WORCESTER STATE UNIVERSITY

PROGRAM ASSESSMENT HANDBOOK

Sarah L. Strout, Ph.D.

Assistant Vice President

Office of Institutional Assessment and Planning

.

American Association for Higher Education's Principles of Good Practice for Assessing Student Learning

1. The assessment of student learning begins with educational values.

2. Assessment is most effective when it reflects an understanding of learning as multidimensional, integrated, and revealed in performance over time.

3. Assessment works best when the programs it seeks to improve have clear, explicitly stated purposes.

4. Assessment requires attention to outcomes but, also, and equally to the experiences that lead to those outcomes.

5. Assessment works best when it is ongoing, not episodic.

6. Assessment fosters wider improvement when representatives from across the educational community are involved.

7. Assessment makes a difference when it begins with issues of use and illuminates questions that people really care about.

8. Assessment is most likely to lead to improvement when it is part of a larger set of conditions that promote change.

9. Through assessment, educators meet responsibilities to students and to the public.

Introduction

This handbook was designed as a resource for Worcester State University programs to assist in program level assessment. This handbook addresses the basics of student learning assessment, the benefits of assessment, and concerns relating to assessment. This handbook was made specifically for WSU faculty, and thus should not be distributed or printed for external publication. I hope that you find this handbook helpful. For further assistance, The Office of Assessment and Planning is always happy to meet with faculty and programs. While this handbook focuses mainly on program level assessment, the Center for Teaching and Learning has a library of resources on course level and classroom assessment techniques. For more information, contact the Director of the Center for Teaching and Learning.

What is the Assessment of Student Learning?

Assessment has been a part of education for as long as there has been teaching. Assessment can refer to any activity that is designed to collect information on student learning. As part of the teaching process, faculty examine what improves student learning and what does not, and uses those observations to make changes to their courses, pedagogy, and programs. Formal assessment (as discussed in this handbook) simply makes those informal activities more systematic, focused, effective, and public.

Program assessment is an ongoing process to monitor and improve student learning within a program. Program assessment explores how the varied parts of the curriculum come together to result in improvements of student learning (Allen, 2004). Program assessment is not meant to evaluate individual students, faculty, courses or staff, rather to look at the skills and knowledge obtained by graduates of the program overall.

> "Assessment is the systematic collection, review, and use of information about educational programs undertaken for the purpose of improving student learning and development."
>
> Palomba & Banta, 1999

Why Assess Student Learning?

Improving Student Learning in the Program

The results from assessment can help programs identify specific content and skills that students are weak in, and target curriculum to those areas. Programs can also identify the point in the program when students learn skills and content. Assessment data can also allow faculty to identify discrepancies between what faculty want students to learn and what they actually learn (Allen, 2004).

Program assessment can help rationalize the design of new courses, or changes to perquisite courses and course sequencing. Program assessment can also justify resources needed to improve programs (Allen, 2004; Dunn, McCarthy, Baker, & Halonen, 2011).

Program assessment can also provide a means of communicating expectations. Students will be aware of what skills and knowledge the program values, and what they can expect to learn in the program. In addition, program learning outcomes can help programs develop a common language among full-time faculty, part-time faculty, and students (Allen, 2004).

> "If assessment doesn't help improve teaching and learning activities, why bother with it?"
>
> Suskie, 2004

Equity

By the year 2020, 45% of high-school graduates in the United States will be minority students. By 2050, the United States will become a majority-minority nation. It is projected that the number of white high school graduates will decline in the next 10-20 years and the number of high-school graduates of color will rise (Bransberger & Michelau, 2016).

> "Assessment, if not done with equity in mind, privileges and validates certain types of learning and evidence of learning over others, can hinder the validation of multiple means of demonstration, and can reinforce within students the false notion that they do not belong in higher education."
>
> Montenegro & Jankowski, 2017

As a public institution, Worcester State University will continue to see an increase in the diversity of students including, but not limited to, students of color, LGBTQ students, and first-generation students. It is important to note that students enter college with various degrees of familiarity with college practices. For first-generation students and students from low-performing high schools, it is especially important to be as transparent as possible in expectations of student-work, to give all students an equal chance to succeed.

To address this problem, assessment practitioners have developed guidelines for inclusive assessment, (also known as culturally responsive assessment. Culturally Responsive Assessment is student-centered assessment that takes into account the diversity of the student population the university serves rather than using a one-size-fits-all approach (Montenegro & Jankowski, 2017).

How to Create Inclusive and Culturally Responsive Assessments

- Poorly written and ambiguous student learning outcomes make it difficult for students to understand what is expected of them and thus, demonstrate their learning. SLOs should be written with the student in mind, being as clear and explicative as possible.
- Assessment approaches should allow for multiple ways for students to demonstrate mastery, rather than a one-size-fits-all approach.
- Clearly defined rubrics allow students to see exactly what is expected of them, and help instructors avoid bias when assessing student work.
- Instructions for student work should include as much information as possible, so that students do not need to 'read between the lines.' Many first-generation students and students from low-performing high-schools do not have experience in decoding what 'college-level' work is, so the more clearly an instructor defines what is expected of the student, the more likely a student is to succeed.

- Do not assume that students have entered your course or program with specific skills already developed. Writing, critical thinking, and information literacy skills (just to name a few) may not have been emphasized in the students' high-school education.
- Disaggregating assessment data when possible allows programs to see if particular demographic groups are struggling or succeeding in relation to other groups, and allows faculty to develop interventions to help struggling students.
- Avoid comparing students against peers. Having a standard judgement (for example, a rubric) allows instructors to compare students against a pre-determined standard.
- When students do not perform well on assignments, ask them why. It may be that the instructions were not as clear as the instructor thought, or the students did not understand a concept and need review.

Discipline-Specific Accreditation

Another reason for assessment is discipline-specific accreditation. It has become standard for discipline-specific accreditors to require assessment of student learning both for initial accreditation and for continuing accreditation (Dunn, McCarthy, Baker, & Halonen, 2011). Many programs will need to show evidence of a continuous assessment process before being accredited.

Accountability and Transparency

While the primary purpose of student learning assessment is to improve student learning, assessment also serves other purposes, including providing accountability and transparency to stakeholders, and meeting regional accreditation requirements.

Public higher education was built on the premise that access to a college education was both a public good for society and a private good for students. Taxpayers accepted the obligation to provide adequate operating funding to public colleges and universities while expecting they would keep tuition relatively low. Over time, as fiscal issues grew, demands for accountability grew as well (Baer, 2017). This has led to an expectation that the public will have access to information about the success of students attending public colleges and universities.

> "Like most compacts, the one between American society and higher education became strained when rights and responsibilities moved from vague generalities to specific demands and competed for funding without public services."
>
> Burke & Minassians, 2004

The Secretary of Education's Commission on the Future of Higher Education (2006) offered a strong indictment of American higher education. The commission focused on costs that were too high, graduation rates that were too low, especially among low income and minority students, and learning outcomes that remained a mystery. Added to these reports, for-profit institutions have been confronted by additional outcomes measures, including the rate at which students default on loans post-graduation and the relationship between the price of education and post-completion earnings (Gainful employment regulations). These measures are now beginning to be examined in not-for-profit institutions as well (Baer, 2017).

In terms of regional accreditation requirements, Worcester State University is accredited by The New England Commission of Higher Education (NECHE) formerly known as the Commission of Institutions of

Higher Education in the New England Association of Schools and Colleges. In the 2016 Self Study Guide the commission states:

> "Within the standard on *Educational Effectiveness,* assessment is a core component. Specific to assessment of student learning, institutions demonstrate that they meet the Commission's criteria for institutional effectiveness through a process that articulates expected student outcomes, gathers and interprets evidence regarding what and how students are learning, and uses these interpretations to enhance teaching and learning and to inform institutional decision-making in order continually to improve students' learning experiences and their outcomes."

Deficiencies in student learning outcomes assessment are the most common shortcoming in regional accreditors' evaluations of institutions. Accreditors are asking increasing numbers of institutions for follow-up actions or reports focused on these issues (Provenzis, 2010). Staff from the New England Association of Schools and Colleges (now NECHE) indicate that 80 percent of institutions in their region have been asked for follow-up actions related to assessment (Banta & Palumbo, 2015).

The following are NECHE Standards relating to assessment:

> 8.3 Assessment of learning is based on verifiable statements of what students are expected to gain, achieve, demonstrate, or know by the time they complete their academic program. The process of understanding what and how students are learning focuses on the course, competency, program, and institutional level. Assessment has the support of the institution's academic and institutional leadership and the systematic involvement of faculty and appropriate staff.

> 8.4 The institution with stated goals for students' co-curricular learning systematically assess their achievement.

> 8.5 The institution uses a variety of quantitative and qualitative methods and direct and indirect measures to understand the experiences and learning outcomes of its students, employing external perspectives including, as appropriate, benchmarks and peer comparisons.

> 8.6 The institution defines measures of student success and levels and achievement appropriate to its mission, modalities and locations of instruction, and student body, including any specifically recruited populations. These measures include rates of progression, retention, transfer, and graduation; default and loan repayment rates; licensure passage rates; and employment.

> 8.7 The institution uses additional measures of student success, such as further education, civic participation, religious formation, and others, as appropriate to its mission, to understand the success of its recent graduates. Information from students and former students is regularly considered.

> 8.8 The results of assessment and quantitative measures of student success are a demonstrable factor in the institution's efforts to improve the learning opportunities and results for students.

8.9 The institution devotes appropriate attention to ensuring that its methods of understanding student learning and student success are valid and useful to improve programs and services for students and to inform the public.

8.10 The institution integrates the findings of its assessment process and measures of student success into its program evaluation activities and uses the findings to inform its planning and resource allocation and to establish claims the institution makes to students and prospective students.

Common Concerns about Assessment

Will this process affect my academic freedom?

Nothing inherent in the assessment process interferes or violates the academic freedom of the instructor. Program outcomes, assessment measures, scoring procedures, and the pedagogies used are decided by the faculty in the program. Resources on best practices are available for faculty, but in no way prescriptive.

Will this be more work for me?

Depending on where in the assessment process a program is, yes and no. It takes time and effort for programs to decide on mission statement, program outcomes, and assessment measurements. However, once a program has made these decisions, assessment should not take too much time out of your daily work. The Office of Assessment and Planning is available to help programs plan assessment that will take the least amount of time and effort as possible, while still providing rich data.

Will assessment information be used to evaluate faculty?

Absolutely not. This process is about assessing the effectiveness of programs, courses, and services, not individuals. In fact, because faculty and program chairs collect assessment artifact themselves and aggregate the data, as well as decide what data to share, administrators will not have access to information that can be tied to a specific faculty member.

Won't this eventually lead to standardized testing?

On the contrary, program assessment allows us to avoid standardized testing by given evidence that we are assessing student learning on our own. In fact, recent studies show that colleges that have moved to assessing student learning have decreased their use of standardized testing (49% in 2008 to 38% in 2015 - AAC&U, 2015).

Will the results have complete statistical validity?

While it is possible to create educational assessments that have high reliability and validity, the amount of time and effort it would take to create these assessments would outweigh the benefits. Many of the questions that faculty are interested in (Do our students remember information from one class to another? What topics do students have the most difficulty understanding?), do not need to be answered with complex statistical analyses.

Will the results be useful, and be used?

This depends on what information is collected and how the program uses it. If a program collects information 'for the sake of assessment' and then does not use that information, then it will not be useful or used. The best assessments answer questions faculty have about their students and program, and then are used to make decisions. Will the information be used at a higher level? It depends on to what extent programs provide information. If only a few programs are providing assessment information at the university level, it would be hard to use that information to inform change at the university level.

Why do assessment if our program is doing well?

Assessment is always a work in progress because we are always working to increase meaningful student learning. The primary purpose of program assessment is to improve the quality of educational programs by improving student learning. Even if the quality of a program is good, there is always room for improvement (Wehlburg, 2007).

I can't assess everything I want my students to learn.

It is a common misunderstanding that assessment requires that everything can be reduced to statistical measures. The purpose of assessment is to have objective results; but these results do not need not be quantifiable. If the faculty identify an important result that is not quantifiable, the process simply asks them to document the result and how the result was found.

What about Grades?

While grades do measure student learning, traditional grading offers one score or grade to represent student performance on many different outcomes. A grade does not necessarily provide information at the level of detail necessary to link student performance to a specific outcome. For example, a grade of C on an English paper may reflect adequate content, poor mechanics, average synthesis, and good effort, or may reflect poor content, adequate mechanics, and average synthesis.

Another limitation of grading for program assessment is that unless every instructor teaching a particular course assigns final course grades in exactly the same way (same assignments, same exams, same weights, same grading approach), you cannot be confident that one section's A is the same as another section's A.

Although grades have limitations for program assessment, you may be able to use grades for program assessment if they directly relate to the program's learning outcome, do not take into account information other than the outcome, and if grading methods are consistent across program faculty and courses. For more information see: Walvoord, Barbara, & Anderson, Virginia Johnson. (1998). Effective grading: A tool for learning and assessment. San Francisco: Jossey-Bass. This book is available in the CTL Library.

Steps in the Assessment Process

- Develop a program mission statement
- Develop program goals
- Develop program learning outcomes
- Create a curriculum map
- Decide on assessment measures
- Decide when to assess each outcome
- Choose target/criteria for success
- Collect and score artifacts/data
- Examine and analyze assessment results
- Create action plans
- Write an assessment report
- Carry out action plans
- Close the Loop

> "Successful assessment is an ongoing, iterative process. It is undertaken with the knowledge that the assessment process itself will be constantly updated and adapted to meet the changing needs of the institution, students, faculty, and public."
>
> Banta, Lund, Black, & Oblander, 1996

Mission Statement

A mission statement is a broad statement of purpose defining a program's philosophy and often describing values and aspirations. The purpose of a mission statement is to convey to others, both internal and external to the University, an understanding of why a program exists and what it does (Allen, 2004).

The mission statement should be clear, powerful, and broad enough to guide decision-making and provide the foundation for learning outcomes. In addition, the mission statement should be able to stand on its own and distinguish itself from other programs if the program's name were removed. Accrediting bodies also expect that program mission statements are consistent with the University's mission statement and vision (Banta & Palumbo, 2015)

Mission statements typically include the purpose of the program the values or philosophy related to the purpose, who the program serves, and the activities by which the program carries out its purpose. Mission statements should be brief (no longer than one page with most only a paragraph).

Program Goals

Program goals are broad statements that describe the long-term program targets or directions of development. They state in broad terms what the program wants to accomplish (regarding student outcomes) over the next several years (Allen, 2004). Program goals flow from the mission and can provide a framework for determining the outcomes of a program (Walvoord, 2004). It is best to develop between 3-5 program goals. Each program goal will likely have 1-2 program learning outcomes associated with it

Program Learning Outcomes

Program Learning Outcomes (PLOs) are specific statements that describe the expected knowledge, abilities, values and attitudes of a graduate of your program (Huba and Freed, 2000). Whereas program goals are broad statements of intended learning, program learning outcomes describe the specific, measurable behaviors that provide evidence that the goals are being met.

Program outcomes should include knowledge and behaviors that are critical to student success in the program and should distinguish a program from other programs. Each program learning outcome should be linked directly to one or more program goals.

Identifying Program Learning Outcomes

Wiggins and McTighe (2005) advocate a "backward design" when thinking about educational programs: starting from a set of desired results, faculty ask what evidence would demonstrate these results and then plan instruction and activities designed to lead to the outcome. Below are some ways to identify program learning outcomes.

Inventory of Current Assessment Practices

When beginning an assessment process, it is helpful to identify the assessment that is already taking place in the program. The assessment may currently be informal or at take place at the course-level, but can give a good idea of what information a program already has on student learning.

"Ideal Student" exercise

One of the best ways to determine program learning outcomes is to imagine an ideal graduate of your program. What would the ideal graduate know? What would the ideal graduate be able to do? What values would the ideal graduate possess?

Review of Program Documents

Program documents provide valuable information on the learning that occurs in the program currently. Some common program documents to review are:
- Course syllabi
- Course assignments
- Promotional materials
- Exams
- Catalog descriptions
- Accreditation Reports
- Program reviews
- Mission Statement

Use the "25 percent rule"

To figure out what faculty in the program value the most try the following exercise: Imagine that you had to reduce program material by 25 percent. What content or skills would you keep and which would you discard?

Delphi iteration

Delphi is a group technique to reach consensus about the most important themes before attempting to write specific outcomes (you will need an impartial mediator to run this activity)

1. In a brainstorming session, ask each member to build a list of knowledge/attitudes/behaviors/skills that is important for students in your program to learn
2. Have each member anonymously rank each item (e.g.: 1=very important; 2=somewhat important; or 3=not important.)
3. Place the items in rank order and show the (anonymous) rankings to the members
4. Repeat the ranking and discussion process until members converge on consensus.

Other Sources for Program Learning Outcomes

When creating learning outcomes, it may also be helpful to consult professional organizations, similar programs at other universities, methods books, peer institution websites, or banks of learning outcomes online.

Writing Program Learning Outcomes

After identifying the knowledge, skills and abilities that the faculty want to assess, actual learning outcomes are drafted. The foundation of an assessment plan is the learning outcomes, which guide the selection of measurements and assist in the interpretation of results. Therefore, it is extremely important to correctly identify, develop, and define the learning outcomes before other aspects of the assessment plan are initiated.

Many programs choose to use start each outcome with the same wording, for example "Graduates of the X program will be able to.... (action verb).... (specific content). Example: Graduates of the X program will be able to explain how X theory applies to the banking system.

Choosing an Action Verb

Given that learning outcomes focus on observable and measurable actions performed by students, the selection of an action verb for each outcome is crucial. Determining the best verb to use in a learning outcome can be challenging because of its need to accurately reflect the knowledge, skills and abilities being studied.

Program Outcome Tips

Remember to write outcomes in terms of student abilities, knowledge, values and attitudes and not on the process of instruction and learning.

~

Outcome statements should be framed in terms of the program and not individual courses or students

~

Program learning outcomes should be accepted and supported by all members of the program.

~

Developing appropriate and useful outcomes is an iterative process; it is not unusual to go back a number of times to refine outcomes.

~

An outcome statement should not impose restrictions on the type or number of assessment measures that have to be used to evaluate the outcome (i.e. 95% of students will pass the final exam in course XXX)

~

Remember the purpose of assessment is not to evaluate, judge, or compare faculty or programs, but to use what is learned from the assessment activities to improve programs.

~

Certain verbs are ambiguous and subject to different interpretations in terms of what action they are specifying. Some verbs that are linked to learning, such as *understand* or *know*, are poor choices in an outcome/objective statement because of this ambiguity. Instead you should select verbs that focus on observable and measurable action, such as 'describe' or 'list'.

Many programs will use a taxonomy as a reference for choosing action verbs. The first taxonomy of educational action verbs was Bloom's Taxonomy of Educational Objectives (1956). Bloom's taxonomy is a classification system useful for identifying student learning outcomes. The most well-known domains of Bloom's taxonomy is the cognitive domain.

Cognitive Domain

The cognitive domain is comprised of six intellectual skill levels arranged by increasing difficulty, academic rigor, and complexity. The cognitive taxonomy was revised in 2001 by Anderson and Krathwohl (see Table X)

Table 1. Cognitive Domain Taxonomy Revised

CATEGORY	DESCRIPTION
Remember	Ability to recall previously learned material
Understand	Ability to grasp meaning, explain, restate ideas
Apply	Ability to use learned material in new situations
Analyze	Ability to separate material into component parts and show relationships between parts
Evaluate	Ability to judge the worth of material against stated criteria
Create	Ability to put together the separate ideas to form new whole, establish new relationships

Table 2. Examples of Action Verbs for Cognitive Domain

Remember	Understand	Apply	Analyze	Evaluate	Create
Define	Choose	Apply	Analyze	Appraise	Arrange
Identify	Cite	Demonstrate	Calculate	Critique	Assemble
List	Describe	Dramatize	Categorize	Debate	Collect
Name	Determine	Employ	Compare	Diagnose	Compose
Recall	Differentiate	Generalize	Contrast	Evaluate	Construct
Recognize	Discriminate	Illustrate	Correlate	Judge	Create
Record	Discuss	Interpret	Differentiate	Measure	Design
Relate	Explain	Operate	Examine	Rate	Develop
Repeat	Express	Perform	Experiment	Score	Modify
	Interpret	Practice	Solve		Organize
	Restate	Use	Test		Plan
	Review	Utilize			Prepare
	Translate				Produce
	Respond				Propose
					Predict
					Reconstruct
					Synthesize

The Affective Domain

In addition to the traditional cognitive domain, Bloom also addressed the more nebulous affective domain. The affective domain (Krathwohl & Bloom, 1964) refers to attitudes, values, interests, appreciation, and feelings toward people, ideas, places and objects. Examples of affective learning outcomes are examining one's personal belief system, participating in an activity, or internalizing a value system consistent with the values of the program.

Affective objectives are very important, but many find them notoriously difficult to measure. The affective outcomes sought by programs may be best assessed by faculty observation, student comments during discussion, surveys of student opinion, focus groups, interviews, and writing assignments where students are asked to express their personal opinions or beliefs.

Table 3. Affective Domain Taxonomy

LEVEL	DESCRIPTION
Receiving	Being aware of or attending to something in the environment
Responding	Showing some new behaviors as a result of experience
Valuing	Showing some definite involvement or commitment
Organization	Integrating a new value into one's general set of values, giving it some ranking among one's general priorities
Characterization by Value	Acting consistently with the new value

Table 4. Examples of Action Verbs for Affective Domain

Receiving	Responding	Valuing	Organization	Characterization
Accept	Complete	Defend	Codify	Internalize
Attend	Comply	Devote	Discriminate	Verify
Develop	Cooperate	Pursue	Display	
Recognize	Discuss	Seek	Order	
	Examine		Organize	
	Obey		Systematize	
	Respond		Weigh	

The Psychomotor Domain

The last domain is the Psychomotor domain (not part of the original classification system), which describes psychomotor movement, patterns, and responses. As with the other domains, the levels of the psychomotor domain are to some extent hierarchical, describing increasingly difficult levels of skill. The Psychomotor Taxonomy below is adapted from three taxonomies (Simpson, 1972; Dave, 1975; Harrow, 1972).

Table 5. Psychomotor Domain Taxonomy

LEVEL	DESCRIPTION
Observing	Active mental attending of a physical event
Imitating	Attempted copying of a physical behavior
Practicing	Trying a specific physical activity over and over
Adapting	Fine tuning. Making minor adjustments in the physical activity to perfect it.

Table 5. Examples of Action Verbs for Psychomotor Domain

Bend	Grind	Organize
Calibrate	Handle	Perform (skillfully)
Construct	Heat	Reach
Differentiate by Touch	Manipulate	Shorten
Dismantle	Measure	Sketch
Display	Mend	Stretch
Fasten	Mix	Write
Grasp	Operate	

Figure 1. 3-D Model of Cognitive Taxonomy

Outcomes should be 'SMART'

Specific
- Define learning outcomes that are specific to your program

- Be as specific as possible in deciding the action verb you will use (know is more general than explain)

- Do not join elements in one objective statement that cannot be assessed by a single assessment measure

Measurable
- Make sure you can feasibly collect data for the outcome, and anticipate how you would measure the outcome

- Consider your available resources in determining the most reasonable method for collecting data

- Include more than one measurement measure to show students are competent in the outcome

Attainable
- When defining the learning outcomes and setting targets, use targets that will move you in the direction of your vision, but don't try to "become perfect" all at once

Results-Oriented
- You should determine what standards are expected from students in your program. For some learning outcomes, you may want 100% of graduates to achieve them. This expectation may be unrealistic for other outcomes

- You may want to determine what proportion of your students achieve a specific level (e.g., 80% of graduates pass the written portion of the standardized test on the first attempt). If you have previously measured an outcome, it is helpful to use this as the baseline for setting a target for next year

Time-Bound
- When defining the outcomes, it is important to describe where you would like to be within a specified time period (e.g., 10% improvement in exam scores within one year, 90% satisfaction rating for next year, 10% improvement in student communication performance within two years).

The University of Central Florida (February 2004). UCF Academic Program Assessment Handbook.

Curriculum Map

A critical part of the assessment plan in academic programs is the cooperative development of a curriculum map. Curriculum mapping is a technique that allows programs to align their courses and instruction with stated learning outcomes. This technique identifies where in the curriculum the defined program learning outcomes are being covered. This becomes a road map to how students will successfully satisfy the student learning outcomes identified by the program. The curriculum map also provides the program a better understanding of what is being taught, and serves as a tool to help make adjustments to the curriculum (Harden, 2001).

For example, assume that an identified program learning outcome is that students will be able to communicate effectively within the discipline in oral presentations. Construction of a curriculum map ensures that instructors of specific courses know they are responsible for giving students instruction on what makes a good oral presentation within the field, including appropriate visual aids, tone, level of formality, appropriate audience level, etc.

The curriculum map is usually constructed as a grid. Along one axis is a listing of the program's learning outcomes. Along the other axis is a list of the required courses in that program (and may also include elective courses). Within each cell, programs can indicate if the topic is I: Introduced, R: Reinforced, M: Mastered and A: Assessed.

- Introduced (I): Students first learn about key ideas, concepts or skills related to the outcome.

- Developed (D): Students gain additional information related to the outcome. They may start to synthesize key ideas or skills and are expected to demonstrate their knowledge or ability at increasingly proficient levels.

- Mastered (M): Students are expected to be able to demonstrate their ability to perform the outcome with a reasonably high level of independence and sophistication.

- Assessed (A): Students are asked to demonstrate their ability to perform the outcome

A "healthy" map means: Each learning outcome is introduced, developed and mastered, in a logical sequence, at least once across multiple courses. If each cell in the column is filled, it suggests redundancy and overlap related to that outcome in the curriculum. If few cells are completed or are missing an "I," "R," or "M," it is likely the curriculum is not covering that outcome as completely as faculty and/or the program might like. Each course should support at least one and ideally more than one learning outcome. If a required course does not seem related to any program learning outcomes, it provides the opportunity to ask whether the course should be required or whether an important learning outcome is missing from the program learning outcomes. If elective courses are included in the matrix, and an elective course is related to many outcomes, it may be that the course should be a required course. The curriculum map allows for faculty to discuss the current curriculum in relation to the program's mission and values.

Assessment Measures

Assessment measures are the techniques and tools used to assess student learning. Selecting appropriate assessments measures is important because the goal of assessing student learning is to determine the level of students' mastery of particular learning outcomes. If the measurement does not accurate assess the outcome, the results will not be useful to the program (McClendon & Ho, 2016). The most important criterion for any assessment measure is whether measure will provide accurate and useful information (Banta & Palomba, 2015).

Direct and Indirect Assessment

Assessment measures can fall into one of two main categories: direct measures and indirect measures. Direct measures of assessment require that students demonstrate their knowledge, skills, and learning by doing something. The strengths of direct assessments are that they are tied directly to the curriculum, can be integrated into routine faculty workloads, and can be used for grading as well as assessment. The potential limitations include unclear reliability and validity, and that creating direct assessments requires time to develop and to rate (Allen, 2004).

Indirect measures of assessment focus on student perceptions of learning and often involve surveys, interviews, or focus groups to ask students to self-report or reflect on their learning rather than to demonstrate it. The strengths of indirect assessments are that they are easy to create, use, and score, and that many individuals can be assessed at once. The limitations of indirect measures is that self-report may be inconsistent with actual learning (Allen, 2004).

> "You are note trying to achieve the perfect research design; you are trying to gather enough data to provide a reasonable basis for action."
>
> **Walvoord, 2010**

The Office of Assessment and Planning can help programs create surveys to use as indirect assessments.

Examples of Direct Measures

- Writing assignment
- Course Exam
- Capstone course project
- Portfolio
- Juried show/exhibit
- Case study
- Research project
- Standardized exam
- Certification/Licensure exam

Examples of Indirect Measures

- Exit survey
- Alumni survey
- Employer/Supervisor surveys
- Reflection journal/reflective essays
- Institutional level surveys (NSSE, Noel Levitz etc)
- Focus groups
- Interviews
- Activity Log

Qualitative and Quantitative Assessment

Assessment measures can also be categorized as either qualitative measures or quantitative measures. The most common type of measures are quantitative, however, qualitative measures are a legitimate form of assessment, and should be seriously considered in any decision regarding the choice of assessment.

Quantitative measures use numbers (or can be converted to numbers for data analysis); whereas qualitative data use words and are generally reported as a narrative. For quantitative data, the same information is usually collected from each participant in exactly the same way, and different responses are translated into a series of numbers (Allen, 2004).

> "The most compelling evidence is integrative, drawing as needed on both direct and indirect evidence of learning and bringing together quantitative and qualitative information."
>
> Hutchings, Kinzie, & Kuh, 2015

Qualitative assessments emphasize flexibility in data collection and focus on understanding processes and events, rather than precisely measuring them. However, these measures may also be more time consuming to collect and more difficult to assess since there may not be one simple solution, answer, or perspective that would be deemed correct (Banta & Palomba, 2015).

Examples of Qualitative Measures
- Exit interviews
- Formal recitals
- Open-ended questions on surveys and interviews
- Portfolio reviews
- Performance reviews
- Oral examinations
- Behavioral observations

Examples of Quantitative Measures
- Rubric scores
- GPA
- Grades
- Primary trait analysis scores
- Exam scores
- Demographics
- Forced-choice surveys
- Standardized teaching evaluations

Test Blueprinting

Quizzes and exams can also be used for direct learning assessments, providing that the test items are aligned with learning outcomes. The process of aligning test questions and learning outcomes is called test blueprinting.

The primary benefit of test blueprinting is that it makes direct learning assessments quick and easy. If you have already documented that questions 4, 7, 8, and 10 on your mid-term exam addresses your program's learning outcome 1, for example, when it comes time to collect data on how well students are achieving these outcomes you only need to extract the answers to those questions.

If you decide to design a direct learning assessment based on test items from an existing test, the first step is to read through the test with the learning outcome you are assessing in mind. Circle the test items that you believe demonstrate achievement of this outcome. If you are only assessing your course you may be content with your own judgments, but if you are assessing the program it is advisable to have one or more other faculty in your program verify that you have chosen the right items.

Next, you will need to transfer the test items into a table to help organize your data collection. For each outcome, you should list the questions relating to that outcome. Then you can enter what percent of students were able to correctly answer the questions relating to that outcome to use for assessment reporting.

Choosing Assessment Measures

The next step after you decide on the mission and goals of the program is to determine what assessment measures you will use to capture student learning.

It is important to consider your program culture when choosing assessment measures. In a program in which qualitative research is valued, qualitative measures should be incorporated into the assessment plan. Conversely, in a program in which faculty value numerical data, efforts should be made to include quantitative measures. If the data are not meaningful to the faculty, changes on the basis of the information are unlikely to occur.

When choosing your assessment measures, be sure to address the issues of participant attrition/retention, the actual amount of time involved, and cost and/or resources. Longitudinal studies are particularly vulnerable to these challenges. Any effective assessment plan will acknowledge these challenges and incorporate ways to address them within the development and implementation of the plan itself.

Choose assessment measures that allow you to assess the strengths and weaknesses of the program. Effective measures of assessment provide both positive and negative feedback. Finding out what is working well is only one goal of program assessment. Remember, the most effective program assessment plan is one that is closely linked to the curriculum and uses available information and resources to the greatest degree possible.

Multiple Measures and Triangulation

Relying on only one measure to provide information about the program will only reflect a part of students' achievement. Additionally, some program outcomes may be difficult to assess using only one measure. For each program outcome, a combination of direct and indirect assessment measures should be used. For example, responses from student surveys may be informative, however, when combined with students' test results they will be more meaningful, valid, and reliable.

Measurement Tips

Allow time for mistakes and for ongoing faculty input and discussion
~
Encourage and set time aside for faculty meetings to discuss assessment techniques and measurements
~
Comprehensive does not mean "assessing everything"
~
Faculty involvement is critical to the success of assessment. Feedback can be through group or individual discussion, e-mail communication, or other means.
~
Think about the ways in which you can use one source of information for a variety of course and program-level purposes.
~
Measures do not need to be complex. One essay question on an exam may be enough to assess a program outcome
~

Best practices suggest that each learning outcome be evaluated by three measures, a practice called triangulation. This should involve at least one, but preferably two, direct measures of student performance (Banta & Palomba, 2015).

Embedded versus Add-on Assessments

Embedded assessments use course work for both grading and assessment. Embedded assessments can be used to assess competency at the individual student level, the course level, and the program level (Allen, 2004).

Embedded assessment often does not require additional time for data collection, since instruments used to produce student learning information can be derived from course assignments already planned as part of the requirements. The usefulness of embedded assessments is built-in, as the assessment is an expectation rather than an aspiration and relate directly to the materials being taught. These assessments are also more likely to be examined and used by faculty members (Hutchings, Kinzie, & Kuh, 2015).

> "The work assigned by faculty has always been the most meaningful and natural source of evidence for documenting student learning
>
> Hutchings, Kinzie, & Kuh, 2015

Add-on assessments are additional tasks that go beyond course requirements and are usually given outside of the classroom such as during designated assessment days on campus. Generally they involve standardized testing. Because they are not typically part of the course grading structure, students are often less motivated to perform well (Banta & Palomba, 2015).

Authentic Assessments

Authentic assessments involve real-world activities that professionals in the field would encounter. Authentic assessments can be conducted either in the classroom or at fieldwork sites (Allen, 2004).

Local versus Standardized Assessments

Local assessments are instruments developed by faculty members within a program for internal use only. They are helpful in assessing standard-based questions (i.e., whether or not students are meeting objectives within the program), because they can be directly linked to program learning outcomes (Allen, 2004).

Standardized assessments are published instruments developed outside of the institution. They rely on a standard set of administration and scoring procedures and because of this are often times more reliable. These assessments provide information about how students in a program compare to students at other peer institutions or to national/regional norms and standards (Allen, 2004)/.

Selecting a Target/Criteria for Success

After identifying learning outcomes and ways to assess them, the next step is to identify standards of performance or targets. The purpose of identifying targets is to gauge student performance.

If you do not have a clear sense of how you expect students in your program to perform, then it becomes more difficult to evaluate their performance and draw valid conclusions about your results. Setting targets can be thought of as setting an a priori hypothesis.

In higher education, criteria for success are usually expressed in terms of the percentage of students who will meet a specified performance level. This is an arbitrary process since it is essentially a consensus judgment based on faculty's holistic impressions of how they expect students to perform. Because of this, it may take more than one try to set appropriate targets.

When identifying targets, it is important to set criteria that are reasonable in terms of what students are capable of performing. Targets that are too high or too low provide relatively little information.

The target identifies the minimum level of proficiency or a standard that determines if the student learning outcome is achieved or met per student. The target also typically includes the proportion of students who are expected to meet that standard for the academic program to be considered successful. For some student learning outcomes, all students would be expected to meet the minimum level of proficiency. For other student learning outcomes, 100% achievement may be unrealistic. Therefore, a percentage of students would be more appropriate criteria. Some programs choose a target for the percentage of students who are competent and another, smaller percentage of students who should score as highly competent. For previously measured student learning outcomes, the former criteria may become the current baseline, especially if the student learning outcome has been met. Finally, it is important that the faculty are aware that performance data will not be used to evaluate them so that target setting remains unbiased and fair.

Collecting Data

Collecting data involves gathering information from the students who will be assessed. Most commonly, the process of gathering assessments takes place in the last half of the fall and spring semesters (but can also take place at any point during the year).

Suggestions for successful data collection include:
- Strategize on who will gather the data and who will be responsible for monitoring data collection as early as possible
- Set a timeline for gathering data across the program or unit
- Send reminders to collect data
- Inform the people whose data is being collected that their information will be used to improve the program or unit

Ethical Considerations
While assessment activities are generally exempt from human subjects review (when the results are not going to be presented or published) the following should be considered:
- Include a statement on your syllabus that student work may be used for assessment purposes and allow students to opt-out
- Remove identifying information about students and instructors before sharing artifacts for rating
- Keep all artifacts in a secure location, available only to faculty in the program
- Aggregate students and courses to avoid identifying particular students and/or faculty

- If there are any concerns about whether the assessment should be reviewed by the HSRB, consult with the HSRB chair.

Sampling

One of the decisions that must be made about assessment is how to select students to participate. Because we are a relatively small university, we have many programs with small numbers of graduates each year. For this reason, most small programs will not use sampling at all. However, when the population size of the program is large and classes have multiple sections, assessing every student in the program may become an unmanageable task that will require the use of a sampling procedure.

It is important to consider the manner in which students are selected since this has important implications in terms of how results may be generalized to the entire population of students in the program. When a sample is biased it will be difficult to draw valid conclusions about how the program is working.

Sampling bias occurs when sampling procedures consistently under-represent some kinds of groups while over-representing others. To avoid bias, every student ideally should have an equal opportunity of being selected. For this to occur, there must be an unbiased sampling frame, one that does not exclude certain individuals (i.e., the worst students, a particular gender, major or race, etc.).

Simple random sampling provides the best means of obtaining a representative sample. However, in most instances this is difficult to do since access to the entire population of students in a program may not be available. In most cases, programs often rely on multistage sampling that is not truly random. For example, courses or sections of courses are selected and then students are systematically sampled from these classes. This is done by first arranging students' work in alphabetical order, randomly selecting a starting point, and then selecting every kth student.

The main point is to try to make the sample as representative as possible by not excluding any particular group of students. In courses with multiple sections, it is important to include all sections to avoid a possible "professor" effect.

The other important consideration in sampling is the number of students to include in the sample to draw valid conclusions. Obviously the greater the sample size, the more confidence you may have in your conclusions. A very simple rule of thumb is to include all students in the sample if there are there fewer than thirty-five students who can be assessed. If there are thirty-five or more students you may want to use sampling rather than assessing all students.

Scoring Artifacts

Rubrics

In the assessment process, artifacts are the samples of student work produced in response to faculty developed assignments. A rubric is a document that articulates the expectations of an assignment, task, or activity by listing criteria or priorities, and describing levels of quality.

Using rubrics to evaluate summative assessments is valuable for various reasons; they serve as tools for faculty to communicate what is expected of students, guide both faculty and students in the areas and components that are most critical for learning, and describe performance levels to help both faculty and students come to a mutual understanding of what it means to perform at a particular level (McClendon & To, 2016).

Rubrics are criterion-referenced, rather than norm-referenced. Raters ask, "Did the student meet the criteria for level 5 of the rubric?" rather than "How well did this student do compared to other students?" This is more compatible with cooperative and collaborative learning environments than competitive grading schemes and is essential when using rubrics for program assessment because you want to learn how well students have met your standards.

Rubric Tips

Faculty using a rubric should keep in mind that rubrics are seldom perfect. They typically require revisions even when the most careful consideration goes into their preparation, and even when they are developed by experts in the field.

~

When students receive rubrics before starting an assignment, they understand how they will be evaluated and can prepare accordingly.

~

Rubrics can help with equity, allowing all students to start on the same level, regardless of background.

. ~

A rubric scoring guide can be used to evaluate a student's performance based on the sum of a full range of criteria rather than a single numerical score. Rubrics can be used to score essay questions, projects, portfolios and presentations.

Rubrics can be checklists, analytic or holistic, and they can be created for any content area including math, science, history, writing, foreign languages, drama, art, music, etc.

Checklists

Checklists are the simplest type of rubric and list accomplishments that are evident in the students' work. Checklists give a list of content that should be included in students' work. The content may be listed sequentially indicating the order in which it should occur. The rater marks each item on the checklist that the student has completed or included in their work. Checklists do not give an indication of the quality of student work.

Holistic Rubrics

Holistic rubrics describe levels of performance with regards to the overall quality of the paper or project as a whole, without considering the components of student work separately. Holistic rubrics assess the overall quality of student work by providing descriptions of student work at different levels of performance. These descriptions define the overall characteristics of student work at each level of performance. Holistic rubrics provide an overview of student performance and have the advantage of quick scoring. However, holistic rubrics do not differentiate between multiple traits and therefore may not provide as detailed a picture of student performance as an analytic rubric. They are most useful when a single trait is sufficient to define the quality of student work.

The first step in developing a holistic rubric is to identify the components in the student work that are related to the learning outcome. These components should be linked to the student learning outcomes developed as part of the program assessment plan. After the components are identified, the next step is to decide how many levels are necessary to classify the quality of students' work. The descriptors chosen for the mid-point level of the rubric should describe the primary characteristics of the students' work that meet the minimum acceptable program standard.

Analytic Rubrics

Analytic rubrics guide the scoring of student work on multiple traits first, and then sum the individual scores to arrive at a total score. Analytic rubrics provide separate evaluation of student work on multiple traits. They can pinpoint particular areas where students need improvement, which can be used during planning to suggest opportunities to improve instruction. One drawback to the use of analytic rubrics is that they require more time to use than holistic rubrics. The first step in developing a holistic rubric is to identify the components in the student work that are related to the learning outcome. These components should be linked to the student learning outcomes developed as part of the program assessment plan. After the components are identified, the next step is to decide how many levels are necessary to classify the quality of students' work. The descriptors chosen for the mid-point level of the rubric should describe the primary characteristics of the students' work that meet the minimum acceptable program standard.

The next step is to decide how many levels are necessary to classify the quality in student work for each trait being measured. The descriptors chosen for each level of the rubric should describe the primary characteristics of students' work for each of the selected traits. Sometimes it can be difficult to find meaningful descriptors for several levels of performance. Remember, all of the characteristics listed must be reflected in the students' work in order to be scored as meeting that level of the rubric.

Avoiding Bias in Scoring

Rate the performance of all students on one task before going to the next to keep scoring criteria in mind

~

Whenever possible rate performance without knowledge of student's name to avoid halo effects

~

Use multiple raters for high-stakes decisions

~

Train your raters to calibrate their scoring by bringing them together to review their responses in order to identify patterns of inconsistent responses

~

Analytic rubrics can be used for both assessment and grading. Traits not associated with the learning outcomes can be included on the rubric (i.e. correct number of sources) and included in the total grade, but not added to the total score for the outcome.

Analyzing Assessment Results

Though very important, defining and collecting evidence about program learning outcomes is just the beginning of the program improvement process. Program faculty need to draw conclusions about the meaning of the evidence, develop a plan to improve student performance, and then collect more evidence to determine the effect of the changes. Including program faculty in all steps of the assessment process is important to ensure its meaningfulness and effectiveness. The inclusion of faculty insights is probably most important component in interpreting results and identifying strategies for improving student learning.

Quantitative Data Analysis

It is important to make the most out of the information you collect through appropriate analysis and interpretation. Assessment points out the strengths and weaknesses of your program and provides evidence for change. While data analysis can be relatively complex, for the purpose of assessment it is usually basic.

Two important steps should be completed before analyzing data. The first step is to review the data visually. Reviewing data visually has two benefits: It allows for the identification of outliers and possible mistakes, and it enables basic patterns or trends to emerge. For example, it may be clear that all students who took a particular class had difficulty with a particular outcome.

The second step of the process is to determine the appropriate method for analyzing the data. This can range from simply counting the number of successful students to a high powered statistical analyses. The two key factors are first to make sure the analysis method fits the data; and second, to ensure that method aligns with the program's needs.

Types of Data

There are two main types of quantitative data used in assessment: categorical and numerical data.

Categorical data

Categorical data are based on groupings or categories for the evaluation of student performance. For example a simple passed/failed score is categorical because there are two groups into which students can be placed. Often rubrics generate categorical data by using a scale of "exceeding expectations," "meeting expectation," and "failing to meet expectations".

Numerical data

Numerical data are based on scales that reflect student performance. Tests which are scored based on the percentage of questions answered correctly generate numeric data.

Direct measures can generate either categorical or numerical data. Student's papers rated on an assessment rubric may be categorized as "meeting standard" or "failing to meet standard". However, the papers may be scored on a numerical scale indicating the overall quality of the paper with respect to the learning outcome.

Indirect measures can also generate either categorical or numerical data. By asking students on a questionnaire: "Did you have sufficient writing in the program?" a program would compile categorical data based on those saying "yes" and those saying "no." However, by asking students to indicate how strongly they agree with a statement like "there was sufficient writing required in my program", numeric data could be generated by applying an agreement scale. (5 – Strongly agree, 4 – Agree, 3 – Neither, 2 – Disagree, 1 – Strongly disagree).

Assessment's focus on student achievement of learning outcomes typically requires the determination of counts and percentages. Together they show clearly the number of students involved in the activity and the rate of successful display of the outcome. Both categorical and numerical data can be analyzed using counts and percentages. Numeric data has the additional benefit of being able to be analyzed using

descriptive statistics. Mean, median, and mode provide useful information to interpret data by allowing for easier comparison between groups and tests for significant differences.

Distributions

By examining how data are distributed around measures of central tendency, particularly the mean and median, a richer understanding of the data emerges. The standard deviation represents the average deviation of scores from the mean. Small standard deviations in student performance indicate that performance levels varied little across students in the sample. Large standard deviations indicate a greater variability in levels of student performance. Standard deviations are commonly reported with the mean.

Percentiles represent the percentage of a distribution of scores that are at or below a specified value. They are often reported with the median which by definition is the 50th percentile. For example: a median score of 75 on a final exam would be the 50th percentile indicating 50% of students scored above 75 and 50% scored below. By examining the 25th, 50th, and 75th percentiles one can gain a sense of a student's performance relative to the group.

Missing data and valid responses

Working with assessment data, there are many instances when data will not be available for every student. As a general rule, missing data should be excluded from calculations of percentages and descriptive statistics. If a program has ten (10) students, and eight (8) submit a needed paper for the assessment of an outcome; then eight (8) submitters become the basis of the analysis.

Qualitative Data Analysis

The most common type of qualitative data analysis used in assessment is content analysis. Content analysis involves searching for common themes within responses. Content analysis has more credibility if multiple people are involved. Reviewers should summarize common themes and the extent to which there was consensus on those themes in the responses. Faculty may also plan ahead and developing a coding scheme with particular words and phrases and then counting the number of times those words/phrases are used in responses (Allen, 2004).

Reporting Assessment Results

Once data have been gathered and analyzed the information now should be reported. A report might be for an internal audience such as program faculty, administration, or students. External audiences may include accreditors, state programs, or a grant-funding agencies.

Presenting Data

Tables and graphs are useful in presenting analysis because they focus attention to specific results. Tables are useful for reporting multiple percentages and frequencies, comparison of student performance with stated performance standards, and some descriptive statistics. They provide an ordered way for readers to see results quickly for each outcome measure without having to search through text to find a particular result. Graphs can further enhance the visual impact of assessment. Graphical representations of results show differences in variables, which makes graphs highly effective in showcasing assessment results.

When sharing the results of program assessment it may be useful to report each learning outcome and outcome measure paired with the corresponding results of the analyses, which joins the multiple outcome measures (direct and indirect) for each learning outcome. Next, compare the results with the specified performance standard and discuss the implications of the data as they relate to the program. Both strengths and areas for improvement are discussed, because showcasing program success is just as important as identifying areas for improvement, when it comes to making data based decisions about the program.

Direct Quotations

Using direct quotations from student responses to qualitative measures makes the report more powerful as the results represent what the students are telling us. However, if overdone, those reading the report may only read the quotations and skip over the rest of the information in the report.

Four Questions to Address in the Report

Suskie (2011) indicates four questions every assessment report should answer:
1. How do you define a successful student?
2. What's the evidence that students meet your definition of success?
3. Are you satisfied with your results? Why or why not?
4. If you're not satisfied, what could be the cause and what are you going to do about it?

Using Assessment Results

The purpose of assessment is not achieved simply through the collection of vast amounts of data. Rather, assessment's purpose is to answer questions, shape better policies, make better decisions-all designed to improve student success and strengthen institutional performance. Simply performing assessment activities is not the same thing as using assessment results (Kinzie, Hutchings, & Jankowski, 2015).

The results of the assessment should be used to identify changes to improve the program. These changes could be related to the assessment process, the curriculum, or the academic process. Typically these changes are called 'action plans.' Before making any changes, programs should share assessment results with all program faculty and discuss them together, so that any action plans can be decided on collectively.

> "The real challenge begins once faculty, staff, administrators, and students at institutions try to use the evidence to improve student learning."
>
> Kuh, 2011

In addition to benefiting the department, changes based on assessment can be used by the University for accreditation purposes. NECHE (2016) specifies that "the institution needs to go beyond discussing evaluative processes and relate how it actually makes use of the results of its various institutional effectiveness efforts."

The following are areas within the academic environment where changes may be implemented as a result of assessment:

Changes to Assessment Plan
- Revising program outcomes
- Revising assessment measurement
- Collecting additional data
- Changing data collection measures

Changes to Curriculum
- Changing teaching practices
- Changing or enforcing prerequisites
- Changing course sequences
- Revising or updating course content
- Adding/Removing course(s)

Changes to Budget/Resources
- Increasing classroom space
- Adding lab resources
- Hiring or re-assigning faculty or staff

Changes to Academic Processes
- Changing the frequency or timing of course offerings
- Improving technology and/or other instructional aids
- Offering professional development
- Revising advising standards or processes

Changes to Promotion and Marketing
- Communicating and celebrating student performance and success
- Communicating student voices and perceptions to stakeholders
- Industry feedback from external assessments

Evaluating the Process

At the end of an assessment cycle, it is important for the faculty involved in the assessment project to reflect on the process and to share their thoughts with all program faculty so that everyone in the program stays informed and engaged.

Some questions to reflect upon are:
- Did you have a positive or negative experience implementing your assessment measures?
- What were students' reactions to the assessment process?
- What did you find especially effective in the assessment process?
- What did you particularly dislike about the process?
- What would you change about the process? Why?
- What will you do again? Why?
- What do the results suggest for assessment in your program?

Closing the Loop

Once the action plan is implemented and given time to have the desired effect, faculty use the means of assessment to once again determine if students are performing at the desired level. This is called 'closing the loop.' If students are still not demonstrating competency, they develop another action plan, implement it, and assess it. That cycle continues until the criteria are met.

Closing the loop encompasses analyzing the results of action plans by re-assessing outcomes to determine the effect those action plans had on student learning. Closing the loop actually occurs after action plans have been implemented and you have had a period of time to see whether the action plan led to improvements or not. It is important to note that some improvements take time to manifest themselves in the assessment results. For example, if the improvement strategy was to add additional reading materials to an introductory course to be able to expand students' knowledge on a certain topic, then several years may pass before the students are able to take the senior level course in which their knowledge is assessed. Because closing the loop is an evaluation of an action plan once it is implemented, this section of the assessment report cannot be completed until the next assessment cycle (Kinzie, Hutchings, & Jankowski, 2015).

This is one of the major reasons why it is recommended that outcomes are assessed for as long as possible in order to gather longitudinal data that can provide a clearer picture of trends. Outcomes can also be placed under dormancy for a couple of years if there are no major changes in the data and the criteria are being consistently met or exceeded. These outcomes can be reintegrated after a given period in order to reassess whether the outcome is being successfully met.

> "Completing the assessment—that is, measuring the impact of actions taken to improve student learning, and the ultimate stage—gaining evidence of improved student learning is assessment's nirvana, perhaps."
>
> Kuh et al., 2015

References

Allen, M. J. (2004). *Assessing Academic Programs in Higher Education*. San Francisco, CA: Jossey-Bass.

Anderson, L. W., & Krathwohl, D. R., et al. (Eds.) (2001). *A Taxonomy for Learning, Teaching, and Assessing: A Revision of Bloom's Taxonomy of Educational Objectives*. Boston, MA: Allyn & Bacon.

Baer, L. L. (2017). Connecting the dots: Accountability, assessment, analytics, and accreditation. *Planning for Higher Education Journal, 46(*1), 1-16.

Banta, T. W., & Palomba, C. A. (2015). *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education, 2 Ed.* San Francisco: Jossey-Bass.

Bloom, B., Englehart, M., Furst, E., Hill, W., & Krathwohl, D. (1956). *Taxonomy of Educational Objectives: The Classification of Educational Goals. Handbook I: Cognitive Domain.* New York, NY: Longmans, Green.

Bransberger, P., & Michelau, D. K. (2016). *Knocking at the College Door: Projections of High School Graduates, 9th Ed*. Boulder, CO: Western Interstate Commission for Higher Education.

Burke, J. C., & Minassians, H. P. (2004). Implications of state performance indicators for community college assessment. *New Directions for Community Colleges, 126,* 53-64.

CIHE (2016). *Self-Study Guide.* Boston, MA: Commission on Institutions of Higher Education

Dave, R. H. (1970). The psychomotor domain. In (Eds.) R. J. Armstrong et al., *Developing and Writing Behavioral Objectives*. Tucson, AZ: Educational Innovators Press.

Driscoll, A., & Cordero de Noriega, D. (2006). *Taking Ownership of Accreditation: Assessment Processes that Promote Institutional Improvement and Faculty Engagement.* Sterling, VA: Stylus.

Dunn, D. S., McCarthy, M. A., Baker, S. C., & Halonen, J. S. (2011). *Using Quality Benchmarks for Assessing and Developing Undergraduate Programs.* San Francisco, CA: Jossey-Bass.

Harden, R. M. (2001). AMEE Guide No. 21: Curriculum mapping: a tool for transparent and authentic teaching and learning. *Medical Teacher, 23*(2): 123-137.

Harrow, A. J. (1972*). A Taxonomy of the Psychomotor Domain: A Guide for Developing Behavioral Objectives.* New York, NY: D. McKay.

Hutchings, P., Kinzie, J., & Kuh, G. D. (2015). Evidence of student learning: What counts and what matters for improvement. In (Eds.)Kuh, G.D., Ikenberry, S.O., Jankowski, N., Cain, T.R., Ewell, P.T., Hutchings, P., & Kinzie, J. *Using Evidence of Student Learning to Improve Higher Education.* San Francisco, CA: Jossey-Bass.

Krathwohl, D. R., Bloom, B. S., & Masia, B. B. (2006). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 2: Affective Domain.* New York, NY: McKay.

Kuh, G.D. (2011, January). Foreword. In (Eds.) C. Blaich and K. Wise. *From gathering to using assessment results: Lessons from the Wabash National Study* (NILOA Occasional Paper no. 8). Urbana, IL: National Institute for Learning Outcomes Assessment.

Kuh, G.D., Ikenberry, S.O., Jankowski, N., Cain, T.R., Ewell, P.T., Hutchings, P., & Kinzie, J. (2015). *Using Evidence of Student Learning to Improve Higher Education.* San Francisco, CA: Jossey-Bass.

McClendon, K., & Ho, T. (2016). Building a quality assessment process for measuring and documenting student learning. *Assessment Update, 28*(2), 7-14.

Montenegro, E., & Jankowski, N. A. (2017, January). *Equity and assessment: Moving towards culturally responsive assessment* (Occasional Paper No. 29). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Palomba, C. A., & Banta, T. W. (1999). *Assessment Essentials: Planning, Implementing, and Improving Assessment in Higher Education.* San Francisco: Jossey-Bass.

Provenzis, S. (2010). *Regional Accreditation and Student Learning Outcomes: Mapping the Territory* (Occasional Paper No. 6). Urbana, IL: University of Illinois and Indiana University, National Institute for Learning Outcomes Assessment (NILOA).

Simpson, E. (1972). *The Classification of Educational Objectives in the Psychomotor Domain (Vol. 3).* Washington, DC: Gryphon House.

Suskie, L. A. (2004). *Assessing Student Learning: A Common Sense Guide*. Bolton, MA: Anker.

Suskie, L. A. (2011). *Summarizing and Using Assessment Results.* Online Seminar.

The University of Central Florida (2004). *UCF Academic Program Assessment Handbook.*

U.S. Secretary of Education's Commission on the Future of Higher Education. (2006, September 19). *A Test of Leadership: Charting the Future of U.S. Higher Education. A Report of the Commission Appointed by Secretary of Education Margaret Spellings.* Washington DC: U.S. Department of Education.

Walvoord, B. E. (2010). *Assessment Clear and Simple: A Practical Guide for Institutions, Departments and General Education.* San Franscisco, CA: Jossey-Bass.

Walvoord, B. E., & Anderson V.J. (1998) *Effective Grading: A Tool for Learning and Assessment.* San Francisco: Jossey Bass.

Wehlburg, C. M. (2007). Closing the feedback loop is not enough: The assessment spiral. *Assessment Update: Progress, Trends and Practices in Higher Education, 19*(2), 1-3.

Wiggins, G. & McTighe, T. (2005). *Understanding by Design* (2 Ed). Alexandria, VA: Association for Supervision and Curriculum Development ASCD.