What Is Assessment?

Some valuable ideas you’ll find in this chapter

- Assessment is simply deciding what we want students to learn and making sure they learn it.
- Assessment is a cousin of traditional empirical research.
- Assessment today is based on research on effective teaching strategies in higher education.
While the term assessment can be used broadly – we can assess the achievement of any goal or outcome – in this book, the term generally refers to the assessment of student learning. Many assessment practitioners have put forth definitions of student learning assessment, but the best one I’ve heard is in the Jargon Alert box. It’s from Dr. Jane Wolfson, a professor of biological sciences at Towson University (personal communication, n.d.). It suggests that student learning assessment has three fundamental traits.

1. We have evidence of how well our students are achieving our key learning goals.
2. The quality of that evidence is good enough that we can use it to inform important decisions, especially regarding helping students learn.
3. We use that evidence not only to assess the achievement of individual students but also to reflect on what we are doing and, if warranted, change what we’re doing.

Assessment is part of teaching and learning

Assessment is part of a four-step process of helping students learn (List 1.1). These four steps do not represent a one-and-done process but a continuous four-step cycle (Figure 1.1). In the fourth step, evidence of student learning is used to review and possibly revise approaches to the other three steps (see Jargon Alert on closing the loop), and the cycle begins anew.

List 1.1 The Four-Step Teaching-Learning-Assessment Process

1. Establish clear, observable expected goals for student learning
2. Ensure that students have sufficient opportunities to achieve those goals
3. Systematically gather, analyze, and interpret evidence of how well student learning meets those goals
4. Use the resulting information to understand and improve student learning

If the cycle in Figure 1.1 looks familiar to you, it’s the Plan-Do-Check-Act cycle of business quality improvement popularized by Deming (2000): Plan a process, do or carry out the process, check how well the process is working, and act on the information obtained during the Check step to decide on improvements to the process, as appropriate.
Comparing traditional and current approaches to assessment

Faculty have been assessing student learning for centuries, often through written and oral examinations. How do today’s approaches to assessment differ from traditional approaches? Table 1.1 summarizes some key differences between traditional and contemporary ways of thinking about assessment.

Table 1.1: Traditional Versus Contemporary Ways of Thinking About Assessment

<table>
<thead>
<tr>
<th>Traditional Approaches: Assessment is...</th>
<th>Contemporary Approaches: Assessment is...</th>
</tr>
</thead>
<tbody>
<tr>
<td>Planned and implemented without consideration of learning goals, if any even exist</td>
<td>Carefully aligned with learning goals: The most important things we want students to learn (Chapter 4)</td>
</tr>
<tr>
<td>Often focused on memorized knowledge</td>
<td>Focused on thinking and performance skills (Chapter 4)</td>
</tr>
<tr>
<td>Often poor quality, simply because faculty and staff have had few formal opportunities to learn how to design and use effective assessment strategies and tools</td>
<td>Developed from research and best practices on teaching and assessment methodologies (Chapters 3 and 26)</td>
</tr>
<tr>
<td>Used only to assess and grade individual students, with decisions about changes to curricula and pedagogies often based on hunches and anecdotes rather than solid evidence</td>
<td>Used to improve teaching, learning, and student success as well as to assign grades and otherwise assess individual students (Chapters 6 and 26)</td>
</tr>
<tr>
<td>Used only in individual course sections; not connected to anything else</td>
<td>Viewed as part of an integrated, collaborative learning experience (Chapter 2)</td>
</tr>
<tr>
<td>Not used to tell the story of our successes; stories are told through anecdotes about star students rather than broader evidence from representative students</td>
<td>Used to tell our story: What makes our college or program distinctive and how successful we are in meeting societal and student needs (Chapter 25)</td>
</tr>
</tbody>
</table>
Comparing assessment and grading

Obviously there is a great deal of overlap between the tasks of grading and assessment, as both aim to identify what students have learned. There are two key differences, however. The first is that the grading process is usually isolated, involving only an individual faculty member and an individual student. Assessment, in contrast, focuses on entire cohorts of students, and it often considers how effectively many people, not just an individual faculty member, are collectively helping them learn.

The second difference between grading and assessment is that they have different purposes. The main purpose of grades is to give feedback to individual students, while assessment has three broader purposes discussed in Chapter 6: Ensuring and improving educational quality, stewardship, and accountability. Grades alone are usually insufficient to achieve these purposes for several reasons.

Grades alone do not usually provide meaningful information on exactly what students have and haven’t learned. We can conclude from a grade of B in an organic chemistry course, for example, that the student has probably learned a good deal about organic chemistry. But that grade alone cannot tell us exactly what aspects of organic chemistry she has and has not mastered.

Grading and assessment criteria may differ. Some faculty base grades not only on evidence of what students have learned, such as tests, papers, presentations, and projects, but also on student behaviors that may or may not be related to course learning goals. Some faculty, for example, count class attendance toward a final course grade, even though students with poor attendance might nonetheless master course learning goals. Others count class participation toward the final grade, even though oral communication skills aren’t a course learning goal. Some faculty downgrade assignments that are turned in late. Under these grading practices, students who do not achieve major learning goals might nonetheless earn a fairly high grade by playing by the rules and fulfilling other less-important grading criteria. Conversely, students who achieve a course’s major learning goals might nonetheless earn a poor grade if they fail to do the other things expected of them. To better sync grading and assessment criteria, add a professionalism learning goal (Chapter 4) or develop a competency-based curriculum (Chapter 5).

Grading standards may be vague or inconsistent. While many faculty base assignment and course grades on carefully conceived learning goals and standards, others may base grades on inconsistent, imprecise, and idiosyncratic criteria. Faculty may say they want students to learn how to think critically, for example, but base grades largely on tests emphasizing factual recall. Faculty teaching sections
of the same course may not agree on common standards and may therefore award different grades to similar student performance. Sometimes individual grading standards are so vague that a faculty member might, in theory, award an A to a student’s work one day and a B to identical work a week later.

Grades do not reflect all learning experiences. Grades give us information on student performance in individual courses or course assignments (Association of American Colleges & Universities, 2002), but they do not provide information on how well students have learned key competencies, such as critical thinking or writing skills, over an entire program. Grades also do not tell us what students have learned from ungraded co-curricular experiences.

Do grades have a place in an assessment effort? Of course they do! Grades can be useful, albeit indirect (Chapter 3), and therefore insufficient evidence of student learning. Although grades are often too holistic to yield useful information on strengths and weaknesses in student learning, they can be a good starting point for identifying potential areas of concern. DFIW rates – the proportions of students earning a D, F, Incomplete, or Withdrawal in a course – can identify potential barriers to student success.

Grades can be especially useful if courses, assignments, and learning activities are purposefully designed to help students achieve key learning goals (Chapters 5 and 16) by using tools such as test blueprints (Chapter 17) or rubrics (Chapter 15).

Comparing assessment and scholarly research

Assessment, while a cousin of scholarly research, differs in its purpose and therefore in its nature (Upcraft & Schuh, 2002). Traditional scholarly research is commonly conducted to test theories, while assessment is a form of action research (see Jargon Alert) conducted to inform one’s own practice – a craft-based rather than scientific approach (Ewell, 2002). The four-step teaching-learning-assessment cycle of establishing learning goals, providing learning opportunities, assessing student learning, and using evidence of student learning mirrors the four steps of action research: Plan, act, observe, and reflect.

Assessment, like any other form of action research, is disciplined and systematic and uses many of the methodologies of traditional research. But most faculty and staff lack the time and resources to design and conduct rigorous, replicable empirical research.
studies of student learning. They instead aim to keep the benefits of assessment in proportion to the time and resources devoted to them (Chapter 12). If you design your assessments reasonably well and collect corroborating evidence (Chapter 21), your evidence of student learning may be imperfect but will nonetheless give you information that you will be able to use with confidence to make decisions about teaching and learning.

Comparing assessment and evaluation

Is assessment a synonym for evaluation? It depends on the definition of evaluation that is used.

Evaluation may be defined as using assessment information to make an informed judgment on matters such as whether students have achieved the learning goals we've established for them, the relative strengths and weaknesses of teaching strategies, or what changes in learning goals and teaching strategies are appropriate. Under this definition, evaluation is the last two steps of the teaching-learning-assessment process: Interpreting student learning evidence (part of Step 3) and using it (Step 4). This definition points out that student learning evidence alone only guides us; it does not dictate decisions to us. We use our best professional judgment to make appropriate decisions. This definition of evaluation thus reinforces the ownership that faculty and staff have over the assessment process.

Evaluation may be defined as determining the match between intended and actual outcomes. Under this definition, evaluation is virtually synonymous with the third step of the teaching-learning-assessment cycle.

Evaluation may be defined as investigating and judging the quality or worth of a program, project, or other endeavor. This defines evaluation more broadly than assessment. We might evaluate an employee safety program, an alumni program, or a civic project designed to reduce criminal recidivism. While assessment focuses on how well student learning goals are achieved, evaluation addresses how well all the major goals of a program are achieved. An anthropology program, for example, might have goals not only for student learning but also to conduct anthropological research, provide anthropological services to local museums, and conduct its affairs in a cost-effective manner. An evaluation of the program would consider not only student learning but also research activities, community service, and cost-effectiveness.
Comparing assessment and measurement

Just as assessment and evaluation of student learning are sometimes considered synonymous, so are assessment and *measurement* of student learning. But many people have a relatively narrow conception of measurement, thinking of it as placing something on a quantitative scale akin to a yardstick. This book avoids the term *measurement*, because assessment is much broader than this conception. Assessment may generate qualitative as well as quantitative evidence of student learning (Chapter 20); it may generate categorical evidence as well as evidence that can be placed on a scale (Chapter 23); and it does not have the precision that images like a yardstick imply (Chapter 24).

Time to think, discuss, and practice

1. Compare the traits of traditional and contemporary assessment practices in Table 1.1 with those you’re aware of at your college. Do people at your college largely practice traditional or contemporary approaches to assessment? Can you think of any ways to make contemporary approaches more pervasive?

2. If anyone in your group has already conducted an assessment of student learning, ask them to share what was done. Then discuss how, if at all, the assessment would have been done differently if it had been approached as scholarly research.