

PRINCIPLES OF ASSESSMENT - PART 1*

Kelvin Seifert
Nathan Gonyea

This work is produced by OpenStax-CNIX and licensed under the
Creative Commons Attribution License 3.0[†]

Abstract

Principles and basic concepts in assessment

NOTE: Dr. Rosemary Sutton contributed to this module.

1 Basic Assessment Concepts

Best practices in assessing student learning have undergone dramatic changes in the last 30 years. In the past teachers often did not assess students' learning, they tested them on the knowledge and skills taught during the previous weeks. The tests varied little in format and students always did them individually with pencil and paper. Many teachers now use a wide variety of methods to determine what their students have learned and also use this assessment information to modify their instruction. In this module the focus is on the basic principles of assessments.

Assessment is an integrated process of gaining information about students' learning and making value judgments about their progress (Linn & Miller, 2005). Information about students' progress can be obtained from a variety of sources including projects, portfolios, performances, observations, and tests. The information about students' learning is often assigned specific numbers or grades and this involves measurement. **Measurement** answers the question, "How much?" and is used most commonly when the teacher scores a test or product and assigns numbers (e.g. 28 /30 on the biology test; 90/100 on the science project). **Evaluation** is the process of making judgments about the assessment information (Airasian, 2005). These judgments may be about individual students (e.g. should Jacob's course grade take into account his significant improvement over the grading period?), the assessment method used (e.g. is the multiple choice test a useful way to obtain information about problem solving), or one's own teaching (e.g. most of the students this year did much better on the essay assignment than last year so my new teaching methods seem effective).

The primary focus in this module is on **assessment for learning**, where the priority is designing and using assessment strategies to enhance student learning and development. Assessment for learning is often **formative assessment**, i.e. it takes place during the course of instruction by providing information that teachers can use to revise their teaching and students can use to improve their learning (Black, Harrison, Lee, Marshall & Wiliam, 2004). Formative assessment includes both **informal assessment** involving spontaneous unsystematic observations of students' behaviors (e.g. during a question and answer session or while the students are working on an assignment) and **formal assessment** involving pre-planned,

*Version 1.4: Nov 6, 2012 1:47 pm +0000

[†]<http://creativecommons.org/licenses/by/3.0/>

systematic gathering of data. **Assessment of learning** is formal assessment that involves assessing students in order to certify their competence and fulfill accountability mandates. Assessment of learning is typically **summative**, that is, administered after the instruction is completed (e.g. a final examination in an educational psychology course). Summative assessments provide information about how well students mastered the material, whether students are ready for the next unit, and what grades should be given (Airasian, 2005).

2 Types of Assessment

There are many different ways to categorize assessments, some of which we addressed earlier in the module. These include the relationship between the test and instruction (formative vs. summative), the rigor of the assessment procedures (formal vs. informal), the type of student performance desired (maximum vs. typical), the type of test items used (supply vs. selection), whether time is a factor in evaluating performance (speed vs. power), how the test is administered (individual vs. group), how the test is scored (objective vs. subjective), and how grades are assigned (norm referenced vs. criterion referenced). We previously addressed the first two qualities: the relationship between the test and instruction and the rigor of assessment procedures. Assessments can be given during instruction in order to provide feedback to the student and teacher about learning, **formative assessment**, or at the end of instruction as a final evaluation of what was learned, **summative assessment**. The development, administration and scoring of an assessment can be highly structured and systematic, **formal assessment**, or have little structure and be unsystematic, **informal assessment** (Linn & Miller 2005).

Another method of classifying assessments is based on the type of student performance a teacher is trying to evaluate. We can assess our students' **typical performance**, their ability on an average day with no preparation like in a pop quiz. Or, we can assess their **maximum performance**, their peak performance given ample time to study and prepare for the assessment as in a final exam (Linn & Miller 2005).

In addition to classifying assessments based on the type of performance a teacher wants to assess, we can classify assessments based on the way students show their knowledge of information. Specifically, do we want students to recognize information or recall information? An assessment containing **selection** questions asks student to recognize the correct answer that is provided in a list of options. These types of questions include multiple choice, true/false, and matching. **Supply** questions are questions that require students to recall the answer from memory without alternatives being provided – such as in short answer or essay items (Linn & Miller 2005).

The next two methods of classifying assessments relate to the administration of the test. The first is whether or not the amount of time it takes to complete the assessment is important. In a **speed test** the amount of time it takes a person to complete the assessment is important. These tests usually contain a large number of questions and a short time limit. These assessments are often used to measure fluency, or someone's ability to complete simple tasks automatically and quickly (e.g. simple addition or letter identification). In a **power test** the amount of time it takes to complete the assessment is not important. However, this doesn't mean that there isn't a time limit, only that the time limit doesn't impact most students' performance on the exam. For instance, Mr. McMorris, a 5th grade teacher, may give his students 1 hour to complete a 30 question multiple choice end of chapter exam in science. While this test has a time limit, virtually every 5th grade student can complete 30 multiple choice questions in under one hour without rushing (Linn & Miller 2005).

The other method of classifying assessments related to administration is whether the test can be given to more than one person at a time. In an **individual assessment** the teacher must give each student the exam one-on-one. For instance, high school foreign language teachers often assess students' oral language fluency through one-on-one oral exams. In a **group assessment** the teacher can give the exam to multiple students at the same time. The vast majority of tests you have taken in school are group assessments because the entire class takes the test at the same time (Linn & Miller 2005).

The last two methods of classifying assessments deal with the grading. The first deals with the amount of interpretation that is required in scoring. An **objectively scored** assessment contains questions that

require no interpretation on the part of the grader. For instance, on a multiple choice test the option selected is either correct or incorrect. In **subjectively scored** assessments the grader must interpret the response given by a test taker to determine its level of correctness. For instance, in an essay exam there are numerous ways to phrase the same correct answer. The grader must read the test taker's answer and determine to what extent the answer meets the requirements of the question (Linn & Miller 2005).

Once the individual items on an assessment have been scored, the students must be assigned grades. Grades can be assigned using one of two methods – norm referencing or criterion referencing. In a **norm referenced** test, a student's grade is based on how his or her performance compares to other students who took the same exam. For instance, Ms. Tang, a high school math teacher with a class of 10 students, may decide to assign the student who does the best on the algebra exam an A, the next 2 highest scoring students Bs, the middle 4 scoring students Cs, the next 2 highest scores Ds, and the lowest performing students an F. You will notice that there is no mention of the number of questions each student answered correctly, only the relationship between their overall performance and that of the other students in the class. All of the students could have answered 90% or more of the questions correctly or none of the students could have answered more the 50% of the questions correctly (Linn & Miller 2005).

In a **criterion referenced** test, a student's grade is based on how his or her performance compares to an absolute standard. For instance, a B being awarded to students who answer 80%-89% of the questions correctly on an exam. In Ms. Tang's example, she could have implemented a criterion referenced grading system by awarding any student who answered 90%-100% of the questions correctly an A, 80%-89% correct a B, 70%-79% correct a C, 60-69% correct a D, and less than 60% correct an F. You should notice that there is no mention of the number of students who will earn each grade. All of the students could earn As or they could be evenly distributed across the letter grades (Linn & Miller 2005).

3 References

Airasian, P. W. (2005). *Classroom Assessment: Concepts and Applications* (3rd ed). Boston, MA: McGraw Hill.

Black, P., Harrison, C., Lee, C., Marshall, B. & Wiliam, D. (2004). Working inside the black box.: Assessment for learning in the classroom. *Phi Delta Kappan*, 86 (1) 9-21.

Linn, R. L., & Miller, M. D. (2005). *Measurement and Assessment in Teaching* (9th ed). Upper Saddle River, NJ: Pearson.