CoEnv Teaching Support Team
Best Practices for Student Evaluation of Course Learning Objectives
April 2024

Evaluation is a general term signaling that an assignment or assessment (quiz, exam) is designed to address whether students are demonstrating progress toward achieving the learning objectives of the course, and that the instructional team is providing students with useful information on how well they are doing. ***Grading*** is the most obvious method of evaluation. However, ***feedback*** is another method of evaluation which can specifically target strength(s) and weakness(es).

This document is a living resource that ***guides faculty in best practices in the evaluation of student performance and attainment of the learning objectives of the course***, with an emphasis on undergraduate classes. Evaluation of student learning is a fundamental part of teaching. Instructors can spend a great deal of time scheduling and grading assignments, exams and other assessment tools. Those schedules can be rigid and not adaptable to unforeseen circumstances, from problems experienced by individual students to something affecting the entire class. This document provides guidance for establishing a suite of evaluations of student learning that are robust to unforeseen circumstances.

Note that a syllabus usually sets the social contract between the instructional team and the students by providing clarity of expectations regarding assignment and exam timing. Therefore, changing those evaluations within the quarter of teaching could be seen as breaking that contract. Nevertheless, we encourage all instructors to read through this document and assess the degree to which they can change their evaluation strategy to conform more closely to these best practices, *even if the syllabus and Canvas page are already published*. Whether the course is currently being taught, or an instructor is considering changes to a future offering, taking the time to speak to the students about any changes, including how those changes would help the students achieve course learning goals, is strongly advised.

Best Practices for All Classes: Establishing Multiple Approaches for Evaluation

Always make sure that courses have ***multiple individuals, and at least one other faculty member,*** with full access to all files and all grades (***assigned as a "teacher" in Canvas***). In cases where the lead/sole instructor is not a member of the regular faculty (e.g., postdoc, pre-doc instructor, part-time lecturer, or other guest instructor), the chair/director of the unit should have such access.

Piling all of the learning evaluation into the end of the quarter - such as a cumulative final exam - ***can be a recipe for failure***, especially if the expectation is that students will be able to apply their knowledge (i.e., demonstrate synthesis), rather than just regurgitate information (i.e., demonstrate memorization). Therefore, inserting smaller, “building block” assignments, or homeworks, or quizzes, all of which are designed to build and assess skill development, is key. In this type of restructuring, the final exam, or the final research paper, becomes the ultimate test of the skill(s) attained in the course, rather than the *only* demonstration possible.

Creating an evaluation plan in which there are ***multiple instances of each type of assessment***, allows students to suss out what is being asked of them, improve over time, and demonstrate mastery of learning objectives across a completed set of work. Examples include:

- daily prompted writing
- weekly quizzes or homework sets
- three exams rather than two, or one
- At the same time, the “multiple instances” approach allows instructors to:
- drop the lowest grade, accommodating students who may have started underprepared, or who have crises within the quarter
- create a portfolio score that documents improvement, rather than attainment of a pre-set threshold of knowledge/skill

In general, ***evaluating students "early and often,"*** as often as every class meeting, and through ***an integrated set of small, medium and large stakes assignments*** allows instructors to:

- keep track of the degree to which all students are on top of their work
- evaluate how students are improving their skills throughout the quarter, with the possibility of portfolio scoring (see below)
- provide regular (low stakes) feedback which helps keep students engaged

- have a running evaluation of student performance that is substantive enough to allow grading prior to the end of the quarter, if necessary
- accommodate absences while reducing requests to make up assignments (e.g., drop lowest grade from final calculation for that category of assessment)

FAQs Focused on Diversifying and Streamlining Evaluation

Is evaluation of every assignment necessary?

In short, no. Especially in cases of well-scaffolded assignments, assignments with particular structure tied to an evaluation rubric, and/or daily low-stakes assignments requiring students to learn and practice skills, ***students will build knowledge relevant to meeting learning objectives, even if all assignments aren't evaluated*** (i.e., practice makes perfect). Additional active learning work leading to the achievement of learning objectives without explicit evaluation, or even direct oversight by the instructional team, includes:

- student-led discussion sessions prompted by Canvas-posted prompts, with designated roles such as "facilitator," "synthesizer," and "scribe."
- credit/no-credit assignments (i.e., turned in, did not turn in) that are building blocks towards a larger project
- summary sheets
- reflective writing assignments

Is the instructional team solely responsible for evaluation?

Yes, and... Consider integrating ***peer-review*** using a clear and simple evaluation rubric that creates a structure for appropriate feedback (e.g., 1-2 specific things the evaluated student is doing well, 1-2 specific things that need work). While peer-review should not be used to create a grade, it may scale to a relative ranking of student work, and this may itself be used to guide evaluation. At the same time, enacting peer-review provides students with additional skills (e.g., positive critical review). Peer review itself might be evaluated.

Should all evaluation be student-specific?

Yes, and... Group assignments can achieve learning goals (e.g., teamwork) that individuals working alone can not. Consider using random assignment into groups, and varying their composition over the quarter. That said, group work should always:

- explicitly indicate multiple roles to achievement of the assignment (e.g., data acquisition, quantitative analysis, qualitative analysis, graphic design/representation, literature selection and annotated bibliography production, writing, presentation production, presentation delivery), with individual students assuming each role, and rubrics for success tied to task performance across all roles.

- have a clear role/task-based rubric also allows for self-evaluation, and peer-evaluation within the group, allowing the instructional team to better understand who did what without having to attend, for instance, all team meetings.

What if there's not enough time to teach and evaluate?

CoEnv is known for hands-on and field-based activities, but these take a lot of time, and may take instructional team members away from other duties. In cases of potential time management conflicts and where student-led learning activities are not possible, consider tapping other experts to help students achieve learning goals, including:

- other faculty, in or outside of the responsible unit
- members of the research staff
- local experts, especially in the case of place-based, and/or field-based learning

Allocation of others to some of these tasks can free up the instructional team for high-priority tasks, such as student evaluation.

What is the "stakes" approach to evaluation?

"***Small Stakes***" evaluations, such as daily writing, pre- and post-class Canvas quizzes, quizzes using Poll Everywhere, or active learning group work during the class can be used to ***build skills and confidence***; can be ***graded in a short amount of time*** (minutes per student, 2-3 day turn-around) and/or automatically (e.g., through Poll Everywhere quizzes, or Canvas quizzes). These assignments can also ***scaffold, or build to, high-stakes work***.

"***Medium Stakes***" evaluations, such as weekly problem sets, lab reports, drafts or research papers, or quizzes, can be placed throughout the quarter to ***test content knowledge and skills development*** and provide waypoints toward high-stakes work. These assignments can also ***scaffold, or build to, high-stakes work***.

"***High Stakes***" evaluations commonly include a ***mid-term and a final***. While these forms of evaluation are convenient, they may place too much emphasis on a single date, increase student stress, and be prone to failure if class is canceled and/or students need to miss exam days (see stress). Other types of high-stakes assignments include written work or presentations like a ***term paper or research paper***.

A stakes-based grading rubric should tie level (low, medium, high) to the degree of mastery of the skill or concept/content of the course, percent of the total grade, time spent by instructional team on grading and/or feedback, and scaffolding as follows:

	Small	Medium	High
Level of Mastery	awareness of skill/concept	demonstration of skill, application of skill of knowledge	advanced skills or synthesis demonstration
Percent of Grade (per assignment)	little-to-none*	5-10% per assignment	15-30% per assignment
Instructional Team Evaluation/Feedback Time	2-5 minutes per student assignment		can be significant for grading; as a "final" assignment, feedback should be minimal/none
Scaffolding	building blocks for high stakes assignments		

** but see portfolio scoring as an alternate approach to creating a larger fraction of the total course score attached to low stakes assignments*

What if high stakes assignments simply take too much time to evaluate?

High stakes exams can be scored efficiently using Canvas quizzes, Gradescope, or scantrons to automate some or all of the grading. ***Gradescope*** is an application best used to facilitate exam grading when questions are short answer, or other written forms requiring human intervention. This application will:

- compile students' answers to each question into a single grade space, so that they can be easily read and compared.
- group similar responses to speed up grading (e.g., all answers shown in a set receive 3 points).
- allow the instructor to enter words, phrases or coding syntax, with wildcard symbols and logic operators (e.g., phenotypic plasticity, marine protected area*, pyroclastic flow OR tephra) that Gradescope will search for within any question answer.
- automatically grade according to a rubric provided by the instructor (e.g., 3 points for the occurrence of a specified phrase).
- automatically regrade if that rubric is changed (e.g., changing a 3 point value to 5 points) or if a new search term is added.

High stakes assignments such as research papers necessitate significant time to read, process and evaluate. Consider moving written work, and especially work

demonstrating research analysis and synthesis, into a "***real-time evaluation framework***" such as a lightning presentation (1-3 minutes), or a poster, infographic or graphic abstract. Properly scaffolded (see below), and with a clear grading rubric, real-time evaluation can streamline final grading, moving evaluation effort on the part of the instructional team back towards small and medium stakes assignments that build towards the final. At the same time, real-time evaluation can facilitate peer-evaluation.

What is assignment scaffolding, and how does it support learning and efficient evaluation?

Large, complex assignments, such as a single exam, or a research paper, that are usually due towards the end of the quarter, can be difficult for some students to assimilate, both because of the stress inherent in any stand-alone high stakes assignment, as well as the tendency of many students to push off starting high stakes work (inadequate or unrealistic time management). Scaffolding the creation of a high stakes assignment (e.g., a final paper or presentation) by disarticulating both skills needed, and creation of component parts, can strengthen student learning through practice, force regular student work (through continuously due smaller pieces), and provide more opportunities for the instructional team to provide targeted feedback. Should a student not be able to complete the course, a set of scaffolded assignments can still be used to provide feedback in service of learning objectives, and in extraordinary cases, allow final grading even if the final large stakes assignment is missing, late, or incomplete.

Scaffolding Examples:

1. Group-based quarter-long research project:

	Week 2	Week 4	Week 6	Week 7	Week 10
	research objectives	analytical methods proposal	project data	draft report	final report and presentation
feedback:	X	X		X	
grading:	X	X	X		X
rubric:	Y	Y	N	Y	Y[2]
stakes:	L	M	L	M	H

2. Individual multi-week writing project:

	Week 1	Week 4	Week 6	Week 7	Week 8	Week 9	Week 10
	topic	bibliographic search results	annotated bibliography	outline	peer review draft	instructor review draft	final paper
feedback:			X	X	X[3]	X	
grading:	X	X	X	X		X	X
rubric[1]:	N	N	Y	Y		Y	Y[4]
stakes:	L	L	M	L	L	M	H

1. No means points for turning assignment in. Yes, means points allocated into specific categories tied to learning objectives.

2. Rubric carries over from draft.

3. Provided by student peers directly in Canvas, not instructional team

4. Rubric carries over from draft, with single addition re attention to feedback

What is portfolio scoring?

A portfolio is a mass of work (e.g., a series of small stakes writing assignments, quantitative analyses or homeworks), where students have had to perform tasks repeatedly, allowing them to practice then demonstrate skill mastery. A portfolio score uses a rubric to assess the ***degree to which each student improves their performance throughout the course regardless of their starting point***. In this way, a portfolio score rewards improvement (i.e., relative scoring) rather than absolute attainment of a specific level (aka "A work" or "B work"). Instructors can sample student work selectively without having to grade all assignments (e.g., at the beginning, middle and end of the course). Portfolio scoring also allows the instructional team to assign a score to any student who has completed multiple assignments, even if they stop early (e.g., leaving before the course concludes) or start late (e.g., add the course after the first week).

Portfolio Examples:

Autograded weekly homeworks (low stakes) in a beginning stats class:

Week:	1	2	3	4	5	6	7	8	9	10	Mean	Port.
	20%	20%	25%	40%	35%	50%	45%	60%	70%	70%	43%	+50%
	70%	80%	80%	90%	100%	80%	90%	70%	80%	90%	83%	+20%

While student 1 would fail the quiz set (43%), their level of improvement (+50% from the beginning two quizzes to the last two quizzes) is large. A portfolio score would reward that effort. Portfolio scores can be assessed over an entire quarter, or some shorter period of time. For instance, student 1 improved by 27% from the beginning two quizzes to the middle two (#5 and 6), and student 2 improved by 15%.

Daily in-class short writing assignments (low stakes) in a twice weekly course:

	Assignment:	1	2	10	11	16	17	Port.
Student 1	Structure	L	L	M		M	M/H	L to M/H
	Science	L/M	L	L/M		M	M	L to M
	Style	L	L/M	L/M		M/H	M/H	L to M/H
Student 2	Structure	M	M/H	M	M/H		M/H	M to M/H
	Science	M	M	M/H	M/H		M/H	M to M/H
	Style	M	M	M	M		M/H	M to M/H

Out of 17 daily writing assignments, portfolio scoring examines only six, two at the beginning of the quarter, two in the middle, and the final two. Of these, student 1 missed #11 while student 2 missed #16, which doesn't matter because multiple writing samples are assessed. Although student 2's final scores (M/H across the board) are higher than student 1's (M to M/H), student 1 started at a lower level (mostly L) and has therefore improved more over the quarter.