# Outcomes and Assessment Practices Handbook

Produced by SDCCE Student Learning Outcomes Task Force

October 2021

San Diego College of Continuing Education

San Diego College of Continuing Education
Outcomes and Assessment Practices Handbook

**What is Assessment?**

In the words of assessment expert Barbara Walvoord, "Assessment is the systematic collection of information about student learning, using the time, knowledge, expertise, and resources available, in order to inform decisions that affect student learning." (Walvoord, 2) Institutions collect data about student learning to determine curricular effectiveness: Are students learning skills necessary to succeed in a course, a program of study, in their communities, or in employment? These data can help individual instructors modify their courses or assist units with programmatic changes. They can help instructors, staff, administrators, and prospective employers develop more effective educational offerings. Data analysis can identify curricular "bottlenecks," as well as suggest resolutions to reduce those barriers to learning and student success. Analysis of assessment data can also provide powerful evidence of an institution's success to outside accreditors, like the Western Association of Schools and Colleges (WASC), and to internal bodies, like governing boards.

There are several foundational principles of good assessment practice:

- **Assessment is the measurement of student learning**. It is intentionally designed to determine what students learn in an educational institution. Most often, assessment refers to the measurement of student learning within the classroom, or "academic assessment," but assessment also occurs across an institution of learning. As examples, measurement of student learning that occurs within co-curricular units, such as student

affairs or student life and development, may focus on interpersonal or leadership skills development that occurs with engagement in wellness workshops and student clubs, respectively. Measurement of classroom learning ordinarily relies on the "artifacts of learning," particularly analysis of assignments that are intentionally designed (and collected) to elicit evidence of learning outcomes. For instance, if a course learning outcome is "Prepare a brief persuasive speech," the assignment would be designed to ask students to demonstrate that outcome.

- **Assessment is designed to identify both challenges and successes in learning.** Good assessment helps document learning. It provides evidence for what students are learning and what they are not. Good assessment also identifies where improvements need to be made, taking the guesswork out of learning improvement plans. In this sense, assessment results that show what students have not learned is both successful and valuable. Good assessment does not always document only student learning successes.

- **Assessment of student learning is an intentional process.** The foundation of that process is a mission or goal statement, whether at the institutional level or at a unit or at the course level. Institutional members design their assessment of student learning first by identifying learning outcomes. At the institutional level, Institutional Learning Outcomes describe the collective skills, behaviors, and knowledge that every student of the institution will acquire by graduation from a program; the same definition applies

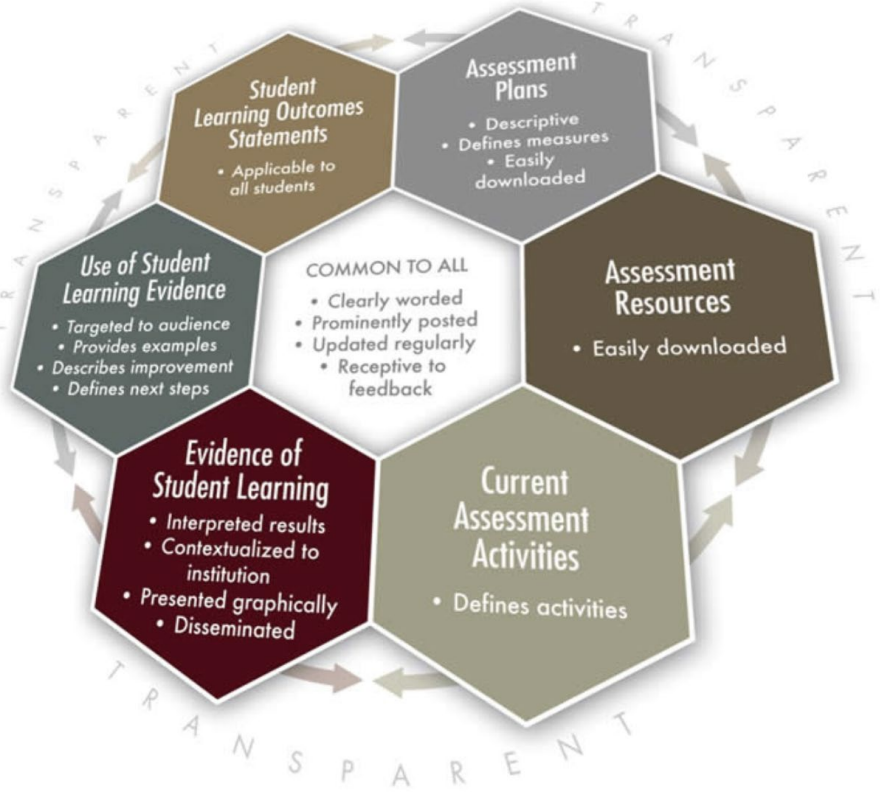
for program and course level learning outcomes, but at those levels of learning, respectively. Learning outcomes communicate clearly to students and other institutional stakeholders, as well as community members, what students will know and be able to do. San Diego College of Continuing Education has four Institutional Learning Outcomes encompassing the skills of Social Responsibility, Effective Communication, Critical Thinking, and Personal and Professional Development.

- **Assessment is an ongoing process.** WASC's accreditation framework notes that assessing student learning is at the core of the accreditation process. The process includes identifying learning goals, planning assessment activities, implementing an assessment plan, analyzing the assessment results, and revising instruction, curriculum, and action plans for improvement. Good assessment processes are ongoing, but also sustainable. Institutions are generally cautioned against assessing all learning outcomes in every assessment cycle.

- **Assessment is distinct from Accountability.** Assessment measures learning. Accountability is the concept within higher education that uses such metrics as graduation rates, time to degree, returns on investment (employment rates and average earnings for graduates), and other such "measurements" to inform students, parents, educators, and the public about an institution's overall effectiveness. The College Portrait, published by the organization Voluntary System of Accountability, provides sketches of public institutions of higher education and their "performance" in not just

learning, but also in such data points as cost of attendance, student debt load upon graduation, and institutional characteristics (particular strengths, average class size, student demographics, and so forth). In general, “accountability” is a term often used by governing boards to determine a program or an institution’s worth. One way to think about these distinctions is this: A four-year public institution may have a phenomenal graduation rate in which 91% of all students graduate within six years, have a low cost of attendance, and a diverse student body, but if assessment is not taken seriously, one may not have a good sense of what students know and can do upon graduation.

In sum, assessment is an ongoing process of intentionally measuring student learning. It is an organized and transparent process. There are no hidden practices; assessment resources, plans, processes and results are visible to institutional members and others. The National Institute for Learning Outcomes Assessment (NILOA) defines a six-dimensional framework for transparency in assessment that can be visualized in the process chart below.




Other institutions, including several community colleges, have modified NILOA's Transparency Framework to better reflect their own institutional practices.

**Purposes of Assessment**

Assessment has multiple and intersecting purposes for institutions. **For institutions,** meeting regional accreditation requirements is the most important purpose of assessment. As WASC notes in its statement on institutional quality and accreditation, "Accreditation is a voluntary dual-purpose process that schools (1) must be worthy of the trust placed in them to provide high-quality learning and (2) clearly demonstrate continual self-improvement." Assessment provides the evidence required to meet these accreditation standards.

Internal practices contribute to assurances of learning: Are students learning the skills and content they are taught? Are faculty members, as individuals and collectively, effective in their work? Does the institution provide the academic, social, financial, and emotional support services needed for learners to be successful? Is the institution effective in its allocation of resources -- not just financial, but also human -- that maximize learning?

**For Faculty,** assessment activities also provide faculty with valuable opportunities to reflect on their classroom practices. Has the curriculum changed in ways that require revision? Might pedagogical innovations lead to better student learning outcomes? Combined with student demographic data, assessment results may indicate areas where teaching practices could benefit from inclusivity and diversity interventions. When coupled with opportunities for faculty and staff development, faculty reflection on student learning can bolster the overall qualifications and effectiveness of teaching at an institution.

**For Community Members,** when institutions publicly list evidence of student learning on assessment or institutional websites, students, potential employers, and community partners can all be assured that the institution is devoted to ensuring student success.

Scorecards may list percentages of students who achieve intended goals, linking those goals to additional studies -- serving as guideposts to student success. Such publicly displayed evidence of learning indicates how student pathways to higher learning are implemented at the institution.

**For students,** publicly-accessible data provides reassurance that they are on the right path to success. Educational research has shown that adult learners can excel when offered opportunities to chart their own learning paths. SDCE's efforts to divide CTE programs into shorter certificate programs align with assessment's focus on documenting discrete skills. Listing a certificate program's assessment results and student achievement levels ("scorecards") offers students, both prospective and current, information about program expectations, thereby providing them with opportunities to self-direct their learning experiences. Students who achieve a defined level of mastery of specific SLOs might be provided with a digital "badge" that signifies their learning, as part of a more comprehensive certificate program. Students' incremental acquisition of such badges can lead to greater sense of self-efficacy, higher levels of motivation, and a better sense of belonging to the institution -- all recognized to be powerful predictors of student learning. Providing evidence of student success may also boost equity initiatives, especially for students who may have personal histories of past schools' poor performance. Using assessment to identify barriers to success can help students prepare for success by normalizing student support resources, such as utilizing campus-based tutoring services when enrolled in a difficult math class required for a student's major or certificate. Additionally, assessment can help chart multiple pathways for success through basic skills, core curriculum, and CTE and academic programs.

**Assessment Fundamentals: Departments and Units: The Basics**

**Student Learning Outcomes.** As described, student learning outcomes (sometimes called student learning objectives, or SLOs) are statements that indicate what students will understand, know, and be able to do once they complete a discrete learning experience -- in a course, program, or certificate or associate's degree program. Good student learning outcomes are the cornerstones for good assessment; without clearly written SLOs, assessment of learning is difficult. A good student learning outcome statement has several characteristics:

- It is clearly written.
- It does not contain jargon or overly technical vocabulary.
- It uses an action verb, as from Bloom's taxonomy of learning.
- It identifies knowledge or a skill that is discrete.
- It is specific to a course, program, or institution.
- It describes an outcome that can be measured.

SLOs are powerful ways to communicate to students what they will learn at an institution. Above all, SLOs are communication devices -- they signal what is essential learning. They indicate to students what the institution considers both important and achievable. SLOs signal an institution's clear commitment to student learning. Importantly, assessment professionals distinguish *outcomes* from *goals* in this general way: Outcomes are measurable and achievable, goals are aspirational and what you hope students will gain from a program. In their influential text, *Learning Assessment Techniques: A Handbook for College Faculty,* authors Elizabeth Barkley and Claire Howell Major make this analogy: A learning goal is seeing the "target," while a learning outcome is aiming and shooting a "bullseye."

**Writing Student Learning Outcomes.** As with other assessment processes, writing student learning outcomes is an intentional and collaborative activity. Writing SLOs should never be a top-down administration-led process. Rather, faculty and staff members use their intimate knowledge of the many forms of learning to direct the SLO writing process. There is not one prescribed method for writing SLOs, but in general, at most institutions, stakeholders generally follow a process as below:

- Working in specific units, faculty and staff gather documents to determine learning: mission statements, syllabi, samples of student work, informational brochures and other such material that help to identify desirable learning outcomes.
- Identify appropriate achievement levels for each outcome, using Bloom's Taxonomy, a framework of verbs used for characterizing six levels of learning, from basic knowledge acquisition to advanced skills application. This site at UNCC has a good illustration of basic differences between achievement levels, using specific examples from K-12.
- Faculty and staff participate in defining outcomes at the unit level; activities may include brainstorming, using "backwards design" techniques (pioneered by Jay McTighe and Grant Wiggins), and outlining expectations for achievement, among other methods. Collaboration is key. Collaboration works much like peer review, strengthening the clarity and meaning of SLO statements.
- Once faculty and staff have drafted SLOs, share with other stakeholders, especially students. Are the LOs clear to students? Students should be able to understand what is expected of them, using the SLOs as a guide. Systematically collect feedback.

- Collaboratively revise and redraft LOs, based on feedback. Share revised LOs with assessment committees and other decision-making institutional members.
- Revise as needed, with curriculum and programmatic changes, trends in workforce, and changes in disciplines.

**Determining Appropriate Targets for Assessment.** Once SLOs have been finalized, assessment activities proceed, often on an annual basis. Institutional, programmatic, and course needs often determine what learning should be assessed. New institutional requirements -- a core English communication course, for example -- would prompt focus on learning in that course at the end of the first year or two of the requirements. Changes in curriculum would also prompt a focus on learning in the new curriculum courses. Accreditation requirements to demonstrate learning of institutional objectives

**Methods for Capturing Data.** Before deciding on a method to capture data, it is essential to first clarify the types of data needed to determine learning efficacy and to articulate acceptable criteria for achievement. Assessment professionals generally characterize two types of assessment measures: Direct and Indirect measures. Both types of assessment can measure student skills, behaviors, attitudes, and knowledge acquisition. Direct assessment methods measure actual learning using student "artifacts," such as assignments, tests, and demonstrations of knowledge. Indirect assessment methods involve observation or self-report of learning, and as such, are usually regarded as less reliable measures of actual learning. The following clarifies these distinctions, giving examples for each type of measurement of data.

**Direct Assessment.** To collect data of what students know and can do, there are a variety of ways to capture student demonstration of knowledge and skills. Common assessment methods in higher education include results of standardized (state or national) exams, end of course or end of program examination, assignment, or skills demonstration. For instance, a passing score on the National Council Licensing Examination for Practical (Vocational) Nursing (NCLEX) is a direct assessment, as it measures practical nursing skills necessary to practice safely. Other direct assessments include the GED, HSE (HiSET), and TASC exams that determine high school graduation equivalencies. Such national exams are consistently monitored for validity and accuracy to provide objective measurements of knowledge and skills attainment. However, most disciplines do not have national exams to utilize for end-of-program assessment of student learning.

In their place, most programs use various types of end of program (or end of course) assessment methods, using stand-alone (often locally-developed) exams or demonstrations of knowledge and skills. Some institutions, for instance, require students to complete a capstone course (at SDCCE, this may be an "advanced" or culminating course in a program), conduct and finish a project (such as complete a business plan), perform skills demonstration (design a functional mobile app), or collect evidence of learning throughout the program, using a portfolio approach. In co-curricular assessment, submission of an updated resume might be a valid measurement of learning for a career services unit.

In academic areas, valuable evidence of learning may also be collected using other types of assignments and exams, which may be administered at any point in the course. "Embedded assessment" is the practice of using classroom assignments as direct assessments of learning.

Embedded assessments may be either summative (end of module, course, program) or formative (low-stakes assessments that identify barriers to learning in process). When using assignments as embedded assessment, the purpose of the assignment must be clearly aligned with the learning outcome to be measured; evaluation should be informed by a rubric or other standard measurement tool; in best practice, raters are normed in evaluation decisions. That is, evaluators meet together with the same sample assignments, rate them individually, discuss scores, and come to a collective decision about levels of student learning. Evaluators who have participated in norming efforts may then independently evaluate student work, although some institutions prefer collective assessment activities be held on certain "assessment days."

**Pitfall: Grades v. Evaluation.** In using direct assessment methods, especially when using embedded assessment methods, faculty may be tempted to substitute assignment or course grades for evaluation of SLOs as assessment. At first glance, the two might seem equitable. But they are not. Think about alignment of SLOs: An assignment may require that students demonstrate achievement of two or more SLOs, but only one SLO is being assessed in the annual assessment cycle. There are other confounding factors, as well: A student may not turn in an assignment, or may turn in a partially complete assignment, or a student may have done well on an assignment but turned it in late and incurred a late penalty. In the last case, a student might receive a C (satisfactory) but actually have demonstrated a skill at the advanced level. Or in the first case, a student might receive an F, but in assessment practice, a missing assignment is a null case; failure to demonstrate achievement does not signify failure. The following chart illustrates some of the differences between grades and evaluation of assignments.

| Grades | Evaluation (Assessment) |
|---|---|
| Evaluation of a single student | Evaluation of a group of students |
| Evaluation of a single or multiple SLOs or tasks | Evaluates performance of a single artifact |
| May not be aligned with a SLO to be assessed | Aligned with a single SLO to be assessed |
| Evaluation tool may or may not be utilized | Evaluation tool (answer key, rubric) is used, raters are normed |
| Grade may reflect other factors external to performance of a specific task | Measures achievement of a specific outcome, so no "radio interference" |

**Indirect Assessment.** Indirect assessment might be best characterized not as distinct measurements of learning, but as measurements of beliefs, thoughts, behaviors, and attitudes about learning. Indirect assessment techniques include a wide variety of methods useful to determining students' thoughts and behaviors about learning: Surveys and questionnaires, interviews, focus groups, observations, among other types of methods. Faculty, staff, and assessors may use indirect assessments to gauge students' satisfaction with educational offerings, compare student views of learning with direct assessments of their achievements, and determine students' sense of belonging within the institution, a critical factor contributing to student success. Indirect measures are often used in co-curricular and student affairs assessment. For instance, a financial aid office (FAO) might track the number of times a student visits the office as a rough measure of appropriate learning behavior; however, such a measure does not indicate if a student actually learned about financial assistance (though one could reasonably assume that frequent visits would translate into actions to obtain assistance). Comparing FAO visit frequency to students' actions (filing a FAFSA on time, applying for a

CalGrant or other state educational assistance, obtaining SNAP or veteran's educational assistance, and so forth) would provide a better measure of student learning about financial aid opportunities.

Student perception of learning surveys are among the most frequently used indirect assessments: The National Survey of Student Engagement (NSSE) asks students in four-year higher education institutions to evaluate their engagement with key indicators (how many times a student uses math to solve problems, writes essays, visits faculty and the library, and so on) that are generally associated with student success; the survey does not measure student learning directly. This is true for most survey tools: They may provide insight into students' attitudes about institutions and learning, but do not measure learning directly. However, data collected in indirect assessments are often invaluable, as they indicate where students perceive they may need support. Good assessment practices utilize a mix of direct and indirect measures to assess learning and student success.

**Alignment.** In assessment practice, alignment is a crucial issue: The assessment method used must align with (or match) the SLO to be assessed. If alignment is not present, the assessment results will not reflect intended goals. Assessment methods must measure specific SLOs. A useful analogy is this. If the primary SLO of an introductory public speaking class is to "Prepare and deliver a two-minute informational talk to an audience of peers," then the assessment instrument must be designed to match that SLO. For instance, an inappropriate measurement of student achievement would include a quiz on how to deliver a speech, evaluation of a written draft of the speech, or a questionnaire that asked students to rate their comfort in

giving their speeches. Instead, appropriate assessment methods might include analysis of a video of the student giving a short speech or faculty rating of the speech completed during observation of the student's talk.

**Pitfalls: Misalignment and Methodological Issues.** It is sometimes tempting to design shortcuts to assessments that on the surface may seem to appropriately measure student learning. Take this example of the assessment of a Writing Center's tutoring program for students. The Center was tasked with evaluating whether or not student attendance at its tutoring sessions was effective in promoting student success, measured by course grades. As we know, this design itself is problematic as grades and learning are not always synonymous. The Center collected data on student engagement each time the student attended a tutoring session. Tutors noted names, writing problems addressed, interventions and services offered, and noted, where applicable, the outcome (assignment grade) achieved. Students, however, did not consistently report back to their tutors grades they received on their writing assignments. All of this data would have provided a rich repository from which to draw preliminary conclusions about the tutoring program's effectiveness. Instead, Center representatives designed a survey and distributed it to students; the survey asked students to respond to this question: "Did the tutoring service help you be more successful on your writing assignments?" Students were provided with only two possible answers: "No, I still struggle with writing" and "Yes, it was great, tutoring saved my paper and my grade!" Notably, the way in which the survey was designed imposed a framing bias (one that essentially endorsed a rigid either/or determination) by allowing just two answers and thus cannot be considered an appropriate or valid assessment

of learning. Moreover, the survey did not measure learning at all -- only attitudes about services provided to augment learning.

In sum, there are many different methods to collect data about student learning: Direct assessments that measure student learning through evaluation of "artifacts" (student work), examinations, formative assessments (measuring learning in process), student performances of skill (building things, music recital), clinical observation of skills, and so forth. Indirect assessments include such items as surveys, focus groups, questionnaires, attendance records, and the like, that provide insight into attitudes and behaviors about learning, but may not measure actual demonstration of achieving student learning outcomes. What one hopes to measure will determine, in part, the method used to collect assessment data.

**Data Capture and Analysis.** Regardless of method, in assessment practices, data capture needs to be intentional, systematic, and consistent. When collecting artifacts or conducting alumni surveys, for instance, for assessment purposes, determining appropriate sample size is key. In small programs, it might be optimal to assess every student's performance; results might give a complete and clear picture of learning. However, assessing all students' work is not always practical. Instead, a statistically representative sample of the "population" (student group) can be calculated using standard methods such as Slovin's formula. Generally, a representative sample with a confidence level of 95% (calculated as 0.05) is sufficient, often requiring a sampling of about 10% of a population, chosen randomly. Randomizing data to be selected can

ensure integrity of representation. Use a random number generator to randomize selection of data.

How does this work? A scenario: An institution plans to assess an ILO (written communication) in first-year writing courses. After determining the ILO/SLO to be addressed, faculty convene to determine appropriately aligned assignments to be evaluated. The composition faculty design a signature assignment, in this case a five-paragraph expository essay, to utilize for assessment. The composition department offers 15 sections of the first-year writing course each term, with an average of 15 students per section, for a total yearly population of 450. Evaluating 450 different essays, even if they are brief, is unsustainable -- and largely unnecessary. Instead, the faculty chose to use a sample of these essays to evaluate; they calculate that at a 95% confidence level, they needed to select 46 essays. Rather than asking each faculty member to submit 1-2 essays per section (which may not be representative of all student performance), the composition program collates enrollment sheets, uses a random number generator to generate 46 random numbers (between 1 and 450), then selects those essays as indicated by the generator. (In this case, the random number generated is 9, so the program selects every ninth essay for evaluation, across all sections. Alternatively, a random integer generator can identify 46 unique integers between 1-450, then faculty select those essays from the collated group, such as essay numbers 8, 35, 81...and so forth.) Then, with the stack of anonymized (but numbered) essays in hand, the composition faculty (who have already participated in a norming session) use an evaluation rubric to assess each essay, with two faculty members evaluating each essay and averaging their scores. Raters record their scores on

an excel file or in the LMS. Essay authors may be identified by student ID numbers, but not names to limit potential evaluator bias, especially if there are concerns about equity gaps in learning. In this case, unique identifier numbers can be matched with student demographic data. Many institutions, however, use enrollment and learning management software packages that can identify equity gaps using grades to augment assessment results.

In other disciplines, machine-scored examinations can be selected using these same methods. A less precise method is simply to ask faculty to choose every 5th or 10th (or any other number) artifact from each section for evaluation, though this method might result in problems of representation, given last name distribution if chosen alphabetically.

**Rubric Use.** Using a collaboratively designed rubric, aligned with the SLOs to be assessed, is ideal. In some cases, assessment committees may design an assessment rubric. Rubrics take the "guesswork" out of evaluating student work. Well-designed rubrics describe discrete levels of learning, using performance descriptors. For instance, in the above case, the composition faculty have determined that grammar and spelling are one component, or dimension, of written composition. In this case, the rubric dimension might look like this, below:

| Criteria | Exceeds Expectations | Meets Expectations | Needs Improvement | Does not meet minimum |
|---|---|---|---|---|
| **Writing mechanics - Spelling and Grammar** | Two or fewer errors; Uses college-level writing strategies. | Usage is generally correct, though several minor errors may be present. Errors do not distract from essay quality. | Multiple errors diminish essay clarity. | Writing strategies and usage are not successful. Essay contains errors that make meaning unclear. |

In this case, the rubric will likely include additional dimensions (argument, supporting evidence, organization, and so on) for the evaluation of the written communication ILO. Importantly, assessment rubrics need not have more than three levels of learning (exceeds expectations, meets expectations, does not meet expectations), though most faculty assessors specify four or more dimensions to identify more discrete achievement levels.

**Special considerations in Indirect Assessment.** When carefully designed, an indirect assessment can sometimes be used to evaluate student achievement of SLOs in special cases. Here is such a scenario: A highly technical program has an SLO that requires students to identify and use discipline-specific technical language (in this case, specific words). During an exit interview (indirect assessment method), prospective graduates are asked to explain how they would solve a technical problem. To accurately explain the methods for solving the problem, students must employ four terms correctly (let's say "aardvark, ligonberry, tuba, and zebra").

As evaluators conduct their interviews, they make a notation when a student uses each term correctly. Evaluation sheets for such interviews might look like this:

| | **Names technical word** | **Uses word correctly in context** |
|---|---|---|
| **Aardvark** | | |
| **Ligonberry** | | |
| **Tuba** | | |
| **Zebra** | | |

Evaluators would simply need to make tally marks for each category; performance levels may have been identified prior to the interviews (e.g., to meet expectations, students must name all four words and use each correctly in context at least once). In such a scheme, determining assessment results is mostly just an exercise in counting tally marks.

These are just two examples of assessment methods. For a more extensive list of methods, see Barkley and Major, *Learning Assessment Techniques: A Handbook for College Faculty* (Jossey Bass, 2016).

In the analysis of assessment data, many institutions adhere to specific achievement benchmarks (70 - 80%) to determine effective learning. That is, at least 70% of all students assessed must score at the "Meets Expectations" level or higher. In some cases, an institution may set a higher benchmark, especially if warranted in CTE curriculum, by use of external assessment tools (like exams), or in the case of programs in which students must meet certain performance levels to continue.

**Pitfalls and Problems: Legal considerations.** Because of FERPA, legal requirements may dictate how data are stored and evaluated. Institutions typically store data on institutionally-supported drives, software, or cloud services. Use of external storage solutions, such as the commercial service Dropbox (or even Google Drive), can be problematic for storing students' personally identifiable artifacts. Use of institutional learning management software to store student work is preferred when that software incorporates data management tools. Consultation with institutional IT departments and appropriate administrative personnel may be necessary to ascertain how "raw" assessment data should be maintained.

**No Frills Assessment for Busy Faculty.** In most cases, exams and embedded assessments provide the most readily available set of data. However, in both cases, faculty and staff need to consider alignment in order to generate valid results. Another no-frills assessment option is faculty or staff observation of skills demonstration or specific behaviors within classroom environments or support offices. Walvoord defines a no-frills assessment system as one in which units collect data about learning in their own courses, especially when assessing ILOs (or general education SLOs). Because faculty and staff are primarily responsible for data collection and analysis, institutions are encouraged to eschew complicated and expensive assessment methods and tools. Locally-developed, tested, and implemented assessments are preferable. A caution about "pre- and "post-" assessment schemes. A decade ago, many institutions encouraged faculty to employ "pre-course and post-course" assessment methods. In such schemes, students would often take a brief quiz about key course contents early, usually week

1 in the term, then at the end of the term, they would retake the quiz. Comparing pre-course with post-course achievement is inherently difficult; further, assessment professionals now question the usefulness of such measures, as this method is unable to directly indicate that classroom learning leads to improved scores.

Assessment committees and faculty often ask "How much evaluation is enough?" Best practices in assessment caution against attempting to assess every SLO every assessment cycle (usually annually). Instead, consider the purpose of assessment: Is it just to document learning? Is it to assess whether a pedagogical or curricular revision has improved learning? Is it to test whether an aberrant or unexpected result from the prior cycle is valid? The purpose of assessment activities should drive considerations of what SLOs should be evaluated, how often, and on what scale.

**Closing the Loop.** In assessment, the phrase "closing the loop" refers to the practice of designing assessments, systematically collecting and analyzing data, then making adjustments to educational practices and reassessing for improvement. Assessment alone does not improve learning, as Keston Fulcher, et. al., noted in the 2014 NILOA Occasional Paper *A Simple Model for Learning Improvement: Weigh Pig, Feed Pig, Weigh Pig*. Instead, learning improvement gains are made through a process of assessment, effective interventions, and reassessment, or "closing the loop."

Effective interventions that lead to improvement in student learning require that faculty and staff accurately identify and describe learning gaps in courses and gaps in curriculum that impede or negatively affect student learning. Sometimes, assessment results unmistakably identify these gaps. Assessments in a high-failure introductory computer programming course, for example, may reveal that students' math skills are insufficient to facilitate learning. Other times, though, gaps are not clearly demarcated by assessments. Faculty, staff, and assessment committee members can boost identification of these gaps by adopting some of the following strategies. When interpreting assessment results and crafting assessment reports, assessors can

- Discuss results with all faculty (including adjuncts and paraprofessional staff)
- Obtain student views on assessment results using focus groups or classroom discussions
- Involve co-curricular and student support staff in discussions of results
- Consult with assessment committees or other knowledgeable faculty
- Become familiar with key assessment terms, especially bottlenecks, gateway courses, and hidden prerequisites
- Do research on learning improvement and pedagogical innovations in one's field
- Consult professional organizations, community members, and advisory boards
- Engage in meaningful self-reflection as individuals and members of an institutional community
- Pursue faculty development opportunities; even regularly reading *The Teaching Professor* and *The Scholarly Teacher* blogs can be immensely helpful

- Check out NILOA's vast repository of occasional papers and reports

**Action Issues.** As assessors write up their results, it is even more difficult to pinpoint effective interventions to improve learning, or "action" issues. Identifying barriers to student success then designing effective interventions at times seems like monumental tasks. In one well-studied case, student learning was directly related to city bus schedules: The transportation company changed the bus schedule, leading to later arrival time for the first bus of the day. In turn, the change in the bus schedule meant that many students were late to class and thus missed out on essential learning experiences. More often, a variety of factors have an impact on learning: student preparation, instructional quality (especially assignment design), instructional modality (passive versus active learning), and student well-being, among others. Questions of equity and social justice are also of note. See the 2017 *New Directions in Higher Education* issue on "Improving Teaching, Learning, Equity and Success in Gateway Courses" for information on new interventions for addressing equity. Similarly, the TILT-Higher Ed protocol for designing transparent assignments can assist faculty with implementing effective learning interventions.

**Curriculum Mapping and Development.** Engaging in course and curriculum mapping is a critical part of assessment practice and designing learning improvement interventions. Curriculum mapping identifies where in a program's curriculum essential knowledge and skills are taught. Program learning outcomes are listed, then courses are identified where students are introduced to, can practice, and then demonstrate mastery of those PLOs. Curricular learning gaps are evident when a program may have introductory and mastery level courses, but nowhere in the curriculum is there space for students to practice skills. The Eberly Center at

Carnegie Mellon University has developed a curriculum mapping tool for use in mapping program curriculum. At a more granular level, the course mapping guide, an initiative of UC San Diego's Teaching + Learning Commons, is a comprehensive guide, including templates, for faculty use. Systematically defining and mapping course and curriculum outcomes is also useful for more clearly illustrating student pathways to success.

**Reporting Requirements.** Normally, institutions determine their reporting requirements based on agreed upon assessment cycles. WASC generally expects institutions to engage in assessment on an annual basis, though specific requirements are left to institutions. Some programs may choose to complete multiple assessments each year, depending on department or unit needs. Most institutions develop an assessment reporting form of 2-3 pages, in which programs and units identify SLOs assessed, methods used, data collected, findings and interpretation, and identify action issues. Assessment committees, acting in conjunction with faculty, staff, and leadership, may be tasked with reviewing program and unit assessment reports.

Assessment reports are valuable ways to "tell your story" to both internal and external stakeholders. An assessment report should not just function as an internal document. Thus, when writing reports make them accessible by adopting some of these practices:

- Craft the report in non-technical language, where possible
- Consider a wide audience: students, community members, district board, potential employers, and the public

- Reformat the report into brochure or newsletter form for external distribution, digitally or in hard copy format
- Use techniques for visualizing data -- colorful graphs instead of data tables, for instance
- Create infographics that highlight particular successes
- Advertise success
- Post results on the SDCE SLO Website or on institutional or program pages

**The Assessment Cycle.** SDCE adheres to a standard academic assessment cycle, featuring annual integrated planning processes incorporating staff and faculty from all programs (basic skills, CTE, Emeritus, Disability and other student support services), as well as representatives from appropriate administrative offices. A new assessment cycle formally begins during Spring Flex days (Spring Assessment Week) in Spring term, when representatives from programs and units meet to discuss courses taught, review Student Learning Outcomes, select one SLO that all agree to assess; they select an activity or artifact to be assessed, then determine a method of assessment. Rubrics or other evaluation tools may already exist for some SLOs; if not, then faculty and staff will need to devise such tools. As noted, valid measurement of student learning can occur using a variety of tools -- from standardized tests to instructor-scored checklists of observable behaviors. The alignment between the SLO to be assessed, the artifact or activity assessed, and the tools used needs to be clear. During Assessment Week, personnel may also determine or revisit acceptable baseline and target levels of learning. A baseline simply identifies where students are, while a target specifies where programs want them to be at completion.

At SDCE, formal assessment of student work or behaviors generally commences after Spring Flex days, when faculty or staff rate student work. In this sense, most SDCE academic personnel utilize embedded assessments, in which an SLO is evaluated using existing course work (assignments) or artifacts. Faculty rate SLO achievement using evaluation tools as they grade assignments; alternatively, assessment may precede or follow grading. It is imperative that evaluators not conflate grading an assignment with evaluating a specific SLO. It is also important that evaluators create assessment tools that accurately and equitably measure student achievement.

During the summer and fall terms, SDCE faculty and staff analyze assessment data and complete their annual reports for submission via survey to Anthology, a software package that incorporates student enrollment management, institutional effectiveness, student success, retention, and engagement, as well as alumni and advancement (development) functions. Additionally, in the Fall term, units and programs inform the Planning, Research, and Institutional Effectiveness Committee how they will incorporate those data into plans for learning improvement; as the academic year progresses, instructional, curricular, or programmatic changes may be implemented to this end. All of these assessment activities during an annual cycle are included in formal program reviews.

**An example: English Language Learning Program.** As an example, in a fictional English Language Learners program, during Assessment Week, the program faculty have decided to assess student achievement of their program's SLO 1. *Students will write a five-paragraph*

*academic essay of 450-500 words, using appropriate grammar, spelling, and punctuation at an accuracy level of 70%.*

Faculty develop a rubric, similar to that on the top of page 20. They decide that they will employ an embedded assessment method, with all faculty participating. Student work from six different ELL writing courses are scheduled to be assessed: ELL 101, 102, 103, and 201, 202, 203. Near the end of the term, faculty meet to discuss use of the rubric and practice evaluating sample assignments (a norming activity). As they grade final essays, instructors rate students using the rubric; their results are tabulated and submitted to the program assessment coordinator or chair. Initial assessment shows that only 20% of all learners demonstrate achievement SLO 1. (*Students will write a five-paragraph academic essay of 450-550 words, employing appropriate rhetorical strategies and using appropriate grammar, spelling, and punctuation at an accuracy level of 70%.)*

Here, the baseline (20%) is well below the target (70%) level of SLO achievement. There is no need for alarm here, as there may be multiple explanations for shifting baselines. At mid-point in the annual assessment cycle, program faculty revisit their assessment strategies, realizing that assessment results need to be disaggregated by course level. At the top level, ELL 203, 73% of learners meet or exceed the assessment target. But in ELL 101, none meet or exceed the target. Such a result points to effective learning -- students start at low levels and proceed to higher levels of achievement. While a 20% baseline seems too low, it is acceptable given the sampling method and aggregation of data.

There may be other explanations for low baseline levels: statistical anomalies or differentials, demographic or population differences, instructional and pedagogical factors, as

well as methods used to sample or evaluate student artifacts. For instance, a baseline might appear artificially low if enrollments are much higher in lower-level ELL courses than in, for instance, ELL 203. Even employing sampling techniques could skew results lower if far more ELL 101 students were assessed. Students enrolled in ELL 203 may be more motivated and successful learners, as they complete their sixth term in ELL courses, or they may represent a different student demographic than those enrolled in ELL 101 courses (longer term residents of the U.S. for example, and thus more familiar with the English language).

During the Fall term, the ELL faculty decide to disaggregate all the results, by course level, to identify any possible barriers to learning. This action constitutes appropriate mid-program assessment methods, wherein student learning at varying levels is assessed. What might the faculty learn from this activity? Often, they might find fault with the rubric used for evaluation or misalignment between the SLO and student artifacts. In the former case, additional learning levels added to the rubric might provide more precise analysis of learning. In the latter case, students in ELL 101, 102, and 103 are not asked to write essays of longer than 150 words, so their achievement of the SLO that specifies 450-500 words simply cannot be measured. A reexamination of artifacts for students in these courses may show that for brief paragraph-length writing assignments (of fewer than 100 words), students achieve 3%, 8%, and 19% accuracy in usage in the three courses, respectively. Learning gains are evident here. Nonetheless, faculty teaching ELL 102 and 103 have decided to implement assignment changes to help students prepare to write longer essays: They scaffolded a longer essay by assigning brief writing assignments, emphasize multiple revisions and review, then ask students to collate these assignments to create longer essays. Additionally, the program faculty developed a rubric

that incorporated intermediate levels of learning to capture mid-program student achievement. Here are appropriate use of assessment results at the mid-program level.

This example points to the importance of planning, collaboration, and discussion of results, as well as preparing for learning improvement. Intentionality is key: These are well-planned and executed assessment activities, with meaningful discussion of results among faculty motivated to improve student learning.

**Advanced Methods in Assessment.**

When planning future assessments, commit your unit's plans to paper, even informally. An organic approach to assessment planning is preferable to a more rigid model. For instance, consider this scenario: Program A has three SLOS; each year, Program A assesses a different SLO, for a full SLO assessment turnaround time of three years. Program B has three SLOs but the faculty agreed that because the previous year's assessment of SLO 1 identified major barriers to learning, for which the program implemented new learning interventions, in year two, SLO 1 should be reassessed, then in year three, SLO 2 would be assessed, and so on - thus falling "behind" Program A's SLO assessment schedule. Both assessment schedules are acceptable, just so long as action issues are identified and acted upon in both programs. Assessment for assessment's sake is not useful.

When planning for assessment, in addition to WASC and institutional demands, programs should take into consideration their unique goals. Has a program developed new courses? Revised its Program learning outcomes? Experienced increases or declines in enrollment? Had an infusion of students from different student demographics? Encountered disciplinary or trade innovations, as may be the case in IT and CTE? Are there other considerations that warrant investigation during the annual assessments? Another consideration: Has the institution developed and employed a universal assessment planning document? If so, may programs augment or modify those assessment planning templates? Or in other words, what *must* be assessed in the current cycle and what might wait until next year?

Getting faculty and staff buy-in for assessment can be difficult, especially when faculty and staff experience assessment activities as add-ons to their primary duties. Reframing

assessment as an everyday activity and inviting wide participation can help to create a culture of assessment. Focusing on documenting -- and celebrating -- student learning can shift the spotlight away from faculty grievances. Also key to improving buy-in is formally and informally acknowledging the work and efforts of those who do participate in program or institutional assessment.

**Learning Domains.** Educators recognize three major learning domains: Cognitive (signifying knowledge), Psychomotor (Skills), and Affective (attitudes and perspectives). Different programs will emphasize different domains and to different degrees. The Emeritus Department may emphasize cognitive and affective domains, while Business Information Technology may prioritize Cognitive and Psychomotor domain. When mapping courses and curriculum, assessors should identify relevant learning domains, as approaches to assessment and methods employed will vary from one domain to another. Explicitly identifying domains also signals to learners what is prioritized in a program and allows them to better consider program "fit." Explaining learning domains may also help students to identify personal strengths while fostering student learning in general.

**Creating a Learning System.** In their book *Degrees that Matter,* authors Natasha Jankowski and David Marshall, both affiliated with NILOA, argue that well-developed assessment structures contribute to the creation of a "learning system." This learning systems approach centers student learning, not teaching, at the core of an institution's work. Building out assessment plans by inviting participation of community members, students, and employers helps to create

coherent learning pathways. Assessment can help faculty identify more appropriate or different course sequencing, or provide helpful learning interventions to improve student success. Devoting time to developing sustainable assessment structures in which there are clear and consistent expectations for assessment, faculty and staff are empowered to identify goals and plan assessment activities and in which wide participation is encouraged is worthwhile.

In conclusion, assessment is a set of intentional practices that can improve student learning while making more transparent what works and how students can better achieve their goals.

**Sustainable Assessment Structures.**

For some time, faculty and staff have often viewed assessment as time-intensive, mandated but unfunded activities required by institutions. Assessment should not be an add-on activity; it is fundamental to the jobs of faculty and staff. Creating sustainable assessment structures -- methods and practices that can be employed seamlessly in the everyday work of faculty and staff members -- is key to managing what can otherwise become a burdensome assessment workload.

In establishing Spring Assessment Week, SDCE has implemented a model that addresses sustainable structures: Personnel are given the time and resources to develop sustainable structures. Additionally, the goals of SDCE speak to current ideas of sustainability in higher education. Recent literature on "sustainable educational practices" defines "sustainability" in higher education as a set of practices that prepare students for lifelong learning, rather than an idea that graduation from a program signals an end to learning.

To further develop a sustainable assessment structure, annual assessments need to be linked closely to improvement of and for learning. As such, good assessment is a continuous process; formative assessment activities, in particular, are consistently incorporated into everyday teaching practices, in a way that provides both students and instructors with information about what students are learning and how they are learning. Sustainable assessment leads students to self efficacy and boosts students' metacognition. Faculty may employ classroom pedagogical practices (process analysis, IDEALS techniques, problem recognition skills work, exam wrappers) and other classroom assessment techniques, such as

those described in Angelo and Cross; Barkley and Howell, throughout the academic year. Programs can draw upon this larger body of evidence to make data-informed decisions about formal assessment activities, as well as using the results of classroom assessment as valuable data about how and what students learn.

Calendaring assessment activities, incorporating classroom assessment techniques, developing well-defined roles and practices for staff and faculty, providing time and space for collaboration, and empowering faculty and staff to determine what and how to assess, while also encouraging the use of assessment results to improve learning all contribute to sustainable assessment structures. Collaboration is key; no single person should be responsible for all assessment in a unit. Leveraging faculty time and effort by embedding assessment in a variety of operations and processes is essential. Another element of sustainable structures includes creating a culture of learning improvement in which students, staff, and faculty share responsibility for documenting learning. Institutional recognition of effective assessment is also important; exemplary practices and programs should be commended publicly. Finally, providing resources for faculty and staff development is critical to developing sustainable assessment structures, as the landscape of assessment in higher education and our understanding of the science of learning is rapidly changing.

**Appendix: Brief Examples of Assessments in Adult and Continuing Education**

Adult Basic Education:

"Muddiest Point" activity. Angelo and Cross, *Classroom Assessment Techniques.* Outlined in this Purdue University Handout, n.d.

Northstar Digital Literacy assessments, public versions (free).

"Invent the Quiz" activity. K. Patricia Cross Academy (free).

Career and Technical Education:

"Creation of and Reflection on Personal Development Plan" for ePortfolio. Lehman and Rogers-Cooper, *The Occupational Therapy Assistant Program: A Case Study*, NILOA Assessment in Practice bulletin, August, 2021.

"Formative Performance Review" activity. McGaughy, Hopper-Moore, and Farkas, *Career and Technical Education Professional Development and Formative Performance Reviews*, EPIC, October 2013.

Student Support Services:

"Develop Your Brand" personal development activity. Mathews, *An Outcomes Based Approach to Career Development*, NILOA Assessment in Practice bulletin, September, 2019.

Student Needs Assessment questionnaire. From *Needs Assessment Toolkit*. California Community Colleges.

Senior Citizen (Emeritus) Education:

BERI Student Engagement Checklists. (Can be adapted for use in any classroom setting.) Lane and Harris, "A New Model for Measuring Student Engagement in Large University Classes," *Research and Teaching,* 2015.

Weekly Budget. Collins and Holden, "Measuring the Impacts of Financial Literacy: Challenges for Community Based Financial Education," *New Directions for Adult and Continuing Education,* Spring 2014, 79-88.

**Bibliography**

Angelo, Thomas, and Cross, K. Patricia. *Classroom Assessment Techniques: A Handbook for College Teachers.* 1993.

Banta, Trudy W. and Palomba, Catherine A. *Assessment Essentials: Planning, Implementing and Improving Assessment in Higher Education.* 2014.

Barkley, Elizabeth, and Major, Claire Howell. *Learning Assessment Techniques: A Handbook for College Faculty.* 2016.

Condon, William, et. al., *Faculty Development and Student Learning: Assessing the Connections.* 2016.

Dirksen, Julie. *Design for How People Learn.* 2e. 2016.

Eyler, Joshua. *How Humans Learn: The Science and Stories beyond Effective College Teaching.* 2018.

Gannon, Kevin. *Radical Hope: A Teaching Manifesto.* 2020.

Henning, Gavin, and Flower, Darby. *Student Affairs Assessment: Theory to Practice.* 2016.

Hundley, Stephen P. and Kahn, Susan. *Trends in Assessment: Ideas, Opportunities, and Issues for Higher Education.* 2019.

Jankowski, Natasha, and David Marshall. *Degrees that Matter: Moving Higher Education to a Learning Systems Paradigm.* 2017.

Kaplan, Matthew, et.al., *Using reflection and metacognition to improve student learning: Across the Disciplines, Across the Academy.* 2013.

Kuh, George D., Ikenberry, Stanley, et.al., *Using Evidence of Student Learning to Improve Higher Education.* 2015.

Metzler, Eric, and Kurz, Lisa. *Assessment 2.0: An Organic Supplement to Standard Assessment Practice.* 2018. (Addresses new methods of sustainable assessments.)

National Institute for Learning Outcomes Assessment publications.

Suskie, Linda. *Assessing Student Learning: A Common Sense Guide.* 3e. 2018.

Walvoord, Barbara E. *Assessment Clear and Simple: A Practical Guide for Institutions, Departments, and General Education.* 2e. 2010.