of the instrument
  1. Students’ adequate experience
  2. Coverage of sufficient material
  3. Reflect the degree of emphasis

- **Face validity** – outward appearance of the test, the lowest form of test validity

- **Criterion-related validity** – the test is judged against a specific criterion

- **Construct validity** – the test is loaded on a “construct” or factor
Something reliable is something that works well and that you can trust.
A reliable test is a consistent measure of what it is supposed to measure.

Questions:
- Can we trust the results of the test?
- Would we get the same results if the tests were taken again and scored by a different person?

Tests can be made more reliable by making them more objective (controlled items).
Reliability is the extent to which an experiment, test, or any measuring procedure yields the same result on repeated trials.
Equivalency reliability is the extent to which two items measure identical concepts at an identical level of difficulty. Equivalency reliability is determined by relating two sets of test scores to one another to highlight the degree of relationship or association.
Stability reliability (sometimes called test, re-test reliability) is the agreement of measuring instruments over time. To determine stability, a measure or test is repeated on the same subjects at a future date.
Internal consistency is the extent to which tests or procedures assess the same characteristic, skill or quality. It is a measure of the precision between the observers or of the measuring instruments used in a study.
Interrater reliability is the extent to which two or more individuals (coders or raters) agree. Interrater reliability addresses the consistency of the implementation of a rating system.
RELIABILITY – CONSISTENCY, DEPENDABILITY, STABILITY WHICH CAN BE ESTIMATED BY

- Split-half method
- Calculated using the
  i. Spearman-Brown prophecy formula
  ii. Kuder-Richardson – KR 20 and KR21
- Consistency of test results when the same test is administered at two different time periods
  i. Test-retest method
  ii. Correlating the two test results
The concept that assessment should be 'fair' covers a number of aspects.

- Student Knowledge and learning targets of assessment
- Opportunity to learn
- Prerequisite knowledge and skills
- Avoiding teacher stereotype
- Avoiding bias in assessment tasks and procedures
Learning assessments provide students with effective feedback and potentially improve their motivation and/or self-esteem. Moreover, assessments of learning gives students the tools to assess themselves and understand how to improve.

- Positive consequence on students, teachers, parents, and other stakeholders
Something practical is something effective in real situations.
A practical test is one which can be practically administered.

Questions:
- Will the test take longer to design than apply?
- Will the test be easy to mark?

Tests can be made more practical by making it more objective (more controlled items)
- Teacher Familiarity with the Method
- Time required
- Complexity of Administration
- Ease of scoring
- Ease of Interpretation
- Cost

Teachers should be familiar with the test,
- does not require too much time
- implementable
The problem:

- The more reliable a test is, the less valid.
- The more valid a test is, the less reliable.
- The more practical a test is, (generally) the less valid.

The solution:

As in everything, we need a balance (in both exams and exam items)
8. ETHICS

- Informed consent
- Anonymity and Confidentiality
  1. Gathering data
  2. Recording Data
  3. Reporting Data
ETHICS IN ASSESSMENT – “RIGHT AND WRONG”

- Conforming to the standards of conduct of a given profession or group
- Ethical issues that may be raised
  i. Possible harm to the participants.
  ii. Confidentiality.
  iii. Presence of concealment or deception.
  iv. Temptation to assist students.