A3

Office of the President

TO MEMBERS OF THE ACADEMIC AND STUDENT AFFAIRS COMMITTEE:

DISCUSSION ITEM

For Meeting of March 16, 2022

INNOVATIONS IN ASSESSMENT AND GRADING AT THE UNIVERSITY OF CALIFORNIA

EXECUTIVE SUMMARY

The University of California is exploring innovations in instructional delivery, assessment, and grading to improve learning outcomes, promote academic integrity, and advance educational equity. Traditional grading practices, such as averaging grades across assignments or grading on a curve, have merit in some cases but there may be more effective options for advancing achievement and educational equity. UC campuses are engaged in a variety of initiatives to reexamine assessment and grading practices, some of which are illustrated in this written item.

BACKGROUND

Prior presentations to the Board of Regents have described ongoing UC efforts to assess, innovate, and adapt instructional methods to support learning outcomes. The University is following the same approach in evaluating assessment and grading methods to improve learning outcomes, promote academic integrity, and advance educational equity.

As part of UC 2030 goals, some UC campuses are looking at curricular innovations to help eliminate equity gaps. UC campus undergraduate deans, teaching and learning centers, and institutional research offices are creating dashboards for deans, chairs, and faculty highlighting courses with greater equity gaps. With this insight, these individuals are working together on ways to improve courses to address those gaps, with one key strategy to innovate assessment and grading.

Grading with equity

Faculty strive to be fair and eliminate bias from assessment, but standard grading practices may perpetuate bias. For example, averaging grades across multiple assignments may seem fair, except after considering differences in academic preparation. A student from an under-resourced high school may perform poorly on initial assignments but then master the final. That student will always have a lower grade overall because of their preparation starting point, compared to another student from a well-resourced high school that started strong and does as well on that final, but has never progressed in the course.

Other traditional assessment models, like the bell curve, distribute A through F grades to students at the top, in the middle, and at the bottom. The assumption is that students at the top of the curve have mastered the material, while students at the bottom may have not. However, as illustrated with the prior example, a student may have mastered the material by the end of the class but, due to low grades in the early or middle portion of the course, that student's final grade may put them in the middle or even potentially at the bottom of the curve. Moreover, the approach does less to show what students have learned via the course than it does in showing how students compare on the basis of preparation for the course.

During the pandemic, faculty sensitivity to inequity in educational backgrounds was heightened through observing differences in their students' educational opportunities and experiences, such as witnessing a student participating in the course from a closet because it was the only quiet space in the house or from a car because that student drove to a location with Wi-Fi. This experience has ramped up UC efforts to improve instruction, assessment, and grading.

Many campuses are examining assessment and grading practices that promote mastery learning as opposed to merely providing indices of whether or how much was learned. Mastery learning occurs when students have acquired sufficient practice, with instructor feedback, to master a skill or concept before being tested on it. Mastery learning maintains that students must achieve a level of mastery (e.g., 90percent on a knowledge test) in prerequisite knowledge before moving forward to learn subsequent information. So there may have been many tests in the process of achieving mastery, including formative practice tests or high stakes summative tests, where a student did not initially achieve mastery. The key difference is that the student gets repeated chances with different questions on the same topic and does not move on until that student does well enough to demonstrate understanding the concept.

While there is no clear consensus around the best approach, especially given that what is effective for one discipline may require a different approach in another discipline, emphasizing mastery learning and clearly articulating expectations at the beginning of the class are important and can contribute to better learning outcomes. This item will highlight some of the innovations being implemented on UC campuses to improve learning outcomes, emphasize mastery learning, and increase educational equity.

Revising assessment after remote instruction

Responding to concerns about the rapid shift to remote instruction and the pandemic, almost all UC campuses relaxed pass/no pass regulations. This included allowing classes that normally do not allow pass/no pass grading to implement it, as well as making the deadlines to change grading options more flexible. Some campuses are examining whether there are circumstances where continuing these practices might support first-year retention and student success.

In 2020–21, MIT implemented a new "Flexible Pass/No Record" grading policy[1] for entering first-year undergraduates. This policy allowed first-year students the option to designate up to 48 units to be graded on a pass/no record basis. Under this policy, students needed to complete the subject and receive a grade. The student then had the option to keep the grade or request that it be converted to a pass/no record. Letter grades of A, B, or C were pass and D or F were no record. The policy was intended to reduce pressure in the first year of study while allowing for more major exploration. Initial feedback was the new policy had been well-received by students, with 78 percent feeling "very positive" about this new grading policy.

Johns Hopkins has also revised its Freshman First-Semester Grading Policy[2], in which letter grades earned in the first semester are not reported on the transcript and instead are converted to satisfactory (for C- or higher), unsatisfactory with credit (D+ or D), and unsatisfactory with no credit (F). No first semester grades are used to calculate a student's cumulative grade point average, but they are provided to advising offices and faculty advisors to ensure students are making satisfactory academic progress. Students can also request a release of covered grades if needed (e.g., scholarship eligibility).

UC campuses are in various stages of evaluating pass/no pass policies and first year grading. For example, UC Irvine's Division of the Academic Senate has extended the deadline to choose pass/no pass grading to the tenth week of instruction. They have also met with Massachusetts Institute of Technology (MIT) colleagues and are evaluating options that might make sense, recognizing the differences between MIT and UC Irvine. For example, MIT has a much more uniform first-year curriculum.

UC Berkeley's College of Chemistry is in the planning stages around pass/no pass grading for first-year students. Two main motivations for considering a change are (1) seeing how using the pass/no pass option over the last three semesters helped relieve stress on students and (2) believing there is an equity issue for relatively underprepared students who enter the degree program who generally finish the year with lower GPAs than better-prepared peers. Options could include switching a class grade in the first semester or year to pass/no pass, with the College currently assessing the workload impact needed on college advisors and the registrar.

When classes moved to remote instruction, some UC campuses began using oral exams as a way to address academic integrity concerns and it may also be effective in advancing educational equity. UC San Diego's Teaching + Learning Commons Engaged Teaching Hub staff and engineering faculty are collaborating on a National Science Foundation–related grant for oral exams. Both instructor-to-student interaction and opportunity to give real-time feedback and coaching to students have been important benefits. The research team also found that faculty had greater empathy and a deeper understanding of student learning, where support was needed, and how the process improved student perceptions about an instructor's approachability.

---

[1] "Flexible P/NR Grading Option," MIT Registrar's Office, accessed on February 13, 2022, https://registrar.mit.edu/classes-grades-evaluations/grades/grading-policies/flexible-pnr-grading-option
[2] "Undergraduate Student Handbook Grading Policies," Johns Hopkins University, accessed on February 13, 2022, https://pages.jh.edu/design/oliver/academic_manual2010-2011/grades.html

While oral exams are not new in higher education, the context is unique. In this case, oral exams are being used in high enrollment, introductory engineering courses. This approach is an underutilized assessment strategy in science, technology, engineering, and mathematics (STEM) fields, despite a mounting body of research identifying myriad benefits derivable from the pedagogical practice. These include greater student effort and intellectual engagement, increased motivation to learn at a deeper level, improved technical speaking competence, improved assessment power, and security against cheating.[3]

UC Davis's mathematics faculty also used oral exams for upper division courses with enrollments between 50 and 80. They found this approach to assess students more authentic. Students also expressed appreciation for this approach, including a student that said "the final exam showed me I am capable of both understanding and demonstrating my understanding. It made me a lot more excited about math and the opportunities that it has to offer. I am usually someone who gets testing anxiety and I never end up showing my true knowledge on exams because it gets in the way. This exam allowed me to show my knowledge to my full potential."

Choosing and adapting assessments

UC campus teaching and learning centers are at the center of supporting curricular reform, innovations in assessment, and educational equity. The centers supported faculty with the shift to remote instruction, leveraging research and resources on effective teaching practices and strategies. They partnered with faculty and administrators to conduct and collect research on best programs and promising strategies within and outside the UC system.

Teaching and learning centers provide extensive resources on ways to ensure classrooms are equitable and inclusive spaces and that assessment practices support student learning and help instructors gauge understanding, provide feedback, and assign grades. UC Santa Cruz provided the following tips to consider when determining the right assessment method[4].

[3] M. Lubarda *et al*., "Oral exams for large-enrollment engineering courses to promote academic integrity and student engagement during remote instruction," *2021 IEEE Frontiers in Education Conference (FIE)*, 2021, pp. 1-5, doi: 10.1109/FIE49875.2021.9637124.

[4] "Assessing Student Learning – Keep Teaching," UC Santa Cruz, accessed on February 12, 2022, https://keepteaching.ucsc.edu/assessment

KEY TIPS

- Think of **assessment as opportunities for students to learn** with **equity and inclusion** as the foundation to create effective learning spaces for all students.
- To determine the most appropriate types of assessment to use, consider the **purpose of assessment** and the **learning goals of the course**.
- Allow students to **demonstrate learning in a variety of ways** by incorporating **formative assessments**, ample **feedback**, **alternative summative assessments**, **peer review**, and many **opportunities to reflect, correct, and revise**.
- Focus on **higher order skills** of analysis, critical thinking, problem solving and communication of thought process and solutions, as opposed to factual recall.
- Spread assessment throughout the term, rather than relying on a small number of high-stakes exams or final papers.
- Provide **opportunities for and encourage collaborative work**.
- Allow (or require) students to **consult many sources to answer questions**, as we do in the real world
- Talk to your students early and often about **academic integrity**.
- Consider the **grading labor** for each assessment.

UCLA's Online Teaching and Learning and Center for Education Innovation and Learning in the Sciences collaborated with HumTech, Center for Advancement of Teaching and SEASNet to produce an Alternative + Remote Assessments guide. The guide includes the following summary table of assessment adaptations with guidance on how to increase equity and effectiveness and reduce the workload associated with grading.

Summary Table of Assessment Adaptations

Alternatives	Increase Effectiveness	Increase Equity	Critical Equity Considerations	Reduce grading
Open book/ open note/ take home exams	Ask students to provide explanations, interpretations or justifications Use TurnItIn to ensure they use own words	Ask questions that have students make connections to their own experiences	Offer at least 24 hrs to complete (CCLE can make 3-hr exam within a 24-hr range) Provide example problems before the exam to help students build these skills	Use Gradescope to streamline grading *(and to make grading more transparent and therefore more equitable)*
Allow Collaboration on Exam	**Try two-stage exam:** group portion follows individual portion	Turns exam into learning and community-building experience	Assign groups b/c marginalized students are often excluded Provide other opportunities (e.g. discussion section) for groups to work together	Make second stage multiple choice and/or just one combined group exam
Weekly Quizzes	Shuffle questions and responses Add a question at end where students need to explain reasoning for each answer	Consider allowing multiple attempts (at least in the beginning to help students get used to expectations)	If only administered at specific and limited times, use anonymous survey to assess access issues	Allow students to drop lowest 10-20%, no questions asked (also accommodates students with caregiving/work duties)
Projects, Case Studies, Papers and Present-ations	Provide meaningful options: real life case studies, briefing, annotated research bibliography, fact sheet, Op-Ed piece, student-proposed project	Connects to student interests and builds useful skills Use clear rubrics to clarify expectations of students	For larger papers and projects, collect drafts or sections over time to structure progress Synchronous presentations would be challenging if internet access limited If group presentations, assign groups based on student schedules or topics of interest	Use CCLE Workshop tool to allow peer- and self-grading (for drafts and/or final submission) Provide well-structured group projects, including clear milestones, reflection on group functioning and individual accountability
Foster Metacog-nition (e.g. Exam Wrapper)	Students complete "exam wrapper" after first exam to reflect on what to improve. And/or have them read an article on how to learn, and reflect/discuss	This supports the many students who are still learning HOW to learn, especially remotely	Assign (small number of) points so that students with caregiving/work commitments still make time for this. Consider allowing students to get back fraction of lost exam points by submitting corrections and reasoning	Just grade for completion and indicate you will spot-check
TOOL: Gradescope Rubric-based online grading platform that allows TAs to easily click through every exam question they are grading Reduces bias by hiding names Students can easily see which rubric points were gained and lost	**Grading Schemes to Consider:** 1. Straight scale so students always know where they stand 2. "Specifications grading" where all items are graded as satisfactory/unsatisfactory and grade is based on # of assignments completed satisfactorily 3. Ungrading: Eliminate points to foster intrinsic motivation			

A number of these practices are being used at other UC campuses, including:

- Two-stage exams: UC Davis's Introductory Biology lecture course uses a standard individual exam, followed by a collaborative group exam where students work together to solve the same or related questions. Students who preferred that approach said it

provided an opportunity to debate and arrive at a better answer. Students also received immediate feedback on individual exam responses from peers.

- Exam wrappers (i.e., process to review performance on exam to adapt future learning): UC Santa Cruz leveraged its learning management system to carefully monitor students who struggled early in the quarter to target interventions in areas where the student was struggling. Many instructors let students retake quizzes or provided structured opportunities to go back over the exam to determine what was incorrect and why.
- Gradescope: UC Riverside is making this online grading platform available for all instructors as part of the campus's educational technology portfolio. As noted in the UCLA guide, this approach not only reduces workload, it also reduces bias by hiding student names and provides timely feedback to students on where rubric points were gained or lost.

Formative and summative assessments

UC Merced's Teaching Commons provides instructors a series of student-centered pedagogy guides on course design and frameworks, assessments and feedback, and engaged learning experiences.

It makes a distinction between formative assessment that diagnoses, monitors, and directs student learning and summative assessment that evaluates overall learning.

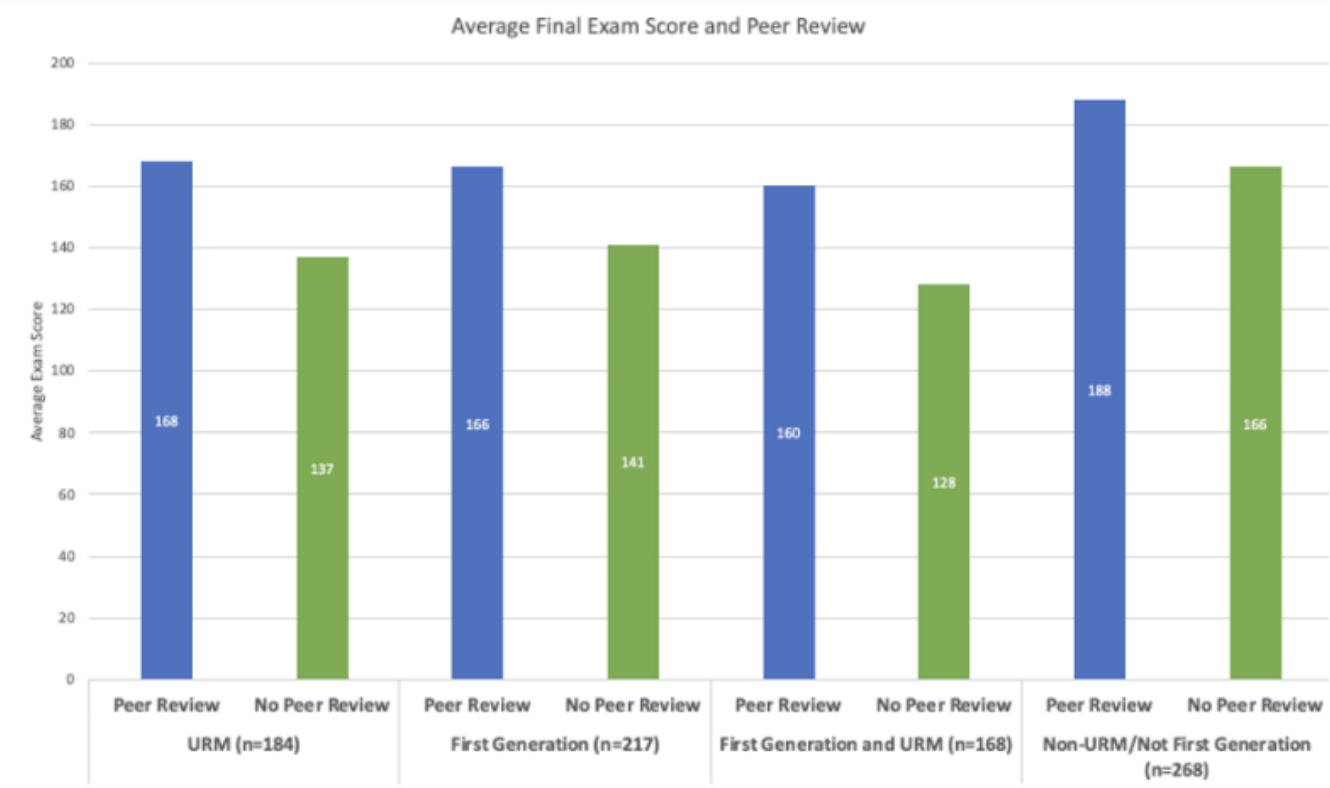
Formative strategies can include things like in-class learning activities, brief reflections, low-stakes weekly quizzes, open book exams, or small group assignment. Summative strategies grade student performance in relation to learning outcomes and can include mid-term exams and finals, special performances, or open-ended, complex projects or design challenges.

Type of Assessment		Purpose or Use of Data is to:	Best Serves
Formative	**"FOR"** Learning	To **diagnose** needs **or monitor** student progress and inform upcoming instructional planning/decisions.	Instructor
	"AS" Learning	To **enhance** student development as self-directed learners, focus learners, and help students identify effective study strategies.	Students
Summative	**"OF"** Learning	To **evaluate/judge** a) student performance on CLOs and/or lesson objectives, and b) the effectiveness of instructional materials and activities.	Instructor

UC Santa Barbara's Writing Program uses both formative and summative strategies to support students. The Writing Program uses portfolio grading where students write formative drafts, followed by longer semi-summative ones that the student can revise for a portfolio which constitutes a large portion of the grade. Many UC writing programs follow this model.

To support the workload associated with this approach, UC Santa Barbara leverages technology and peer-reviewed writing assignments, sometimes known as "writing to learn." Instructors use a peer review and feedback platform to implement writing activities in class. Within its Introductory Biology course, there can be between 240 to 800 students per section. Faculty used peer review and feedback platform to develop writing assignments focused on difficult course concepts. After completing the assignments, students then use the platform to review the work of other students.

Research shows that each stage of this process—writing, receiving feedback, and providing feedback—reduces equity gaps. To date, more than 6,000 students have completed writing/review activities using learning review, showing promising results with gaps in average final exam scores decreasing for underrepresented and first-generation students who completed writing and peer review assessments using the peer review and feedback platform.

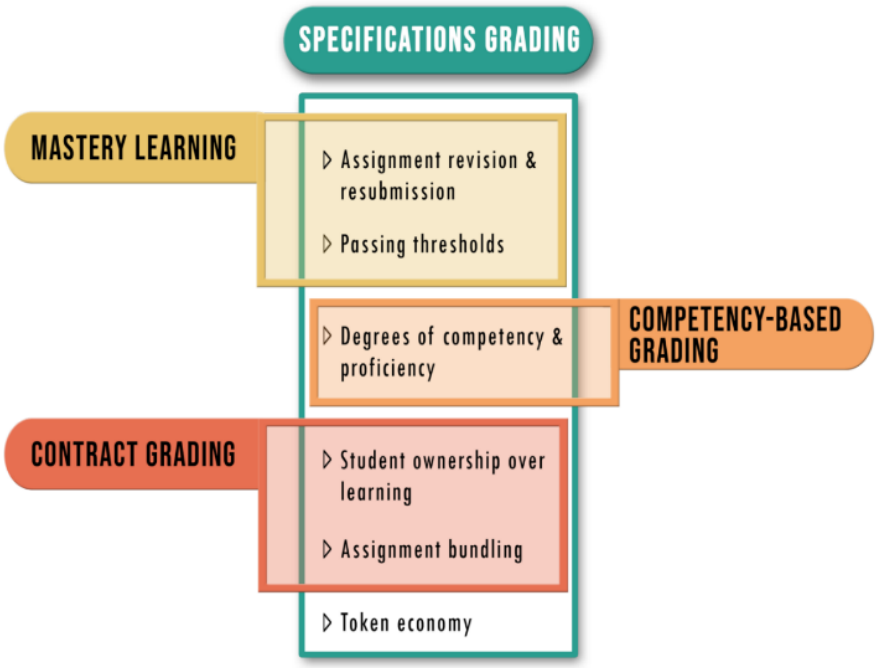


Center for Innovative Teaching, Research, and Learning experts worked closely with faculty at every stage of the process—creating assignments, including identifying key course concepts, designing short writing activities, creating structured peer review rubrics, and using data gathered from students' writing/review to improve instruction.

Contract and Specifications Grading

The UC Davis Department of Mathematics is examining contract grading where a student can choose how to be graded. Within its calculus course, one instructor provides three different grading options for distributions of grades across different types of assessments. Students can choose which option at week four and nine of the quarter. One sample grading option could be four short tests (40 percent), homework (15 percent), group work (20 percent), class engagement (10 percent), and final exam (15 percent). When surveyed, almost nine of ten students preferred the ability to choose a grading scheme.

In some STEM courses in the past, some instructors have begun their courses by saying something that may be familiar to many, "look to your left, look to your right—two of you aren't going to pass." Instructors within UC Irvine's Organic Chemistry class, however, have been adapting grading strategies with a goal that everyone could not only pass, but get an A. These instructors are currently using specifications grading[5] which builds on mastery learning, along with competency-based and contract grading. While each has value, some can be harder to use in large courses (e.g., those with more than 1,000 students enrolled).



Instructors provide very clear, detailed specifications on what it takes to get certain grades. If manageable, students can select assignment bundles to complete as in contract grading; otherwise they are set in advance by the instructor. Aspects of mastery learning and competency-based grading can also apply, where students can have limited options to correct and resubmit work that does not meet the "specs," or specifications, set to reach a satisfactory level. This

[5] Julia Winter, "Specifications Grading at UC Irvine with Renee Link," Ideas that Matter, accessed on February 12, 2022, https://www.alchem.ie/blog/specifications-grading-at-uc-irvine-with-renee-link

approach helps foster an equitable classroom by leveling the playing field for students who may have come in with insufficient preparation when compared to their peers.

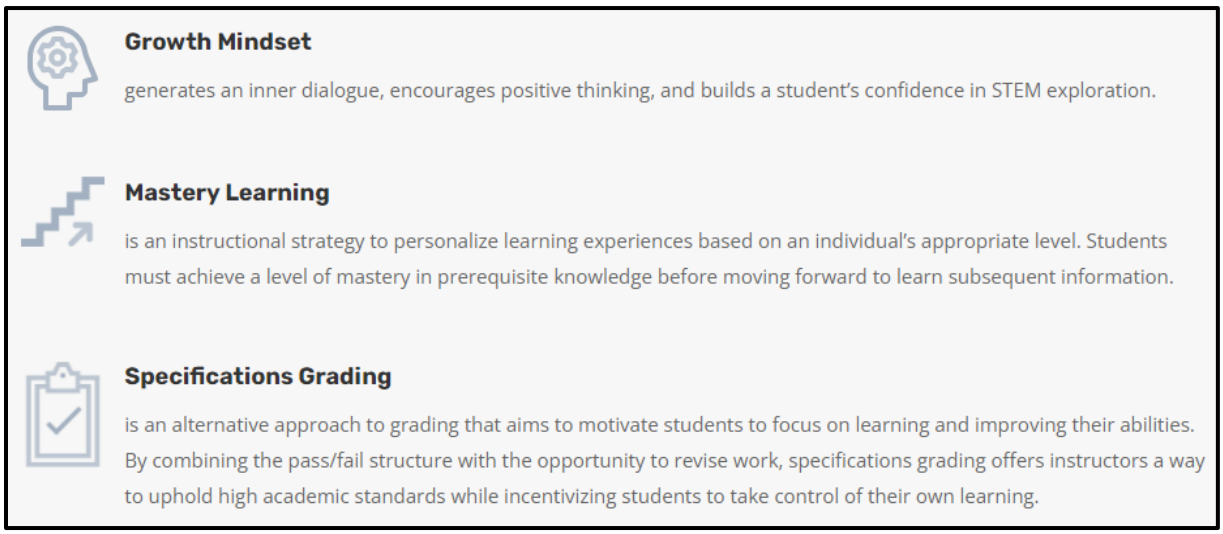
But specifications grading is not limited to STEM disciplines. UCLA's School of Theater, Film and Television uses it in its large General Education Drama and Diversity course. Most of the student's work is graded pass/no pass on the basis of detailed specifications provided in advance. A student has ownership over learning, selecting which assignments to do within certain parameters and which to have graded, then putting in the amount of effort required to reach that grade. Students are encouraged to bring in relevant materials from current events, new plays, or personal interest to discuss with classmates. For every satisfactory participation in class (i.e., work that is "up to spec"), a student earns between one and three points depending on the activity and when the student accumulates 63 points, that student earns an A.

UC Riverside is also using specifications grading in writing courses and teaching statistics in psychology.

Intersegmental efforts

UC campuses are also involved in intersegmental efforts examining innovations in grading and assessment. The following two examples are funded in part by the California Education Learning Lab.

UC Irvine, California State Polytechnic University, Pomona, and San Jose State University have created a cohort-based faculty development team, the Teaching Experiment Academy (TEA), looking to redesign STEM curriculum with mastery learning structure and a specifications grading approach to enhance learning and foster a student's growth mindset.



UC Berkeley, California State University (CSU), Long Beach, and El Camino College have launched a collaborative pilot project to improve student equity and diversity and institutional resilience in computer science through a mastery learning project called "Reorienting Formative

and Summative Assessment Towards Mastery Learning for Learner Success, Student Equity, and Institutional Resilience."

This project adapts the University of Illinois at Urbana Champaign's PrarieLearn platform is open-source software that supports question generation and grading for mastery learning in computer-based assessments. Researchers are looking at the impact of using proficiency-based learning methods, especially paradigm-based question generators (PQGs) where instructors can create problem types (e.g., circuit diagrams, algebra problems) and PRQs can generate endless variants to those problems that can encourage and support "mastery" over foundational computer science concepts tailored to each student versus fast-paced learning and evaluation.

Using PQGs, students can practice as much as necessary. The same system can be used to deliver homework or exams. Because exams are complex to prepare and administer, instructors usually create few of them, so a small number of summative assessments often determine a large part of the student's final grade. However, in this model, by using software, instructors can generate random questions for quizzes, tests, or homework, with relative ease.

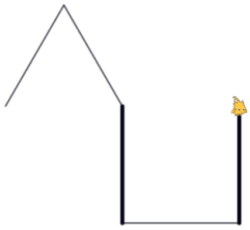
For example, in an exam in an Intro to Computer Science class that was offered in fall 2020 at UC Berkeley, virtually every student was presented with a different example. Below, Figure 1 and Figure 2 show examples of two questions that could be generated for two different students. Given the picture, the students are being asked to author the computer code to draw the fractal shown (this example is from a UC Berkeley's computer science class, CS10).

Figure 1

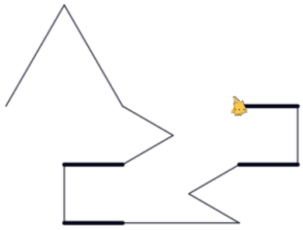
Though this may look daunting at first, it isn't that bad. Remember, every fractal has a base case (**n = 1**) and recursive case. We've drawn the fractal with **bold** lines to indicate the parts of the drawing that recurse; the other parts of the drawing at **n = 2** are just lines. (You don't have to copy our bold/normal style, it's ok if it's all the same.) Look at how the straight line at **n = 1** transforms into the **n = 2** case – this happens for every bold line at the next level. The *five* n = 2 lines are *1/2 the length of* the **n = 1** line, *and the sprite ends facing the same way it began*. For **n = 2**, you can assume the "triangle" formed by the 2 lines is an equilateral triangle, and the "square" formed by the 3 lines has 90 degree angles.

You can click on `Save & Grade` below, since you need to upload your solution to Gradescope.

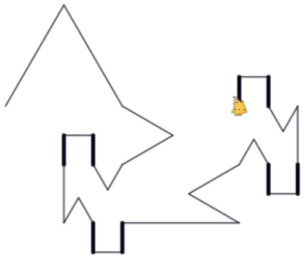
n = 1



n = 2



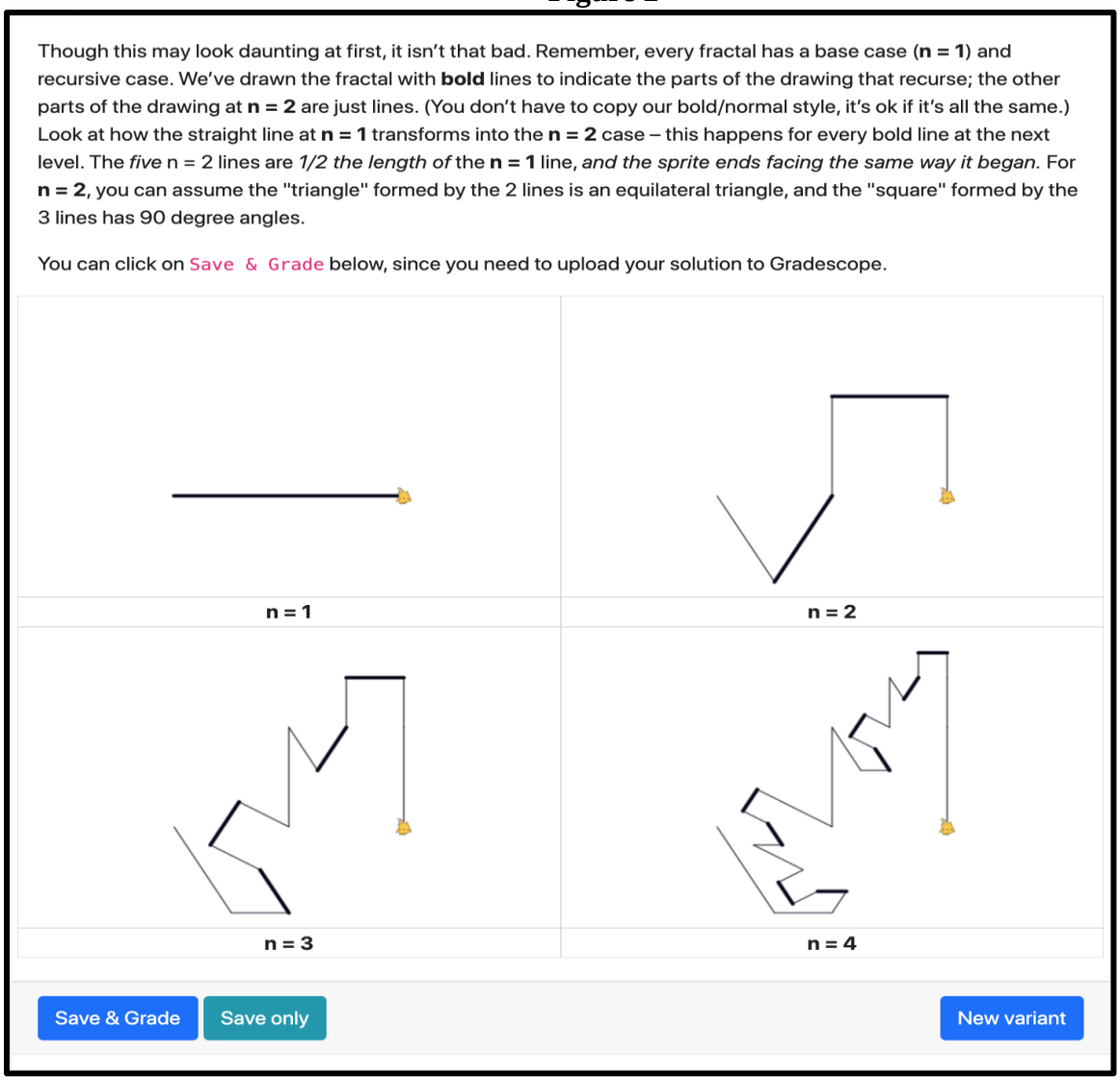
n = 3



n = 4

Save & Grade | Save only | New variant

Figure 2



Each student receives a different problem but all the problems are equally difficult. Since the questions and solutions are all different, there is also the added benefit that cheating is not an option, which contributes to institutional resilience. Through its collaborative project, these questions can also be shared across institutions.

The project's hypotheses are that mastery learning using PQGs will contribute to higher retention, stronger learning outcomes, more effective use of instructor time, and an increase in successful participation in computing for traditionally underrepresented students. The hope is that the findings of the project will open the door towards new learning approaches to expand diversity and inclusion in computer science, with a goal the pilot project can become a permanent center for promoting mastery learning on these and other California campuses.

Training and development

As the University of California continues to learn and collect best practices, campus teaching and learning centers are hosting discussion series, providing training, and supporting departments to further advance innovations in grading and assessment.

For example, UC Merced hosted an "Anti-Racist Pedagogy Faculty Discussion Series – Impact on Grading and Assessment Practices" that focused on teaching methods employed by instructors that disrupt the traditional pedagogical practices that reinforce structural racism. Anti-racist pedagogy builds on inclusive pedagogy practices to equip learners with the tools necessary to identify and dismantle systems of injustice wherever they are encountered. Anti-racist pedagogy practitioners flatten classroom hierarchy, engage students in active problem-based learning, and intentionally design experiences, grading practices, and assessments that honor student voice and agency. As one outcome from the conclusion of this series, participants will have considered student-centered language choices for syllabi, types of assignments that develop critical consciousness, and asset-focused grading/assessment approaches that together foster a sense of invitation and empowerment.

UC Berkeley's EMPOWERing Engineers for Positive Change program also delivered an interactive three-workshop module on Equitable Grading Strategies that focused on best practices to support student learning and development and promote equity in instruction and grading. These workshops covered:

- *What is in a grade?* Participants investigated the role and value of grades in students' learning processes and experiences, examined ways in which grading practices can exacerbate inequities, and explored "alternative" grading strategies and related research on student learning.

- *What approaches can I use for equitable grading**?*** Participants compared and contrasted equity-minded grading frameworks, including Grading for Equity and Contract Grading, and identified an opportunity to adjust a method of assessment in their teaching contexts.

- *How can I apply equitable grading strategies in my teaching practice?* Participants built on prior conversations to reflect equitable grading practices and collaboratively workshop and develop strategies to determine what to implement in their teaching contexts.

CONCLUSION

Assessment and grading are critical for students to evaluate and advance learning and for instructors to adapt and improve instruction. Traditional assessment and grading practices may perpetuate bias and inequities and the University of California is engaged in a number of efforts to advance initiatives that promote grading with equity, including those that improve and reward mastery of subject matter in a course. While there is no clear consensus around the best approach and recognition of disciplinary differences, UC's efforts and partnerships with other

intersegmental and peer institutions will advance research and identify promising practices that can continue to improve student outcomes and educational equity.

Key to Acronyms

PQGs	Paradigm-based question generators
STEM	Science, Technology, Engineering and Math
TEA	Teaching Experiment Academy