

**Direct Assessment Methods --
A Close-Up Look**

by

Barbara D. Wright

Associate Director,
Accrediting Commission for Senior Colleges and Universities
Western Association of Schools and Colleges
Alameda, CA 94501

bwright@wascsenior.org

January 30, 2009

Portfolios: collections of student work (and sometimes other material) intended to illustrate achievement of learning outcomes. The mantra is “collect, select, reflect, connect.”

Advantages:

- Are adaptable to different
 - *levels* of assessment (i.e. individual student, program, institution)
 - *purposes* (i.e. cross-sectional snapshot; change/progress over time)
 - *kinds* of materials (i.e. written work, tapes of performances, student self- assessments)
- Can tell us where student are *and* how they got there
- Emphasize human judgment, meaning-making
- Provide information likely to be *used*
- Have become extremely popular, hence an easy sell
- Engage students, faculty
- Are *educational* for both students and faculty
- Reduce fears of misuse
- Can be managed by students – to some extent
- Are supported by many different software programs

Disadvantages:

- Can be labor-intensive
- Can be cumbersome to store, navigate through
- Must relate contents to articulated outcomes
- Require carefully defined criteria for review, e.g. rubrics
- Require training for reviewers
- Require distinguishing between usefulness of the portfolio for *students* (e.g., to showcase work, impress prospective employers, inform advisors) and for *assessment of learning*

Solutions/responses:

- Collect *samples* of work, not everything from everybody
- Use electronic storage and retrieval
- Give students responsibility for maintaining the portfolio
- Invest in outcomes, because they're the basis for everything anyway
- Invest in good criteria for education's sake
- Invest in training for faculty development's sake

Capstones: a wide variety of culminating projects, assignments, performances, or even experiences, e.g., faculty-supervised community service, internships

Advantages:

- Are cumulative
- Are integrative
- Are adaptable to demonstration of
 - skills
 - general education
 - professional field or major
 - dispositions
 - institutional outcomes
 - combinations
- Are motivating for students
- Set standards for degree completion, graduation
- Provide an occasion for department-level discussion, interpretation
- Invite external evaluation
- Help students make the transition to
 - self-assessment
 - professional assessment
 - life-long learning

Disadvantages:

- Pose challenge of capturing all students in their final year/semester
- Differences within/among majors demand flexibility plus commonality
- May mean an additional course requirement
- Require coordinating multiple dimensions of learning & assessment
- Can be labor-intensive
- Must relate to carefully articulated outcomes
- Require carefully defined criteria for review, e.g. rubrics
- Require distinguishing between purpose of the capstone for *students* and for *program assessment*

Solutions/responses:

- Require the capstone for graduation
- Introduce as widely as possible across the institution
- Include capstone experiences within existing courses
- Provide resources, staff support
- View resources, labor, as worthwhile investment

Performances: activities, live or recorded, designed to demonstrate specific outcomes, e.g. a poster presentation, conduct of a class, a musical or theatrical performance, client counseling, facilitation of a group discussion, "think aloud" analysis of a text.

Advantages:

- Have face validity in terms of preparation for student's real-life goals
- Put emphasis on what the student can *do* (as opposed to knowing *about*):
 - require application
 - may require spontaneous adaptation, problem-solving
 - are integrative
 - provide a reality check
- Give students with practical intelligence, skills, a chance to shine
- Can elicit affective outcomes, e.g. poise, grace under pressure
- Are motivating, encourage practice, rehearsing
- Put the emphasis on active learning
- Promote coaching relationship between students and faculty, especially when there are external reviewers
- Promote self-assessment, internalization of standards
- Are highly adaptable, even to liberal arts

Disadvantages:

- Can be labor-intensive, time-consuming, expensive
- Must relate to articulated outcomes
- Require careful definition of criteria, e.g. rubrics
- Require careful training of reviewers, including external reviewers
- Require coordination, scheduling, esp. of external reviewers
- May frighten off insecure students

Solutions/responses:

- Review a *sample* of students
- Embed in routine, non-threatening situations (e.g., internship, clinical setting)
- Use digital means to make performances accessible to reviewers
- Regard outcomes, criteria, and training as an educational investment
- Remind students they must demonstrate employability

Common assignments, template assignments, secondary readings, and other embedded assessments: student work produced in response to a course assignment is examined for multiple purposes, e.g., to determine command of course material but also to assess writing skill, information literacy, critical thinking, etc.

- “Common assignments”: the *same* assignment across multiple courses;
- “template assignments”: the same *format* but not identical assignment across multiple courses
- “Secondary readings”: student work is examined “secondarily” for other qualities beyond command of course material.

Advantages:

- Use work produced by students as a normal part of their course work
- Solve the problem of quality of student effort
- Are efficient, low-cost
- Have face validity
- Provide maximally useful information for improvement with minimum slippage
- Encourage discussion, collaboration among faculty & support staff
- Can create campus-wide interest

Disadvantages:

- Require considerable coordination
- Can be time-consuming to create, implement
- Can be time-consuming, labor-intensive to score
- Must be designed in relation to specific outcomes
- Require careful definition of criteria for review, e.g., rubrics
- Require careful training of reviewers

Solutions/responses:

- Focus on what’s important
- Use “common questions” if an entire common assignment is impractical
- Regard outcomes, criteria, and training as an educational investment
- Provide support, “teaching circles” to discuss implementation, findings
- Remember the efficiencies, benefits
- Make the investment

Course management programs: Software that allows faculty to set up chat rooms, threaded discussions, etc., and capture student responses

Advantages:

- Are adaptable to wide range of learning goals, disciplines, environments
- Use work produced electronically by students as a normal part of course participation
- Record threaded discussions, chat, ephemera that are impossible or cumbersome to capture face to face
- Give quiet students an opportunity to shine
- Can preserve a large volume of material, allow sorting, retrieval, data analysis
- Are efficient, low-cost
- Are unintrusive
- Solve the problem of quality of student effort
- Allow prompt feedback
- Develop students' metacognition when assessment results are shared
- Often include tests, quizzes, tasks as part of package, supporting multiple-method approach, convenience

Disadvantages:

- Rely heavily on student writing skill, comfort with technology
- Pose challenges to higher levels of aggregation beyond individual course or student
- May discourage collaboration among faculty, staff, programs
- Managing large volume of material can be difficult, intimidating
- "No significant difference" bias may short circuit improvement
- Tests, quizzes may promote recall, surface rather than deep learning
- Built-in survey tools encourage collection of indirect rather than direct evidence
- Direct observation of student performances is difficult or impossible
- Software may drive the assessment effort, instead of assessment goals and values driving choice, use of the software

Solutions/responses:

- Develop good, focused outcomes, criteria, rubrics
- Use built-in data management tools
- Supplement if necessary, e.g. with "The Rubric Processor"
- Invest in training of faculty, external reviewers
- Use tests, quizzes with caution, supplement with authentic tasks
- Negotiate with the maker, customize the software
- Aim for program-level, not just individual or course-level improvement

Classroom Assessment/Research: an approach to assessment pioneered by K. Patricia Cross and Thomas A. Angelo; provides a large collection of techniques individual instructors can use in their classrooms to discover what students are learning – or not – and to make rapid adjustments.

Advantages:

- Takes place at ground zero of learning process for:
 - maximum relevance, usefulness
 - minimum slippage
- Offers maximum privacy, minimum risk, anxiety
- Is conducted continuously, has formative benefit
- Can provide feedback on *both*
 - what students know and can do
 - *and* how they got there, what helps or hinders
- Motivates students to become more active, reflective learners
- Can also be used by faculty collectively for the bigger picture
- Is faculty-friendly, respectful of privacy, autonomy
- Offers significant resources (e.g., T. Angelo and K. P. Cross, *Classroom Assessment Techniques*, 1992) and support networks, especially for community college educators

Disadvantages:

- Is unstructured, highly dependent on individuals' cooperation for
 - administration of CATs (classroom assessment techniques)
 - reporting of results
- Presents challenge of generalizing to program or institution level

Solutions/responses:

- Provide consistent, careful leadership, oversight
- Get buy-in from faculty, others
- Start with agreement on shared outcomes, goals
- Provide training
- Make assessment a campus-wide conversation
- Emphasize the potential for truly useful information for improvement

Student self-assessment: The student demonstrates the ability to accurately self-assess a piece of work or performance, usually in relation to one or more outcomes and a set of criteria, e.g. rubrics

Advantages:

- The ultimate in active learning, engagement, ownership of one's learning
- Highly adaptable
- Extremely educational for students
- Promotes internalization of intellectual, personal, professional standards
- Is an essential component of ongoing professional, personal development
- Is an essential component of life-long learning
- Faculty can aggregate individual results to identify general findings, trends

Disadvantages:

- Challenging, especially at outset, for both students and faculty
- Requires clear outcomes, criteria (e.g., rubrics), expectations for level of proficiency
- Requires student to assess with candor, not spin
- May cause anxiety, avoidance
- Long-standing habits, personality traits may need to be overcome (e.g., self-consciousness, excessive modesty, unrealistically high self-appraisal)
- Requires tact and true coaching attitude from instructor, ability to critique the work or performance, not the person
- Requires careful management of others who may be present

Solutions/responses:

- Experienced instructors guide, mentor novice instructors
- Students receive orientation, training
- Outcomes, criteria, expectations are clear, widely distributed and understood
- Examples of self-assessment are available
- Process is presented as primarily developmental, formative
- Examples of progress over extended time provide encouragement
- Self-assessment is risk-free

Local tests: tests designed in relation to the specific course, program, or institution's curriculum and learning outcomes, as opposed to generic, commercially available tests. Can be cumulative (e.g. comprehensives in the major) or less encompassing but still cross-cutting. Format may vary; need not be multiple choice, as in most commercial tests.

Advantages:

- Tests are traditional, widely accepted academic practice
- Testing across courses or programs requires active faculty participation
- Can stimulate discussion about alignment of goals, curriculum, pedagogy, etc.
- Can be designed to have content validity
- Can adapt readily to institutional changes in curriculum, outcomes
- Can be open-ended, integrative, highly creative in format
- Can provide good quality of student effort if course-embedded
- Provide directly relevant, *useful* information
- Forestall comparison with other institutions

Disadvantages:

- Run risk of focusing more on surface than deep learning
- Provide no norms for reference
- May contain ambiguous, poorly constructed items
- May offer questionable reliability and validity
- May be expensive if test construction is contracted out
- Will not elicit good quality of student effort if seen as add-on
- Will create misunderstanding of assessment if seen as a threat
- May become a missed opportunity to use more innovative approaches
- May invite finger-pointing

Solutions/responses:

- If norms, benchmarks are important, supplement with purchased test
- Use on-campus expertise
- Be careful, pilot any test before large-scale administration
- Provide a "gripe sheet"
- Accept that assessment is ultimately human judgment, not psychometric science
- Keep the focus on useful information & improvement, *not* test scores *per se*
- Depersonalize issues, avoid finger-pointing

Commercially available, standardized tests:

Advantages:

- Are a traditional, widely recognized & accepted means of assessment
- Require little on-campus time or labor
- Prepare students for licensure, other high-stakes testing
- Are norm-referenced
- Offer longitudinal data, benchmarks
- Are technically high-quality
- May reflect recent, important trends in the field (e.g., ETS Major Field Tests)
- Can be useful as *part* of a multiple-method approach

Disadvantages:

- May offer poor content validity
- Generally do not provide criterion-referenced scores
- Test students' ability to recognize "right" answers
- Reflect students' test-taking ability
- Often elicit poor quality of student effort, particularly as add-on
- Reinforce faculty bias toward "empty vessel" theory of education
- Reinforce student bias toward education as memorizing, regurgitating "right" answers (i.e. "surface" rather than "deep" learning)
- Reinforce everybody's bias toward assessment as testing
- Carry risk of misuse of scores, invidious comparisons
- Provide little insight into students' problem-solving & thinking skills or ability to discriminate among "good" and "better" answers
- Offer no opportunity for test takers to construct their own answers verbally, numerically, graphically, or in other ways
- Give students no opportunity to demonstrate important affective traits, e.g., persistence, meticulousness, creativity, open-mindedness.
- Are less likely than local methods to stimulate productive discussion
- Are more likely to elicit finger-pointing, anxiety, resistance
- Can be very expensive (\$10-\$30/student, plus administration costs)
- Generally do not provide good value (i.e., useful information for cost)

Solutions/responses:

- Test samples of students, use matrix sampling
- Negotiate with test maker
- Supplement with other methods
- Use with caution

Direct or indirect? Some methods can work both ways . . .

Classroom research: Classroom research is included here as a direct method but it can function as either a direct or an indirect method. Of the dozens of classroom assessment techniques (or CATs) developed by Cross and Angelo, some demonstrate what students know and can do, while others elicit reflection, perceptions, and other forms of indirect evidence.

Course management programs: Course management programs make it possible for faculty to capture discussions and other evidence that would be ephemeral in the classroom; hence they are classified here as a direct method. Such programs often include a survey or questionnaire template, however, that makes it easy to construct and administer surveys online. See discussion of surveys in handout on "Indirect Methods."

Focus groups: Focus groups are generally regarded as an indirect method of assessment because students are encouraged to talk about their personal experiences and perceptions. However, they can also function as a direct method, if the topic of discussion is an issue in the major and students are guided by the protocol to demonstrate their command of disciplinary concepts, theories and methods. In this case, students do not receive a grade for their role in the discussion, but the recording is analyzed by faculty to draw more general conclusions about the strengths and weaknesses of the academic program.

Portfolios: Portfolios can function as both a direct and an indirect assessment method. They are direct in the sense that student work is displayed and can be rated, providing direct evidence of knowledge and skills. The reflective essays, in which students look back on various pieces of their work, describe what each represented in terms of challenges or achievements, and evaluate their personal progress as learners, are indirect evidence of a high order.

Student self-assessment: Self-assessment is classified here as a direct method because the performance of self-assessment demonstrates directly how skilled students are at self-assessment. However, the process may be structured to elicit student reflection on how learning occurred, what helped or didn't, etc. In other words, self-assessment can also function as an indirect method.