Janet Fulks, Faculty, Bakersfield College – ASCCC, North Area Representative

Sue Granger-Dickson, Counselor, Bakersfield College

January 16, 2009
Authentic Assessment

What is Authentic Assessment?

- Reflects Explicit Criteria
- Exhibits Reliability
- Represents Valid Content
- Assesses Higher Level Learning
- Simulates real world experiences
- Includes Multiple Domains

What is the difference between assessment and grading?

Paul Dressel (1976) has defined a grade as “an inadequate report of an inaccurate judgment by a biased and variable judge of the extent to which a student has attained an undefined level of mastery of an unknown proportion of an indefinite material.” Miller, Imrie, & Cox 1998, p. 24

Reflects Explicit Criteria

- Provides a clear definition of skills, knowledge or values
- Describes levels of mastery
- Provides feedback on content or skill deficiencies

Exhibits Reliability

- Repeat testing provides similar results
- Testing between students shows same skills or knowledge
- Not dependent on words or situations

Represents Valid Content and Context

- Assesses Higher Level Learning
- Simulates real world experiences
- Includes Multiple Domains
- Samples of the importance of context
Assesses Higher Level Learning

Simulates real world experiences

- Qualitative and quantitative
- Looks, feels and smells like an experience in life
- Includes concepts and decision making
- Something they would see at work

Includes Multiple Domains

- Cognitive
- Skills (psychomotor)
- Affective (beliefs)
Choosing the Right Assessment Tools

<table>
<thead>
<tr>
<th>Assessment Tool</th>
<th>Data Direct or Indirect</th>
<th>Bloom's Knowledge, Comprehension, Application or Analysis/Synthesis/Eval</th>
<th>Pros</th>
<th>Cons</th>
</tr>
</thead>
<tbody>
<tr>
<td>Abbreviation</td>
<td>D or I</td>
<td>K, C, A, ASE</td>
<td>easy to grade objective</td>
<td>reduces assessment to multiple choice answers</td>
</tr>
<tr>
<td><strong>Multiple Choice Exam</strong></td>
<td>D</td>
<td>K, C if carefully constructed A, S, &amp; E</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Licensing Exams</strong></td>
<td>D</td>
<td>K, C, A</td>
<td>easy to score and compare</td>
<td>no authentic testing, may outdate</td>
</tr>
<tr>
<td><strong>Standardized Cognitive Tests</strong></td>
<td>D</td>
<td>K, C, A?</td>
<td>comparable between students</td>
<td>heavily dependent on exposure to topics on test</td>
</tr>
<tr>
<td><strong>Checklists</strong></td>
<td>D</td>
<td>variable</td>
<td>very useful for skills or performances</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>students know exactly what is missing</td>
<td></td>
</tr>
<tr>
<td><strong>Essay</strong></td>
<td>D</td>
<td>K, C, A, ASE</td>
<td>-displays analytical and synthetic thinking well</td>
<td>time consuming to grade, can be subjective</td>
</tr>
<tr>
<td><strong>Case Study</strong></td>
<td>D</td>
<td>K, C, A, ASE</td>
<td>-displays analytical and synthetic thinking well</td>
<td>creating the case is time consuming, dependent on student knowledge form multiple areas</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>-connects other knowledge to topic</td>
<td></td>
</tr>
<tr>
<td><strong>Problem Solving</strong></td>
<td>D</td>
<td>K, C, A, ASE</td>
<td>displays analytical and synthetic thinking well used</td>
<td>difficult to grade due to multiple methods and potential multiple solutions</td>
</tr>
<tr>
<td><strong>Oral Speech</strong></td>
<td>D</td>
<td>variable</td>
<td>easily graded with rubric allows other students to see and learn what each student learned connects general education goals with discipline-specific courses</td>
<td>difficult for ESL students stressful for students takes course time must fairly grade course content beyond delivery</td>
</tr>
<tr>
<td><strong>Debate</strong></td>
<td>D</td>
<td>K, C, A, ASE</td>
<td>provides immediate feedback to the student reveals thinking and ability to respond based on background knowledge and critical thinking ability</td>
<td>requires good rubric more than one evaluator is helpful difficult for ESL students stressful for students takes course time</td>
</tr>
<tr>
<td><strong>Product Creation &amp; Special Reports</strong></td>
<td>D</td>
<td>variable</td>
<td>students can display skills, knowledge, and abilities in a way that is suited to them</td>
<td>must have clearly defined criteria and evaluative measures &quot;the look&quot; can not over-ride the content</td>
</tr>
<tr>
<td><strong>Flowchart or Diagram</strong></td>
<td>D</td>
<td>C, A, ASE</td>
<td>displays original synthetic thinking on the part of the student perhaps the best way to display overall high level thinking and articulation abilities</td>
<td>more difficult to grade, requiring a checklist or rubric for a variety of different answers difficult for some students to do on the spot</td>
</tr>
<tr>
<td>Assessment Tool</td>
<td>Data Direct or Indirect</td>
<td>Bloom's Knowledge, Comprehension, Application or Analysis/Synthesis/Evaluation</td>
<td>Pros</td>
<td>Cons</td>
</tr>
<tr>
<td>---------------------------------------------</td>
<td>-------------------------</td>
<td>--------------------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
<td>----------------------------------------------------------------------</td>
</tr>
<tr>
<td>Portfolios</td>
<td>D</td>
<td>variable K, C, A, ASE</td>
<td>provides the students with a clear record of their work and growth best evidence of growth and change over time students can display skills, knowledge, and abilities in a way that is suited to them promotes self-assessment</td>
<td>time consuming to grade different content in portfolio makes evaluating difficult and may require training bulky to manage depending on size</td>
</tr>
<tr>
<td>Exit Surveys</td>
<td>D, I</td>
<td>ASE</td>
<td>provides good summative data easy to manage data if Likert-scaled responses are used</td>
<td>Likert scales limit feedback, open-ended responses are bulky to manage,</td>
</tr>
<tr>
<td>Performance</td>
<td>D</td>
<td>variable K, C, A, ASE</td>
<td>provides best display of skills and abilities provides excellent opportunity for peer review students can display skills, knowledge, and abilities in a way that is suited to them</td>
<td>stressful for students may take course time some students may take the evaluation very hard - evaluative statements must be carefully framed</td>
</tr>
<tr>
<td>Capstone project or course</td>
<td>D</td>
<td>ASE</td>
<td>best method to measure growth overtime with regards to a course or program - cumulative</td>
<td>focus and breadth of assessment are important understanding all the variables to produce assessment results is also important may result in additional course requirements requires coordination and agreement on standards</td>
</tr>
<tr>
<td>Team Project</td>
<td>D</td>
<td>variable K, C, A, ASE</td>
<td>connects general education goals with discipline-specific courses</td>
<td>must fairly grade individuals as well as team grading is slightly more complicated student interaction may be a challenge</td>
</tr>
<tr>
<td>Reflective self-assessment essay</td>
<td>D, I</td>
<td>ASE</td>
<td>provides invaluable ability to evaluate affective growth in students</td>
<td>must use evidence to support conclusions, not just self-opinionated assessment</td>
</tr>
<tr>
<td>Satisfaction and Perception Surveys</td>
<td>I</td>
<td>C, A, ASE</td>
<td>provides good indirect data data can be compared longitudinally can be used to determine outcomes over a long period of time</td>
<td>respondents may be influenced by factors other than those being considered validity and reliability most be closely watched</td>
</tr>
</tbody>
</table>
EXAMPLES OF AUTHENTIC ASSESSMENT

Student Services

Student focus groups after Probation/Success Workshop

Student Educational Plan

Instruction

Pamphlets

Live patients

Flow Charts

Products
**ASCCC Definitions**

**Assessment:** The root word for assessment is assidere - to sit beside. Assessment is a means of using explicit criteria to determine evaluative measures to help facilitate student success.

**Goal:** A Goal is a statement of intent or vision that is not necessarily measurable.

**Measurable Objectives:** Measurable Objectives are small steps that lead toward a goal.

**Measurability:** Measurability refers to both qualitative and quantitative means of measuring.

**Student Learning Outcomes (SLO):** Student Learning Outcomes refer to overarching specific observable characteristics developed by local faculty that allow them to determine or demonstrate evidence that learning has occurred as a result of a specific course, program, activity, or process.

<table>
<thead>
<tr>
<th>Goals</th>
<th>Objectives</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="image" alt="Goal Image" /></td>
<td><img src="image" alt="Objectives Image" /></td>
<td><img src="image" alt="Outcomes Image" /></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Goals</th>
<th>Objectives</th>
<th>Outcomes</th>
</tr>
</thead>
<tbody>
<tr>
<td>A goal is a statement of intent or vision that is not necessarily measurable. The aim, the vision, usually the catalog description of a course or program.</td>
<td>Measurable Objectives are small steps that lead toward a goal.</td>
<td>SLOs overarching specific observable characteristics, developed by local faculty, to determine or demonstrate evidence that learning has occurred as a result of a specific course, program, activity, or process.</td>
</tr>
</tbody>
</table>
Faculty and Assessment –

- Faculty are NOT motivated to do assessment because of accreditation.
- 93% of faculty responded that measurable improvement was an important motivator to improve student learning and do assessment.
- 82% of faculty reported that discovering what works and what does not in the classroom was an important motivator to do assessment.
- Consistency and fairness in grading was an important motivator for 69% of those surveyed.
- “Post secondary assessment done right must be rooted in the course and in the classroom, in the individual cells, to speak metaphorically, where the metabolism of learning actually takes place” (Wright, 1999)
- “Classroom assessment is the purest form of assessment-for-improvement, because the information gleaned can be immediately used to improve teaching and learning ...the further away from the individual classroom you get, the harder it becomes to turn assessment data into usable information” (Miller, 1997)

Black and Wiliam reported that diagnostic feedback produced an effect size of 0.4 – 0.7, the largest of any innovation in education. Translated this means that the average pupil in international comparative studies in mathematics would have raised the score of a nation in the middle of the pack of 41 countries (e.g. the U.S.) to one of the top five. Or alternatively, an average gain of 30 percentile points, two letter grades, or 100 points on the SAT scale.
The Process of Assessment and SLO development

- Helps to improve teaching practices
- Formalizes thoughts about courses and provides a holistic picture of the course from beginning to end
- Focuses teaching practices, syllabi, daily activities, and assessments on a single target--SLOs
- Validates both what we are teaching and why we are teaching it
- Promotes robust dialogue among the faculty & stimulates productive departmental conversations
- Reduces departmental favoritism because it is based on performance
- Enhances interdisciplinary cooperation
- Produces useful discussions concerning sequential courses
- Contributes to more rigorous curriculum review with a focus on outcomes
- Encourages consistency of standards between sections
- Maintains high standards
- Directs teaching to be more learning-centered
- Improves student learning by focusing on good practices

**Seven Principles for Good Practice in Undergraduate Education**

Good practice in undergraduate education:

1. Encourages contact between students and faculty,
2. Develops reciprocity and cooperation among students,
3. Encourages active (verses passive) learning,
4. Provides prompt feedback,
5. Emphasizes task on time,
6. Communicates high expectations, and
7. Respects diverse talents and ways of learning.

---

1 Chickering and Gamson
Seven Common (Mis) Perceptions about Outcomes Assessment

1. We are doing just fine without it.
   Assessment is only needed when things are going wrong – medicine for the sick.

2. We’re already doing it.
   Assessment is just old wine in new bottles.

3. We’re far too busy to do it.
   Assessment is an “administrivial” job.

4. The most important things we do can't/shouldn't be measured.
   Assessment is too quantitative; we produce long term or ineffable results in education.

5. We would need more staff and lots of money to do assessment.
   Assessment is too complex and expensive.

6. They'll use the results against us.
   Assessment is a trick or Trojan horse, the next thing you know we will be evaluated on ratios.

7. No one will care about or use assessment.
   Assessment is a waste of time.

T.A. Angelo – University of Akron – Institute for Teaching and Learning
The Case for Authentic Assessment: WHAT IS AUTHENTIC ASSESSMENT? By Grant Wiggins

Assessment is authentic when we directly examine student performance on worthy intellectual tasks. Traditional assessment, by contract, relies on indirect or proxy 'items'--efficient, simplistic substitutes from which we think valid inferences can be made about the student's performance at those valued challenges.

Do we want to evaluate student problem-posing and problem-solving in mathematics? experimental research in science? speaking, listening, and facilitating a discussion? doing document-based historical inquiry? thoroughly revising a piece of imaginative writing until it "works" for the reader? Then let our assessment be built out of such exemplary intellectual challenges.

Further comparisons with traditional standardized tests will help to clarify what "authenticity" means when considering assessment design and use:

* Authentic assessments require students to be effective performers with acquired knowledge. Traditional tests tend to reveal only whether the student can recognize, recall or "plug in" what was learned out of context. This may be as problematic as inferring driving or teaching ability from written tests alone. (Note, therefore, that the debate is not "either-or": there may well be virtue in an array of local and state assessment instruments as befits the purpose of the measurement.)

* Authentic assessments present the student with the full array of tasks that mirror the priorities and challenges found in the best instructional activities: conducting research; writing, revising and discussing papers; providing an engaging oral analysis of a recent political event; collaborating with others on a debate, etc. Conventional tests are usually limited to paper-and-pencil, one-answer questions.

* Authentic assessments attend to whether the student can craft polished, thorough and justifiable answers, performances or products. Conventional tests typically only ask the student to select or write correct responses--irrespective of reasons. (There is rarely an adequate opportunity to plan, revise and substantiate responses on typical tests, even when there are open-ended questions). As a result,

* Authentic assessment achieves validity and reliability by emphasizing and standardizing the appropriate criteria for scoring such (varied) products; traditional testing standardizes objective "items" and, hence, the (one) right answer for each.

* "Test validity" should depend in part upon whether the test simulates real-world "tests" of ability. Validity on most multiple-choice tests is determined merely by matching items to the curriculum content (or through sophisticated correlations with other test results).

* Authentic tasks involve "ill-structured" challenges and roles that help students rehearse for the complex ambiguities of the "game" of adult and professional life. Traditional tests are more like drills, assessing static and too-often arbitrarily discrete or simplistic elements of those activities.

Wiggins, Grant (1990). The case for authentic assessment. Practical Assessment, Research & Evaluation, 2(2). Retrieved February 16, 2004 from http://PAREonline.net/getvn.asp?v=2&n=2. Copyright 1990, PAREonline.net. Permission is granted to distribute this article for nonprofit, educational purposes if it is copied in its entirety and the journal is credited. Please notify the editor if an article is to be used in a newsletter
Resources


