
Strategies for Direct and Indirect Assessment of Student Learning

Mary J. Allen, mallen@csub.edu

SACS-COC Summer Institute

July 28, 2008

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 outcomes.
 2. **Indirect** – The assessment is based on an analysis of reported perceptions about student mastery of learning outcomes.
-

Properties of Good Assessment Techniques

- Valid—directly reflects the learning outcome being assessed
 - Reliable—especially inter-rater reliability when subjective judgments are made
 - Actionable—results help faculty identify what students are learning well and what requires more attention
 - 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
-

Strategies for Direct Assessment of Student Learning

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

Examples of Published Tests

Some Examples of Published Tests		
Measure of Academic Proficiency and Progress (MAPP; replaced the Academic Profile in Jan. 2006)	“College-level reading, mathematics, writing, and critical thinking in the context of the humanities, social sciences, and natural sciences” (can be taken and scored online; essay section is optional)	http://www.ets.org
Collegiate Learning Assessment (CLA)	critical thinking, analytic reasoning, writing skills; based on open-ended questions	http://www.cae.org/content/pro_collegiate.htm
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
iSkills	Seven Information and Communication Technology literacy skills, including data searches, email, and software use. Two versions: core (lower-division) and advanced.	http://www.ets.org/Media/Products/ICT_Literacy/demo2/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 outcomes.
 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 outcomes. • 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 outcomes. • 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 outcomes more often emphasize higher-level skills. • If the test does not reflect the learning outcomes 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.

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 outcomes. • Appropriate mixes of essay and objective questions allow faculty to address various types of learning outcomes. • Students generally are motivated to display the extent of their learning if they are being graded on the work. • 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. • 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

- Community-service learning and other fieldwork activities
- Culminating projects, such as papers in capstone courses
- Exams or parts of exams
- Group projects
- Homework assignments
- In-class presentations
- Student recitals and exhibitions
- Comprehensive exams, theses, dissertations, and defense interviews.

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

Consider integrating “**signature assignments**” into the curriculum, i.e., assignments designed to assess specific learning outcomes. Assignments might be developed as “threshold, milestone, or capstone assessments” [AAC&U (2005) *Liberal Education Outcomes: A Preliminary Report on Student Achievement in College*].

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 outcomes. • Out-of-class assignments are not restricted to time constraints typical for exams. • Students are generally motivated to demonstrate the extent of their learning if they are being graded. • Can provide authentic assessment of learning outcomes. • 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.

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 outcomes, 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 outcomes. • 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 outcomes. • 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. • E-portfolios 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 outcomes. • 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 outcomes. • 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 outcomes 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.

Strategies for Indirect Assessment of Student Learning

- Surveys
- Interviews
- Focus Groups

Surveys

- Point-of-contact surveys
- Online, e-mailed, registration, or grad check surveys
- Keep it simple!

Common Survey Formats	
Type of Item	Example
Check list	Please indicate which of the activities you feel competent to perform. <input type="checkbox"/> Develop an investment plan <input type="checkbox"/> Interpret a financial report <input type="checkbox"/> Provide feedback about an employee's performance <input type="checkbox"/> Write a case study
Classification	Organization of the paper: <input type="checkbox"/> Confusing, unclear <input type="checkbox"/> Generally clear, minor points of confusion <input type="checkbox"/> Clear, logical, easy to follow
Frequency	In a typical term, I used the department's computer lab: Never 1-2 times 3-5 times 6 or more times

Importance	How important is it for the department to provide career counseling? Unimportant Slightly Moderately Very Extremely Important Important Important Important
Linear rating scale	Ability to compose paragraphs in standard, written English. Unsatisfactory ___ ___ ___ ___ ___ ___ ___ Excellent
Likert scale	I am able to write a research paper using MLA standards. Strongly Disagree Disagree Neutral Agree Strongly Agree
Open-ended	Please describe the most important concepts you learned in the program.
Partially close-ended	Please check the most important factor that led you to major in engineering. ___ Experience in a specific course ___ Experience with a specific instructor ___ Work experience in this or a related field ___ Advice from a career planning office or consultant ___ Advice from family member or friend ___ Personal interest ___ Other: please explain
Ranking	Please indicate your ranking of the importance of the following learning outcomes by assigning ranks from “1” to “4,” where a “1” is most important and “4” is least important. ___ Computing ___ Critical thinking ___ Speaking ___ Writing

Gap Analysis

Sometimes it is useful to ask respondents to rate a set of items twice: once to indicate their importance and once to indicate the extent of their achievement. Differences (gaps) between the two ratings receive particular attention when interpreting results, especially items that are judged to be important and not well achieved.

Survey Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> ● Are flexible in format and can include questions about many issues. ● Can be administered to large groups of respondents. ● Can easily assess the views of various stakeholders. ● Usually has face validity—the questions generally have a clear relationship to the outcomes being assessed. ● Tend to be inexpensive to administer. ● Can be conducted relatively quickly. ● Responses to close-ended questions are easy to tabulate and to report in tables or graphs. ● Open-ended questions allow faculty to uncover unanticipated results. ● Can be used to track opinions across time to explore trends. ● Are amenable to different formats, such as paper-and-pencil or online formats. ● Can be used to collect opinions from respondents at distant sites. 	<ul style="list-style-type: none"> ● Provides indirect evidence about student learning. ● Their validity depends on the quality of the questions and response options. ● Conclusions can be inaccurate if biased samples are obtained. ● Results might not include the full array of opinions if the sample is small. ● What people say they do or know may be inconsistent with what they actually do or know. ● Open-ended responses can be difficult and time-consuming to analyze.

Interviews

- Interviews can be conducted one-on-one, in small groups, or over the phone.
- Interviews can be structured (with specified questions) or unstructured (a more open process).
- Questions can be close-ended (e.g., multiple-choice style) or open-ended (respondents construct a response).
- Can target students, graduating seniors, alumni, employers, community members, faculty, etc.
- Can do exit interviews or pre-post interviews.
- Can focus on student experiences, concerns, or attitudes related to the program being assessed.
- Generally should be conducted by neutral parties to avoid bias and conflict of interest.

Some Tips for Effective Interviewing

- Conduct the interview in an environment that allows the interaction to be confidential and uninterrupted.
- Demonstrate respect for the respondents as *participants* in the assessment process rather than as *subjects*. Explain the purpose of the project, how the data will be used, how the respondent's anonymity or confidentiality will be maintained, and the respondents' rights as participants. Ask if they have any questions.
- Put the respondents at ease. Do more listening than talking. Allow respondents to finish their statements without interruption.
- Match follow-up questions to the project's objectives. For example, if the objective is to obtain student feedback about student advising, don't spend time pursuing other topics.
- Do *not* argue with the respondent's point of view, even if you are convinced that the viewpoint is incorrect. Your role is to obtain the respondents' opinions, not to convert them to your perspective.
- Allow respondents time to process the question. They may not have thought about the issue before, and they may require time to develop a thoughtful response.
- Paraphrase to verify that you have understood the respondent's comments. Respondents will sometimes realize that what they said isn't what they meant, or you may have misunderstood them. Paraphrasing provides an opportunity to improve the accuracy of the data.
- Make sure you know how to record the data and include a backup system. You may be using a tape recorder—if so, consider supplementing the tape with written notes in case the recorder fails or the tape is faulty. Always build in a system for verifying that the tape is functioning or that other data recording procedures are working. Don't forget your pencil and paper!

Interview Strengths and Weaknesses

Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> ● Are flexible in format and can include questions about many issues. ● Can assess the views of various stakeholders. ● Usually has face validity—the questions generally have a clear relationship to the outcomes being assessed. ● Can provide insights into the reasons for participants' beliefs, attitudes, and experiences. ● Interviewers can prompt respondents to provide more detailed responses. ● Interviewers can respond to questions and clarify misunderstandings. 	<ul style="list-style-type: none"> ● Generally provides indirect evidence about student learning. ● Their validity depends on the quality of the questions. ● Poor interviewer skills can generate limited or useless information. ● Can be difficult to obtain a representative sample of respondents. ● What people say they do or know may be inconsistent with what they actually do or know. ● Can be relatively time-consuming and expensive to conduct, especially if interviewers and interviewees are paid or if the no-show rate for scheduled interviews is high.

<ul style="list-style-type: none"> • Telephone interviews can be used to reach distant respondents. • Can provide a sense of immediacy and personal attention for respondents. • Open-ended questions allow faculty to uncover unanticipated results. 	<ul style="list-style-type: none"> • The process can intimidate some respondents, especially if asked about sensitive information and their identity is known to the interviewer. • Results can be difficult and time-consuming to analyze. • Transcriptions of interviews can be time-consuming and costly.
--	---

Focus Groups

- **Traditional focus groups** are free-flowing discussions among small, homogeneous groups (typically from 6 to 10 participants), guided by a skilled facilitator who subtly directs the discussion in accordance with pre-determined objectives. This process leads to in-depth responses to questions, generally with full participation from all group members. The facilitator departs from the script to follow promising leads that arise during the interaction.
- **Structured group interviews** are less interactive than traditional focus groups and can be facilitated by people with less training in group dynamics and traditional focus group methodology. The group interview is highly structured, and the report generally provides a few core findings, rather than an in-depth analysis.

Sample Focus Group Questions	
Purpose of Question	Examples
Warm-up	<ul style="list-style-type: none"> • I'd like everyone to start out by stating a word or phrase that best describes your view of the program.
Issue 1: Career Preparation	<ul style="list-style-type: none"> • Please tell us what career you are interested in pursuing after graduation. • How has the program helped you prepare for your career or future activities?
Issue 2: Advising	<ul style="list-style-type: none"> • We are interested in your advising experiences in the program. Could you tell us about your first advising experience in the department? • What did you find most useful in your interactions with your advisor? • What would you like our advisors to do differently?
Issue 3: Curriculum	<ul style="list-style-type: none"> • Thinking about the curriculum and the required courses, how well do you think they prepared you for upper-division work? • What should be changed about the curriculum to better prepare you for your career or for graduate school?
Closing	<ul style="list-style-type: none"> • We've covered a lot of ground today, but we know you might still have other input about the program. Is there anything you would like to say about the program that hasn't been discussed already?

Focus Group Strengths and Weaknesses	
Potential Strengths	Potential Weaknesses
<ul style="list-style-type: none"> ● Are flexible in format and can include questions about many issues. ● Can provide in-depth exploration of issues. ● Usually has face validity—the questions generally have a clear relationship to the outcomes being assessed. ● Can be combined with other techniques, such as surveys. ● The process allows faculty to uncover unanticipated results. ● Can provide insights into the reasons for participants' beliefs, attitudes, and experiences. ● Can be conducted within courses. ● Participants have the opportunity to react to each other's ideas, providing an opportunity to uncover the degree of consensus on ideas that emerge during the discussion. 	<ul style="list-style-type: none"> ● Generally provides indirect evidence about student learning. ● Requires a skilled, unbiased facilitator. ● Their validity depends on the quality of the questions. ● Results might not include the full array of opinions if only one focus group is conducted. ● What people say they do or know may be inconsistent with what they actually do or know. ● Recruiting and scheduling the groups can be difficult. ● Time-consuming to collect and analyze data.

References

Allen, M. J. (2004) *Assessing Academic Programs in Higher Education*. Bolton, MA: Sage.

Allen, M. J. (2006). *Assessing General Education Programs*. Bolton, MA: Sage.

Krueger, R.A., & Casey M.A. (2000). *Focus Groups: A Practical Guide for Applied Research* (3rd ed.). Thousand Oaks, CA: Sage.