ASSESSMENT TERMS AND DEFINITIONS

Direct vs Indirect Measures

Direct measures include student products or performances that demonstrate that specific learning has taken place. Indirect measures, on the other hand, may imply that learning has taken place (e.g., student perceptions of learning) but do not specifically demonstrate that learning or skill. Surveys are frequently used as indirect measures of student learning or student understanding of career options or student perceptions of preparation for the workplace, etc. Middle States recommends using both types of measures; direct measures are preferred but can be combined with indirect measures.

Direct Measures include (but are not limited to)

- comprehensive exams
  -- even MC exams can be useful here when individual items are written to measure multiple outcomes (e.g., knowledge, application, analysis, synthesis, etc.)
- research papers, projects, presentations, performances
  -- use of rubrics/rating scales across multiple raters is important here; overall assignment/paper/project/presentation grades are not specific enough
- portfolios
- field experience supervisor evaluations
- standardized tests (such as licensure exams or graduate study preparation exams)

Indirect Measures include (but are not limited to)

- surveys and interviews
- course evaluations
- retention, graduation, & placement rates
- admission rates into grad schools

Formative vs Summative Assessment

Formative assessment is done on an ongoing basis to determine if changes should be made at the course or program level in an almost immediate manner. For example, course assessment procedures may indicate that a particular teaching strategy is not having the intended effect (i.e., students do not appear to be learning what we want them to learn or they do not appear to be invested in the class or material). With this information in hand, a course instructor may decide to alter the format of the course or make other adjustments to course requirements in an attempt to improve learning. At the program level, a formative assessment procedure may be used midway through in order to determine if students are acquiring the prerequisite skills they need to eventually meet end-of-program goals. If students are not progressing as expected, changes may be instituted at the program level. Summative assessment, on the other hand, occurs at the end of a course or program. Summative assessment allows us to determine if our overall student learning outcomes have been achieved (at the course or program level).

Objective vs Subjective Measures

“An objective assessment is one that needs no professional judgment to score correctly (although interpretation of the scores requires professional judgment). Subjective assessments, on the other hand, yield many possible answers of varying quality and require professional
judgment to score” (Suskie, 2004, p. 99). Thus, multiple choice tests are common objective measures; essay exams and other written assignments are common subjective measures.

Quantitative vs Qualitative Measures
Quantitative evidence of student learning is represented numerically (e.g., a test score or % of students passing a comprehensive exam on the first attempt). Quantitative scores make comparisons and general statements about performance easy. Qualitative evidence of student learning, on the other hand, includes narratives or other non-numerical information (e.g., student responses to open-ended survey items or information gathered via focus groups). Qualitative measures are more difficult to summarize and make comparisons difficult but can provide a wealth of useful information.

Rubrics & Rating Scales
“A rubric is an instrument based on a set of criteria for evaluating student work. Rubrics help a professor or other evaluator to make explicit, objective, and consistent the criteria for performance that otherwise would be implicit, subjective, and inconsistent if a single letter grade were used as an indicator of performance. Rubrics delineate what knowledge, content, skills, and behaviors are indicative of various levels of learning or mastery. Ideally, grading rubrics are shared with students before an exam, presentation, writing project, or other assessment activity. Conscious awareness of what he or she is expected to learn helps the student organize his or her work, encourages self-reflection about what is being learned and how it is being learned, and allows opportunities for self-assessment during the learning process” (Middle States Manual, p. 42). Suskie (2004) describes three basic types of rubrics:

a. rating scales
b. descriptive rubrics
c. holistic scoring guides

Validity & Reliability of Measures
Validity addresses the question of whether or not an instrument or test measures what it is intended to measure (or is something else inadvertently being measured). Reliability addresses the question of how consistently the instrument or test measures what you intend to measure.