

Direct Methods to Assess Student Learning

WASC/AAHE Collaborative Workshop on Building Learner-Centered Institutions—Developing
Institutional Strategies for Assessing and Improving Student Learning

Mary J. Allen, March 2004

Two Basic Ways to Assess Student Learning:

1. **Direct** – The assessment is based on an analysis of student behaviors or products in which they demonstrate how well they have mastered learning objectives.
 2. **Indirect** – The assessment is based on an analysis of reported perceptions about student mastery of learning objectives. The perceptions may be self-reports by students, or they may be made by others, such as peers, alumni, fieldwork supervisors, employers, or faculty.
-

Strategies for Direct Assessment of Student Learning

1. Published Tests
 2. Locally-Developed Tests
 3. Embedded Assignments and Course Activities
 4. Competence Interviews
 5. Portfolios
 6. Collective Portfolios
-

Properties of Good Assessment Techniques

- **Valid**—directly reflects the learning objective being assessed
 - **Reliable**—including inter-rater reliability when subjective judgments are made
 - **Actionable**—results point reviewers toward challenges that can be approached
 - **Efficient and cost-effective** in time and money
 - **Engaging to students and other respondents**—so they'll demonstrate the extent of their learning
 - **Interesting to faculty and other stakeholders**—they care about results and are willing to act on them
 - **Triangulation**—multiple lines of evidence point to the same conclusion
-

Published Tests

Some Examples of Published Tests		
GRE General Test	verbal, quantitative, and analytical skills	http://www.gre.org/
GRE Subject Tests	specific disciplines, such as Biochemistry and Psychology	http://www.gre.org/
Major Field Tests	undergraduate student achievement in a number of disciplines, such as Biology, Business, Chemistry, and History	http://www.ets.org/hea/mft/
Praxis Series tests	aspects of teacher competence	http://www.ets.org/praxis/index.html
Academic Profile	"college-level reading, critical thinking, writing, and mathematics in the context of materials from the humanities, social sciences, and natural sciences"	http://www.ets.org/hea/acpro
Collegiate Assessment of Academic Proficiency (CAAP)	"assesses college students' academic achievement in core general education skills" (writing, reading, math, science reasoning, and critical thinking)	http://www.act.org/caap/index.html
ACCUPLACER	reading, writing, and mathematics	http://www.collegeboard.com/highered/apr/accu/accu.html
COMPASS e-Write	writing	http://www.act.org/e-write/index.html

Steps in Selecting a Published Test

1. Identify a possible test.
2. Consider published reviews of this test, such as reviews in the *Mental Measurements Yearbook*.
3. Order a specimen set from the publisher.
4. Take the test and consider the appropriateness of its format and content.
5. Consider the test's relationship to your learning objectives.
6. Consider the depth of processing of the items (e.g., analyze items using Bloom's taxonomy).
7. Consider the publication date and currency of the items.
8. How many scores are provided? Will these scores be useful? How?
9. Look at the test manual. Were test development procedures reasonable? What is the evidence for the test's reliability and validity for the intended use?
10. If you will be using the norms, consider their relevance for your purpose.
11. Consider practicalities, e.g., timing, test proctoring, and test scoring requirements.
12. Verify that faculty are willing to act on results.

Published Test Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • They generally are carefully developed, highly reliable, professionally scored, and nationally normed. • They frequently provide a number of norm groups, such as norms for community colleges, liberal arts colleges, and comprehensive universities. • Online versions of tests are increasingly available, and some provide immediate scoring. • Some publishers allow faculty to supplement tests with their own items, so tests can be adapted to better serve local needs. 	<ul style="list-style-type: none"> • Students may not take the test seriously if test results have no impact on their lives. • These tests are not useful as direct measures for program assessment if they do not align with local curricula and learning objectives. • Test scores may reflect criteria that are too broad for meaningful assessment. • Most published tests rely heavily on multiple-choice items which often focus on specific facts, but program learning objectives more often emphasize higher-level skills. • If the test does not reflect the learning objectives that faculty value and the curricula that students experience, results are likely to be discounted and inconsequential. • Tests can be expensive. • The marginal gain from annual testing may be low. • Faculty may object to standardized exam scores on general principles, leading them to ignore results.

Locally-Developed Tests

Common Test Item Formats	
Item Type	Characteristics and Suggestions
Completion	These items require students to fill-in-the-blank with appropriate terms or phrases. They appear to be best for testing vocabulary and basic knowledge, and they avoid giving students credit for guessing by requiring recall, rather than recognition. Scoring can be difficult if more than one answer can be correct.
Essay	Essay questions are very popular and can be used to assess higher-order thinking skills. They generally ask for explanations and justifications, rather than memorized lists. Key words in essay questions are <i>summarize, evaluate, contrast, explain, describe, define, compare, discuss, criticize, justify, trace, interpret, prove, and illustrate</i> (Moss & Holder, 1988).
Matching	Usually these questions are presented as two columns, and students are required to associate elements in column B with elements in column A. Such items are easy to score, but they are relatively difficult to construct and they seem best suited for testing knowledge of factual information, rather than deeper levels of understanding.
Multiple-Choice	Multiple-choice questions are popular because they can measure many concepts in a short period of time, and they generally are better than other objective questions at assessing higher-order thinking. They are easy to score, and item banks associated with popular textbooks are often available. Writing good items takes time, and there is strong temptation to emphasize facts, rather than understanding.
True-False	True-false items are relatively easy to construct and grade, but they appear to be best at assessing factual knowledge, rather than deep understanding.

Consider augmenting or replacing traditional tests with

- Authentic tests. For example, students may be asked to respond to real-world case studies or to work with peers on a problem-based learning task, or an essay question might give students a task, a role, and an audience:
 1. You have a Ph.D. in Forensic Psychology and are asked to provide training to members of the State Parole Board to help them make better decisions about when to keep convicted criminals in prison and when to grant them parole. Your presentation must be accurate and must contain advice consistent with the literature. The State Parole Board is a group of intelligent laymen without training in psychology, so your presentation must be in language they can understand, with key psychological terms defined. Write the script for your presentation here.
 2. You are an expert in social psychology with a specialization in persuasion. An administrator at the central office for the high school district asks you to help them develop a campaign to reduce the amount of litter students leave on campus during lunch and after school. The administrator has no formal psychological training, but is very interested in the research underlying your suggestions. As an expert on persuasion, what advice will you give? Be sure to include a summary of the relevant research.
- Performance tests

Locally-Developed Test Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • Appropriate mixes of essay and objective questions allow faculty to address various types of learning objectives. • Students generally are motivated to display the extent of their learning. • If well-constructed, they are likely to have good validity. • Because local faculty write the exam, they are likely to be interested in results and willing to use them. • Can be integrated into routine faculty workloads. • Campuses with similar missions may decide to develop their own norms, and they may decide to assess student work together or to provide independent assessment of each other's student work. • The evaluation process should directly lead faculty into discussions of student learning, curriculum, pedagogy, and student support services. 	<ul style="list-style-type: none"> • These exams are likely to be less reliable than published exams. • Reliability and validity generally are unknown. • Creating and scoring exams takes time. • Traditional testing methods have been criticized for not being "authentic." • Norms generally are not available.

Embedded Assignments and Course Activities

- Classroom assessment activities (Angelo, T. A., & Cross, K. P. (1993). *Classroom assessment techniques: A handbook for college teachers* (2nd ed.). San Francisco, CA: Jossey-Bass.)
- Community-service learning and other fieldwork activities
- Culminating projects, such as senior theses and papers in capstone courses
- Exams or parts of exams
- Group projects and presentations
- Homework assignments
- In-class presentations
- In-class writing assignments
- Portfolios (course or program) and collective portfolios
- Poster presentations and student research conferences
- Reflective essays
- Student recitals and exhibitions

Assignments and activities are purposefully created to collect information relevant to specific program learning objectives. Results are pooled across courses and instructors to indicate program accomplishments, not just the learning of students in specific courses.

Embedded Assignments and Course Activities Strengths and Weaknesses
--

Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • Out-of-class assignments are not restricted to time constraints typical for exams. • Students are generally motivated to demonstrate the extent of their learning. • Can provide authentic assessment of learning objectives. • Can involve CSL or other fieldwork activities and ratings by fieldwork supervisors. • Can provide a context for assessing communication and teamwork skills. • Can be used for grading as well as assessment. • Faculty who develop the procedures are likely to be interested in results and willing to use them. • The evaluation process should directly lead faculty into discussions of student learning, curriculum, pedagogy, and student support services. • Data collection is unobtrusive to students. 	<ul style="list-style-type: none"> • Requires time to develop and coordinate. • Requires faculty trust that the program will be assessed, not individual teachers. • Reliability and validity generally are unknown. • Norms generally are not available.

Competence Interviews

- Can be structured or unstructured interviews
 - Faculty can work alone or in teams
 - Use of open-ended vs. close-ended questions
 - Obvious applications in foreign language courses and professional programs
 - Requires careful attention to the purpose of the interviews, the development of interview protocols, and interviewer training.
 - Can be used to supplement or follow-up on written exams or other products (e.g., theses)
-

Competence Interview Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • The interview format allows faculty to probe for the breadth and extent of student learning. • Can be combined with other techniques that more effectively assess knowledge of facts and terms. • Can involve simulated interactions with clients. • Can also provide for direct assessment of some student skills, such as oral communication, critical thinking, and problem-solving skills. 	<ul style="list-style-type: none"> • Requires time to develop, coordinate, schedule, and implement. • Interview protocols must be carefully developed. • Subjective judgements must be guided by agreed-upon criteria. • Training of interviewers takes time. • Interviewing using unstructured interviews requires expertise. • Not an efficient way to assess knowledge of specific facts and terms. • Some students may be intimidated by the process, reducing their ability to demonstrate their learning.

Portfolios

- Showcase vs. Developmental Portfolios: best work vs. evidence of growth
- Workload and storage demands for large programs can be overwhelming!

Some Questions to Answer Before Assigning Portfolios

1. What is the purpose of the requirement—to document student learning, to demonstrate student development, to learn about students' reflections on their learning, to create a document useful to students, to help students grow through personal reflection on their personal goals?
 2. When and how will students be told about the requirement, including what materials they need to collect or to produce for it?
 3. Will the portfolios be used developmentally or will they be submitted only as students near graduation?
 4. Will portfolios be showcase or developmental?
 5. Are there minimum and maximum lengths or sizes for portfolios?
 6. Who will decide which materials will be included in portfolios—faculty or students?
 7. What elements will be required in the portfolio—evidence only from courses in the discipline, other types of evidence, evidence directly tied to learning objectives, previously graded products or clean copies?
 8. Will students be graded on the portfolios? If so, how and by whom?
 9. How will the portfolios be assessed to evaluate and improve the program?
 10. What can be done for students who have inadequate evidence through no fault of their own?
 11. What will motivate students to take the portfolio assignment seriously?
 12. How will the portfolio be submitted—hard copy or electronic copy?
 13. Who “owns” the portfolios—students or the program?
 14. Who has access to the portfolios and for what purposes?
 15. How will student privacy and confidentiality be protected?
-

Portfolio Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • Students are encouraged to take responsibility for and pride in their learning. • Students may become more aware of their own academic growth. • Can be used for developmental assessment and can be integrated into the advising process to individualize student planning. • Can help faculty identify curriculum gaps, lack of alignment with objectives. • Students can use portfolios and the portfolio process to prepare for graduate school or career applications. • The evaluation process should directly lead faculty into discussions of student learning, curriculum, pedagogy, and student support services. • Webfolios or CD-ROMs can be easily viewed, duplicated, and stored. 	<ul style="list-style-type: none"> • Requires faculty time to prepare the portfolio assignment and assist students as they prepare them. • Requires faculty analysis and, if graded, faculty time to assign grades. • May be difficult to motivate students to take the task seriously. • May be more difficult for transfer students to assemble the portfolio if they haven't saved relevant materials. • Students may refrain from criticizing the program if their portfolio is graded or if their names will be associated with portfolios during the review.

Collective Portfolios

Some of the benefits of traditional portfolios, with much less work!

Collective Portfolio Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> • Can provide direct evidence of student mastery of learning objectives. • Students generally are motivated to display the extent of their learning. • Workload demands generally are more manageable than traditional portfolios. • Can help faculty identify curriculum gaps, lack of alignment with objectives. • Students are not required to do extra work. • The evaluation process should directly lead faculty into discussions of student learning, curriculum, pedagogy, and student support services. • Data collection is unobtrusive to students. 	<ul style="list-style-type: none"> • If assignments are not aligned with the objectives being examined, evidence may be problematic. • If sampling is not done well, results may not generalize to the entire program. • Reviewing the materials takes time and planning.

Most of the materials in this handout are based on or extracted from:
 Allen, M. J. (2004). *Assessing Academic Programs in Higher Education*. Bolton, MA: Anker.