CHAPTER 5: COLLECTING ASSESSMENT INFORMATION

There is no more critical juncture in implementing a successful assessment of the major than the moment of methods selection.

(Johnson, McCormick, Prus, & Rogers; 1993; p.153)

Topics Presented in Chapter 5

- ♦ Direct and indirect measures of student learning
- Description of direct assessment methods
- Description of indirect assessment methods
- ♦ Using existing student work
- ♦ Rubrics

Direct and Indirect Measures of Student Learning

It is important to distinguish between direct and indirect methods of collecting assessment information since units must use at least one direct measure. Suskie (2009) explains that direct methods provide demonstrations of what students know and can do that can be evaluated objectively.

Examples of Direct Measures of Student Learning

- Course-embedded assessments
- Ratings of student skills by their field experience supervisors
- Scores and pass rates on appropriate licensure or certification exams
- Capstone experiences, such as research projects, presentations, theses, dissertations, oral defenses, exhibitions, performances, scored using a rubric
- Other written work, performances, and presentations, scored using a rubric
- Portfolios of student work
- Scores on locally designed multiple-choice or essay tests such as final examinations in key courses, qualifying examinations, and comprehensive examinations
- Score gains (referred to as *value added*) between entry and exit on published or local tests or writing samples
- Observations of student behavior (such as presentations and group discussions), undertaken systematically and with notes recorded systematically
- Summaries and assessment of electronic class discussion threads
- Think-alouds, which ask students to think aloud as they work on a problem or assignment

- Classroom response systems (*clickers*) that allow students in their classroom seats to answer questions posed by the instructor instantly and provide an immediate picture of student understanding
- Feedback from computer-simulated tasks such as information on patterns of action, decisions, and branches
- Student reflections on their values, attitudes, and beliefs, if developing those are intended outcomes of the program

Indirect measures, on the other hand, are often used to collect information from students on what they believe they learned and how and why they learned it. (Suskie)

Examples of Indirect Measures of Student Learning

- Course grades and grade distributions
- Assignment grades, if not accompanied by a rubric or scoring criteria
- Retention and graduation rates
- Admission rates into graduate programs and graduation rates from those programs
- Scores on tests required for further study (such as the GRE) that evaluate skills learned over a lifetime
- Quality and reputation of graduate programs into which alumni are accepted
- Placement rates of graduates into appropriate career positions and starting salaries
- Alumni perceptions of their career responsibilities and satisfaction
- Student feedback of their knowledge and skills, and reflections on what they have learned over the course of their program
- Questions on end-of-course student evaluation forms that ask about the course rather than the instructor
- Student, alumni, and employer satisfaction with learning collected through surveys, exit interviews, or focus groups
- Student participation rates in faculty research, publications, and conference publications
- Honors, awards, and scholarships earned by students and alumni

The University of Virginia (n.d.) notes that:

Direct and indirect evidence can complement each other and when used in tandem can become more than the sum of the two. Indirect evidence may yield insights into students' experiences, ideas for assessment, or information that helps to interpret assessment results or guide application of results. Direct evidence can be brought to bear to test the validity of students' opinions or self-assessments. Student learning is sufficiently complex that multiple approaches may be needed.

The University of Central Florida (2008) provides the chart below that highlights use of both direct and indirect methods for assessment of several learning outcomes:

Use of Direct and Indirect Methods for Assessment of Learning Outcomes						
Outcomes	Graduating Senior Survey	Capstone Course	Portfolio	Focus Group		
Satisfaction with advising	Direct			Indirect		
Acquiring necessary skills and knowledge		Direct	Direct	Indirect		
Proficiency in written communication skills		Direct	Direct	Indirect		

Description of Direct Assessment Methods

The University of Central Florida also provides the following descriptions of assessment methods:

Capstone Course Assignments or Projects

Capstone course assignments or projects can be useful tools for program-level assessment. The assessment of important program learning outcomes can be integrated into a capstone course or project. Assessments structured into the capstone experience can include one or more of the following: exams, integrative papers or projects, research projects, reflective essays, oral reports, surveys, and focus groups. Capstone courses or projects are typically discipline-based and may be designated as a *senior seminar* or an *assessment course*. Graduates from a program demonstrate their competence in several areas and their ability to synthesize learning in the major with a product or performance. Projects are generally judged by a panel using prespecified scoring rubrics for the purpose of identifying where to improve the program.

Example: A panel of faculty members acts as evaluators of performances by music students, theatre students, etc., using a rubric that focuses on the important performance criteria and the quality of each. This method of assessment provides the student a chance to demonstrate the ability of absorbing and integrating their experiences and knowledge.

Advantages

- When capstone courses or projects are required, they can provide an ideal data collection opportunity because seniors are accessible.
- Assessments can provide an opportunity to motivate students through the curriculum. Also they can provide quality data that permit meaningful reflection on the program.
- Seniors are well into the curriculum and can reflect on their learning experience and the curriculum.
- These assessment methods provide seniors with an opportunity to provide meaningful feedback when they believe that their opinions are respected and valued.

- Students get feedback on their accomplishments, and student responsibility is encouraged.
- They can be used for both student evaluation (assess seniors' overall ability and knowledge gained from the program) and program evaluation (annual, continuous evaluation of curriculum from student feedback).
- They support program coherence.
- They provide an opportunity to create local assessment instruments that can be used in conjunction with other methods, such as surveys and standardized tests.
- Many faculty members are engaged in planning the topics and the design of the capstone experience.
- This assessment allows flexible course content (i.e., adaptable to different courses).

Disadvantages

- Capstone surveys could yield invalid or misleading feedback, particularly when responses are not anonymous.
- Student performance may be impaired due to "high stakes" of the project.
- A faculty member may develop the idea that the capstone course or project should only involve him or her.
- Successfully completing the capstone course may be a requirement for graduation which may generate some anxiety for both faculty and students.

Considerations

- Ensure that the course assignments or projects accurately represent the major or program requirements.
- Use checkpoints to prevent difficulties, especially towards the end, which may affect a student's graduation.
- Maintain the curriculum and evaluation of assignments across all sections.
- Ensure that students understand and value the importance of the capstone experience and take it seriously.
- Secure administrative support before implementing a capstone experience since there are usually high costs associated with it because of the small class size required to maximize the faculty-student interaction.
- Design capstone course or project to assess curriculum goals and outcomes.

Case Studies, Simulations, and Hypothetical Situations

A case study is a focused, systematic examination of one instance of a phenomenon such as an event, program, process, or person. Typically, case studies involve collection of qualitative and quantitative data such as observations, surveys, and interviews for an in-depth study of the phenomenon. Students can conduct case studies and/or respond to hypothetical situations.

Advantages

- Can be used to assess student work of both a quantitative and qualitative nature
- Are useful when a student learning outcome is to comprehensively study and understand a phenomenon of particular interest to the field
- Provide an opportunity for students to apply learned skills in context

Disadvantages

• Tend to be expensive, labor-intensive, and time-consuming, which can be prohibitive within a course

Considerations

- Single or multiple cases (collective case study) may be investigated.
- Different approaches may be used such as a highly structured approach or an unstructured process.

Content and Embedded Assessment Approaches

Course-Embedded Questions and Assignments

Course-embedded questions are predetermined questions that measure student learning in specific areas and can be used to assess students' knowledge, skills, behavior, and attitudes within a scheduled test. The test is typically a locally developed test. Often instructors of a particular course use the same questions within their unique course tests at a particular point in the course (e.g., midterm or final). Growth in discipline-specific knowledge, skills, or attitudes may be gauged using the same set of embedded questions in tests for different courses throughout the curriculum.

Portfolio Assignments

A portfolio is a collection of samples of student work. The contents can vary widely, from a collection of photographs, to written assignments, to a collection of computer programs. Sometimes an electronic portfolio is used to facilitate storage and access of the samples of student work. A rubric may be used to evaluate a collection of students' work (e.g., writing, homework, etc.) over a period of time. This method of assessment can provide longitudinal data to gauge growth of particular skills or understandings, as well as an opportunity for student reflection. Typically, each assignment included in a portfolio has been reviewed and graded. A committee or a designated group of faculty members may review portfolios in a program for the purpose of identifying where improvements in the program are needed.

Assessment of Papers, Projects With Standard Scoring Rubrics

A rubric is an assessment tool that can be used to specify scoring criteria for a paper, project, performance, or other method of assessment. Usually all of the key elements of an assignment and their weighting on the total score are identified. A rubric is most effective when it is shared with students prior to the start of an assessment assignment. For more information on developing rubrics, see http://rubistar.4teachers.org/index.php.

Research Paper

This is an assessment method which can be used to evaluate students' abilities to analyze, synthesize, and/or evaluate information that has been taught. A scoring rubric makes evaluation criteria clear when assessing research papers. On the program assessment level it could be part of a capstone project or a tool used in the senior year to determine if students have achieved programmatic learning outcomes.

Essays

Essays may be designed to measure specific learning outcomes (e.g., writing skills, appreciation for art, appreciation of diversity, etc.). These essays are scored using rubrics established by a panel of faculty. The rubrics may be reviewed for the purpose of identifying elements needing more emphasis in the academic program.

Direct Observation by Instructor, Expert Evaluators

A panel of individuals or an expert (e.g., supervisor) can score student performance in practice (e.g., music, communications, clinical). The panel may include members of the faculty, advisory board members, experts in the field, etc. Scoring rubrics are often used to improve inter-rater reliability.

Direct Observation by Peer

In-class exercises can be assessed by peers or peer panels using scoring rubrics. This provides a first-hand familiarity of criteria on the actual rubrics that will be used to assess future work. Students can use rubrics to assess examples of work not produced by classmates (e.g., use of a rubric to assess a videotaped speech for specific elements of speech).

Examinations and Tests

Standardized Examinations and Tests

National Test

Exams available nationally with standardized scores and sub-scores can be used to determine where to improve the program

Examples: Educational Testing Service (ETS) Field Exams, Psychology Area Concentration Achievement Test (PACAT), The Chauncey Group DANTE (Statistics Exam)

State Test

Exams available within the State of Florida with standardized scores and sub-scores that can be used to determine where to improve the program

Examples: Florida Certification Exam in School Psychology, Florida Educational Leadership Exam, Florida Teacher Certification Exam

Local Examinations and Tests

Local Tests

Exams are designed by members of an academic program or administrative program to measure student achievement of specific learning outcomes. The assessment purpose of these tests may be to identify where improvement is needed within the academic or administrative program.

Pre-Post Test

These are a type of locally developed test administered before and after a specified learning experience to measure students' level of knowledge, skills, behaviors, and attitudes. (The learning experience can be a program, course, or unit.) Post-test scores are compared to pre-test scores to determine if the students have learned specific information or concepts.

Certification and Licensure Exam

Certain disciplines (especially in health related disciplines) require that students pass specified certifications and licensure exams. Students' performance on these exams and their sub-scores, when available, are a source of data that can be used to assess student learning.

Description of Indirect Assessment Methods

Surveys

Institutional Level

This category includes locally and nationally developed surveys that focus on evaluating satisfaction with academic programs and service experience, perceived learning outcomes, plans for further education and employment, further education and/or employment placement, and plans of undergraduate and graduating undergraduate and graduate students.

Advantages

• Surveys can be an important tool in understanding student's academic needs and their perception of their educational experience. Additionally, surveys can be used to determine students' satisfaction with the services offered at the university as well as program-specific services such as advising, etc.

Disadvantages

• Surveys are used to gather data regarding the perceptions of individuals about personal experiences. In most instances, this method does not provide direct evidence of knowledge, skills, and abilities. When this method of assessment is implemented, a direct measurement approach should be used as well.

Considerations

- Careful planning for developing and administering institutional level surveys is critical for success. All stakeholders should be included.
- Institutional level surveys have budget implications that should be carefully considered.

Other Indirect Assessment Methods

Focus Group

Individuals who are users of the program or who benefit from the academic preparation made possible as a result of completing the program (e.g., employers, alumni, faculty, parents, etc.) can provide important qualitative data that can be used to identify strengths and weaknesses within the program.

Advisory Committee

Individuals who are experts in the field can assess student preparedness and curriculum content. This method of assessment provides a current and relevant level of analysis which is beneficial to the development of the curriculum as well as the assessment of students' knowledge, skills, and attitudes.

Structured Interview

One-on-one structured interviews with students, faculty, employers, and alumni conducted by a trained interviewer can provide useful information. This information can be used to identify strengths and weaknesses within the program.

Student Activity and Study Log

A log that reflects the amount of time a student spends studying or involved in specific activities can provide important data that can be used to identify opportunities for improvement. This can be managed electronically in a spreadsheet by individuals and combined into a group for assessment purposes.

Institutional Data

Institutional level data such as retention rates, graduation rates, demographics, time-to-graduation, and enrollment in graduate level programs by former graduates can provide useful information regarding the strengths and weaknesses of a program.

Using Existing Student Work

Walvoord (2010) is a strong proponent of using samples of existing student work along with rubrics for assessment.

Advantages

- Information is already available.
- There are no student motivation problems, since students must complete the work for a grade.
- There is no direct cost.
- It reflects what faculty members actually teach, not what's included on standardized tests, so faculty members are more motivated.

Disadvantages

- Evidence is not comparable across institutions.
- Everyone evaluates differently, so common standards or rubrics and training are needed.
- Information is in multiple parts and multiple formats, so it needs to be collected in ways that permit easy access.
- There is quite a bit of work, especially at the beginning.

Rubrics

A rubric is a scoring tool that lays out the specific expectations for an assignment. Rubrics divide an assignment into its component parts and provide a detailed description of what constitutes acceptable or unacceptable levels of performance for each of those parts. Rubrics are composed of four basic parts (University of Connecticut, n.d.):

- A task description (the assignment)
- A scale of some sort (levels of achievement, possibly in the form of grades) (Scales typically range from 3 to 5 levels.)
- The dimensions of the assignment (a breakdown of the skills/knowledge involved in the assignment)
- Descriptions of what constitutes each level of performance (specific feedback)

The University of Connecticut lists the following benefits of using rubrics.

- Rubrics provide timely feedback grading can be done more quickly.

 Since students often make similar mistakes on assignments, incorporating predictable notes into the *descriptions of dimensions* portion of a rubric can simplify grading into circling or checking off all comments that apply to each specific student.
- Rubrics prepare students to use detailed feedback.
 - In the rubric, the highest level descriptions of the dimensions are the highest level of achievement possible; whereas the remaining levels, circled or checked off, are typed versions of the notes/comments an instructor regularly writes on student work explaining how and where the student failed to meet that highest level. Thus, in using a rubric the student obtains details on how and where the assignment did or did not achieve its goal, and even suggestions (in the form of the higher level descriptions) as to how it might have been done better.
- Rubrics encourage critical thinking.

 Because of the rubric format, students may notice for themselves the patterns of recurring problems or ongoing improvement in their work.
- Rubrics facilitate communication with others.
 - TAs, counselors/tutors, colleagues, etc. can benefit from the information contained in the rubric (i.e., provides information to help all involved in a student's learning process).
- Rubrics help faculty refine their teaching skills.
 Rubrics showing a student's continuing improvement or weaknesses over time, or rubrics showing student development over time, can provide a clearer view of teaching blind spots, omissions, and strengths.
- Rubrics help level the playing field.
 - To aid first-generation or non-native speakers of English, rubrics can act as a translation device to help students understand what teachers are talking about.

Sample of a Rubric for a Slide Presentation on Findings From Research Sources (Suskie)						
	(5) Well done	(4-3) Satisfactory	(2-1) Needs improvement	(0) Incomplete		
Organization	Clearly, concisely written. Logical, intuitive progression of ideas and supporting information. Clear and direct cues to all information.	Logical progression of ideas and supporting information. Most cues to information are clear and direct.	Vague in conveying viewpoint and purpose. Some logical progression of ideas and supporting information but cues are confusing or flawed.	Lacks a clear point of view and logical sequence of information. Cues to information are not evident.		
Introduction	Presents overall topic. Draws in audience with compelling questions or by relating audience's interests or goals.	Clear, coherent, and related to topic.	Some structure but does not create a sense of what follows. May be overly detailed or incomplete. Somewhat appealing.	Does not orient audience to what will follow.		

The University of Virginia offers the following guidelines on developing rubrics:

- Clearly define the assignment including the topic, the process that students will work through, and the product they are expected to produce.
- Brainstorm a list of what you expect to see in the student work that demonstrates the particular learning outcome(s) you are assessing.
- Keep the list manageable (3-8 items) and focus on the most important abilities, knowledge, or attitudes expected.
- Edit the list so that each component is specific and concrete (for instance, what do you mean by coherence), use action verbs when possible and descriptive, meaningful adjectives (e.g., not *adequate* or *appropriate* but *correctly* or *carefully*).
- Establish clear and detailed standards for performance for each component. Avoid relying on comparative language when distinguishing among performance levels. For instance, do not define the highest level as *thorough* and the medium level as *less thorough*. Find descriptors that are unique to each level.
- Develop a scoring scale.
- Test the rubric with more than one rater by scoring a small sample of student work. Are your expectations too high or too low? Are some items difficult to rate and in need of revision?

And the University of Virginia offers the following advice on using rubrics:

- Evaluators should meet together for a training/norming session.
- A sample of student work should be examined and scored.
- More than one faculty member should score the student work. Check to see if raters are applying the standards consistently.
- If two faculty members disagree significantly (e.g., more than 1 point on a 4-point scale), a third person should score the work.
- If frequent disagreements arise about a particular item, the item may need to be refined or removed.

A number of rubric libraries are available on institutional web sites. Some of these include:

California State University Fresno http://www.csufresno.edu/oie/assessment/rubric.shtm
University of Delaware http://assessment.udel.edu/resources/rubrics.html
University of Virginia http://www.web.virginia.edu/iaas/assess/tools/rubrics.shtm