

Assessment Methods

Assessment methods are how evidence of student learning is collected. Learning data can come from a variety of sources—exams, student artifacts, performances, projects, reflections, surveys, focus groups, and interviews are some common examples. Before implementing measures to assess program learning outcomes, there are several considerations. First, a program should consider the questions it wishes to answer with the collected evidence. Examples of these questions include:

- How effective is the program curriculum at accomplishing what it sets out to do?
- Does the program meet the standards established internally or by an external entity?
- Can the learning experiences in the program be improved and where do these improvements need to be focused?
- How does the program compare to others in terms of the learning it is producing?

These examples are not exhaustive of the questions assessment methods may attempt to address, but these are generalized examples of key questions programs tend to be interested in answering with assessment data.

Diagnostic, Formative, and Summative Assessments

To these ends, a fundamental consideration is the function of a particular assessment. An assessment instrument can serve one of three purposes:

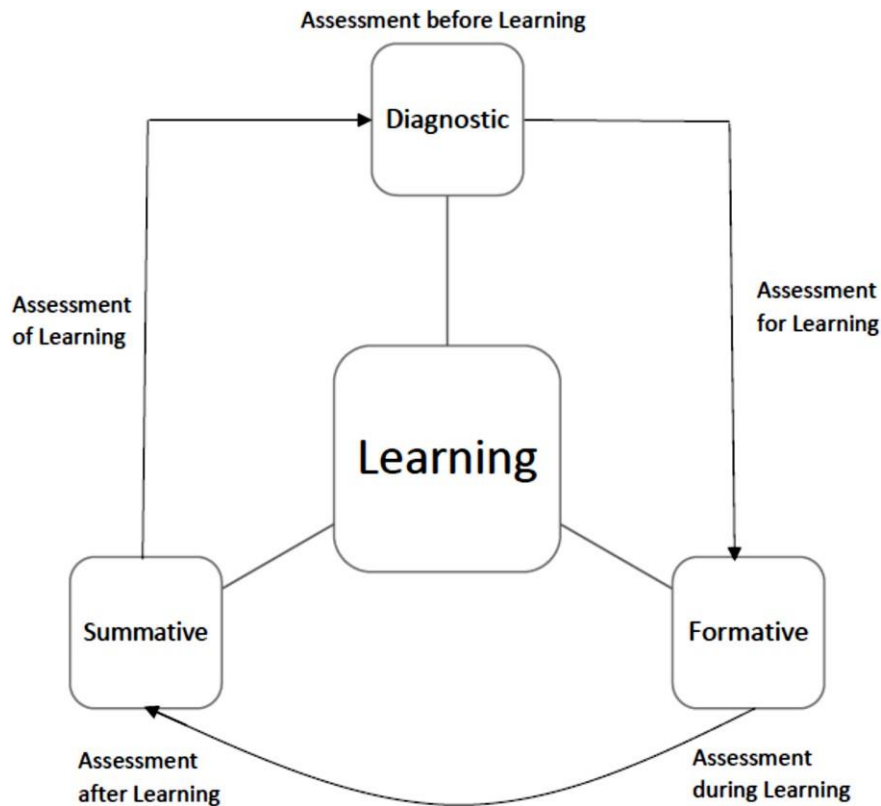
1. **Diagnostic Assessment**—These assessments attempt to ascertain what students know and are able to do upon entering a course or program of study. These assessments provide information concerning the initial state of student learning and can provide insight into student motivation for learning and the knowledge and skill strengths, weaknesses, and gaps that may exist among the students.
 - Examples of diagnostic assessments may include: a low-stakes quiz or diagnostic exam testing students' knowledge upon entering a course; a survey given on the first day of class to establish students' knowledge, interests, and perception of skills; questions from a final exam in a preceding, required course given again to establish the learning students are bringing with them into a subsequent, higher-level course.
2. **Formative Assessment**—Formative assessment is used to monitor the ongoing progress of student learning and to provide feedback to instructors and students concerning the progress of the learning. This information can be used to adjust instructional strategies, content focus, or modify curricular structures to better address learning needs. Provided with formative assessment feedback, students can identify where they have strengths and weaknesses and allows them to focus their efforts in these areas before a summative assessment. For instructors, formative assessment data helps to target interventions, instructional strategies, and content intensity in areas that may require focus.
 - Examples of formative assessments include: weekly quizzes given to students to establish their ongoing progress; chapter exams that test students' knowledge from which final exam questions may be compiled; a midterm survey of students asking them to identify areas of learning that remain problematic or unclear to them; a series of performance-

based assessments in a lab course that provide students to practice their skills before undertaking a summative assessment of those skills.

3. **Summative Assessment**—Summative assessments evaluate student learning at the end of a course or program of study. These assessments are interested in students’ mastery of the learning and are intended to represent the cumulative learning in the course or program and seek to answer the question, “How well have students achieved the programmatic or institutional expectations for learning?”

- Examples of summative assessments may include: a comprehensive final exam in a course; a high-stakes exit written or oral exam given at the end of a program of study; an end-of-course survey or reflection given to students that asks them to identify or rate their own proficiency in the key course objectives; a final skills-based project in a course requiring students to apply all that they have learned to successfully complete the assigned task(s).

When effectively applied, these forms of assessment operate in an integrated cycle where ongoing assessment informs the complete learning process:



Direct versus Indirect Measures

An assessment method tends to be classified as either a direct or indirect measure. There is no such thing as a perfectly direct measure of learning; all measurement involves some level of error. The directness of any measure of learning is predicated on its ability to facilitate the direct demonstration of what students have learned. Conversely, an indirect measure asks students to relate their perception of learning or is proxy data suggestive of learning without direct demonstration. To reiterate:

Direct Measures—Performance tasks requiring the demonstration of learning, behaviors, skills, or values

Indirect Measures—Feedback or perceptions about learning, behaviors, skills, or values or proxy data suggestive of learning having occurred

As a rule, the use of more direct measures is preferable to indirect measures. However, good practice suggests that a mix of direct and indirect measures should be utilized to assess student learning. Direct measures provide tangible, visible data that provides evidence of what students have and have not learned as a product of a course of study whereas indirect measures consist of proxy data and learner perceptions that suggest whether learning has or has not occurred.⁵ Indirect measures may complement the use of direct measures by adding context, nuance, and detail to the results of direct measures.

Direct and Indirect Assessment Definitions and Examples

	Direct	Indirect
Definition	<p>An explicit representation or demonstration of learning via a performance, process, or product</p> <p>Addresses such questions as: What did students learn as a result of one or more educational experiences? To what degree did students learn? Are students meeting our goals for learning?</p>	<p>Students' perceptions of their own learning or proxy data suggestive of learning having occurred</p> <p>Addresses such questions as: Where do students report struggling with the learning? What were the student perceptions of the learning task and the resultant learning? How might the direct methods be modified to improve learning?</p>
Examples	<ul style="list-style-type: none"> • Ratings of student skills by trained observers • Scores or pass rates on licensure or certification examinations • Capstone assessments such as research projects, presentations, theses, dissertations, or performances scored with a rubric • Portfolios of student work • Course-embedded assignments and tests • Standardized tests 	<ul style="list-style-type: none"> • Course grades and grade distributions • Admission rates into higher-level study after program completion • Placement rates of graduates • Alumni perceptions • Exit surveys • Focus group interviews • Employer surveys

⁵ Maki, P. (2004). Assessing for learning: Building a sustainable commitment across the institution. Stylus.

Quantitative and Qualitative Methods

Quantitative assessments tend to be highly structured with predetermined correct answers that may be framed as response options, data entry, or required illustrations that can be evaluated and summarized into numeric data that can then be described and analyzed statistically. Scores on tests, rubrics, surveys, checklists, or other performance indicators are examples of quantitative data, and these assessments are often objective in nature with one correct answer requiring no professional judgement to score (although professional judgement is required to interpret the scores), such as multiple choice, matching, or true-false questions. Quantitative methods can also be subjective and require professional judgement to assign a score, for example student skill demonstrations or performances requiring a rubric score.

Qualitative assessments tend to be more flexible and open-ended and often require extended responses. The analysis of qualitative data is focused on establishing trends and themes in the data that can be summarized descriptively using words. Qualitative assessments are underutilized and are often underappreciated for their ability to explore possibilities and provide insights to discover problems and solutions that cannot be found through quantitative methods alone.¹ Despite their tendency to require extended, open-ended responses, qualitative assessments are structured and systematic and evaluate students using common criteria. These qualities differentiate qualitative assessments from informal or anecdotal observations that are unstructured or lacking common criteria.¹

Combining quantitative and qualitative measures is an effective way to ensure assessment data provides both depth and breadth to the interpretation of student learning. Remember that quantitative methods assign interpretive value on numerical data and tend to assess foundational level knowledge and skills whereas qualitative methods place value on the observer and allow for the assessment of broader dimensions of learning.⁵

Reliability of Assessment Methods

Reliability refers to the consistency of a measure. Simply put, reliability is the ability of an instrument to obtain fair and consistent results over repeated applications. Four types of reliability may be relevant to the application of an assessment method.

Types of Reliability

Reliability Type	Concern Addressed
Internal Consistency	Are the instrument's items understood similarly?
Inter-rater Reliability	Do the raters score the work product or demonstration of learning consistently?
Test-Retest Reliability	How stable are the measures over time relative to the target construct?
Parallel Forms Reliability	Are different forms of the instrument reliable or equivalent?

Varied statistical techniques can be used to test these forms of reliability. Establishing the reliability of an assessment method is necessary but not sufficient for the establishment of validity. For more information concerning testing the reliability of assessment instruments, contact the Director of Assessment.

Validity of Assessment Methods

A thorough discussion of the concerns of validity would require significant space be devoted to this complex topic. In brief, validity in this context can be defined simply as whether an assessment instrument measures the learning construct it purports to measure. A more encompassing definition is concerned with the defensibility of inferences derived from data gathered using a particular instrument.⁶ From a practical perspective, validity is concerned with whether an assessment method can measure a learning outcome accurately. A valid assessment method is reliable and allows for the accurate evaluation of learning as articulated in one or more outcome statements⁵. Validity can be concerned with several aspects of an assessment method, depending on the use of the resulting data.

Types of Validity

Validity Type	Concern Addressed
Content Validity	Do experts agree that the items or elements are appropriate and comprehensive?
Concurrent Validity	Does the instrument effectively group people into appropriate groups?
Predictive Validity	Does the instrument effectively identify a future state or outcome?
Convergent & Discriminant Validity	Does the instrument measure the learning construct and nothing else?
Cultural Validity	Does the instrument measure the same thing across sub-populations?
Incremental Validity	Does the instrument add important information beyond other sources of data?
Face Validity	Does the instrument appear reasonable?
Construct Validity	Is the instrument appropriate and specific enough to evaluate the intended outcome and population?

The concerns of validity are relevant to the varied uses of assessment data and can be established using expert feedback and statistical techniques. In many cases construct and content related validity are the chief concerns of educators in the classroom and tend to provide an assessment instrument with face validity. Disaggregating the resulting data by student sub-groups is necessary to establish cultural validity, and this is good practice to ensure that educational experiences are producing equitable learning for all students. For more information concerning establishing the validity of assessment instruments, contact the Director of Assessment.

Aligning Methods with Outcomes

To possess construct validity an assessment method must provide accurate information related to the learning outcome it purports to measure. An assessment method should have a close correspondence with the operant verb category in the learning outcome(s) it measures. For example, in the prior

⁶ Fraenkel, J., Norman, W. & Hyun, H. (2015). How to design and evaluate research in education. 9th ed. McGraw Hill.

example of the learning outcome “Students will design an experiment to test a hypothesis” the verb *design* in the outcome was determined to require students to demonstrate a “creation” level of learning in the cognitive domain and activate their procedural knowledge. A valid assessment of this outcome would require students to design an experiment and could take the form of a project, research paper, or laboratory demonstration evaluated using a rubric or other criteria-based instrument.

As a summative assessment of the outcome, multiple-choice exam questions focused on experimental design would not be likely to have strong construct validity as this assessment method fails to assign an appropriate and specific task for evaluating students’ ability to design an experiment. Rather, this assessment method would be more appropriate for evaluating students’ understanding of and ability to apply and evaluate experimental design elements. While these concerns are necessary for the creation of an experimental design, they are not sufficient to comprehensively evaluate the *creation* of the procedural task in this example.

Refer to the program’s curriculum map to identify where in the curriculum outcomes are addressed. Identify the level of expectation for the PLO in the venue where it is being assessed. A progression map will help to identify these expectations across courses and illustrate how the learning is scaffolded across the curriculum. Through articulation of the expectation for the learning for the assessed PLO, a valid assessment method may be identified.

The following table provides examples of assessment methods for the levels of cognitive learning in Bloom’s Taxonomy.

Example Outcome Verb (Bloom's Taxonomy)	Example Methods	Example Measures
Remember* Students will be able to: <ul style="list-style-type: none"> • Define • Identify • List • Select • Recall • Recognize 	<ul style="list-style-type: none"> • Objective Test items that require students to recall or recognize information: <ul style="list-style-type: none"> ○ Fill-in the Blank ○ Multiple Choice items with question stems such as, “what is a . . .”, or “which of the following is the definition of . . .”) ○ Labeling diagrams • Reciting (orally, musically, or in writing) 	Accuracy – correct vs number of errors Item Analysis (at the class level, are there items that had higher error rates? Did some items result in the same errors?)
Understand Students will be able to: <ul style="list-style-type: none"> • Interpret • Exemplify • Classify • Summarize • Infer • Compare • Explain 	Papers, oral/written exam questions, problems, class discussions, concept maps, homework assignments that require (oral or written): <ul style="list-style-type: none"> • Summarizing readings, films, speeches, etc. • Comparing and/or contrasting two or more theories, events, processes, etc. • Classifying or categorizing cases, elements, events, etc., using established criteria 	Scoring or performance rubrics that identify critical components of the work and discriminates between differing levels of proficiency in addressing the components

<ul style="list-style-type: none"> • Classify • Categorize • Restate • Describe • Paraphrase 	<ul style="list-style-type: none"> • Paraphrasing documents or speeches • Providing examples or illustrations of a concept, principle 	
<p>Apply Students will be able to:</p> <ul style="list-style-type: none"> • Execute • Implement • Solve • Demonstrate • Organize • Interpret 	<p>Activities that require students to use information and procedures to solve or complete familiar or unfamiliar tasks; may also require students to determine which procedure(s) are most appropriate for a given task. Activities may include:</p> <ul style="list-style-type: none"> • Problem sets • Performances • Labs • Prototyping • Simulations 	<p>Accuracy Scores; Check lists; Rubric Scores</p>
<p>Analyze Students will be able to:</p> <ul style="list-style-type: none"> • Differentiate • Organize • Attribute • Categorize • Order • Calculate 	<p>Activities that require students to discriminate or select relevant from irrelevant parts, determine how elements function together, or determine bias, values, or underlying intent in presented materials. These may include:</p> <ul style="list-style-type: none"> • Case studies • Critiques • Labs • Papers • Projects • Debates 	<p>Rubric Scores; Instructor Observations or Juried Assessments—external clients, employers, internship supervisor, etc.</p>
<p>Evaluate Students will be able to:</p> <ul style="list-style-type: none"> • Judge • Critique • Support • Appraise • Defend • Estimate • Measure • Predict • Argue 	<p>A range of activities that require students to test, monitor, judge or critique readings, performances, processes, or products against established criteria or standards. These activities may include:</p> <ul style="list-style-type: none"> • Journals • Diaries • Critiques • Problem Sets • Product Reviews • Case Studies 	<p>Rubric Scores; Instructor Observations or Juried Assessments—external clients, employers, internship supervisor, etc.</p>
<p>Create Students will be able to:</p> <ul style="list-style-type: none"> • Generate • Plan • Produce 	<p>Constructive activities that require students to formulate, design, plan, produce, or construct a product, process, or performance based on their knowledge and cognitive abilities. These activities may include:</p>	<p>Rubric Scores; Instructor Observations or Juried Assessments—external clients, employers, internship supervisor, etc.</p>

<ul style="list-style-type: none"> • Construct • Write • Assemble • Invent 	<ul style="list-style-type: none"> • Research projects • Musical compositions • Performances • Essays • Business Plans • Designs—websites, sets, plans, project outlines 	
--	--	--

* "Remember" as an outcome is a level of learning expected in introductory and intermediate levels of a program; therefore, it is a suitable learning outcome for the level of a course, but not as suitable for program-level learning at the capstone level.

Characteristics of Effective Assessments

An effective assessment task engages student learning while evaluating it. These assessments often address real-world examples and enduring or emerging issues in the discipline. Consequently, the assessment task is often focused on an ill-defined problem whereby students must activate their factual, conceptual, procedural, and metacognitive knowledge and present a reasoned judgement, innovative solution, or insightful analysis. The following table lists eight characteristics of an effective assessment task.⁷

Characteristic	Description of Quality
Valid	Yields useful information to guide teaching and learning
Coherent	Structured to result in the desired performance or product
Authentic	Asks the student to “do”, is realistic, and requires judgement to apply relevant knowledge and skills
Rigorous	Emphasizes purposeful, sustained reasoning (including metacognition) rather than discreet, isolated skills
Engaging	Activates student interest and persistence in the task
Challenging	Activates and evaluates student learning
Respectful	Provides an opportunity for students to express their uniqueness as learners
Responsive	Provides timely feedback to students to improve learning

⁷ Huba, M., & Freed, J. (2000). Learner-centered assessment on college campuses: Shifting the focus from teaching to learning. Pearson.