Developing and Implementing a Specifications Grading System in an Organic Chemistry Laboratory Course

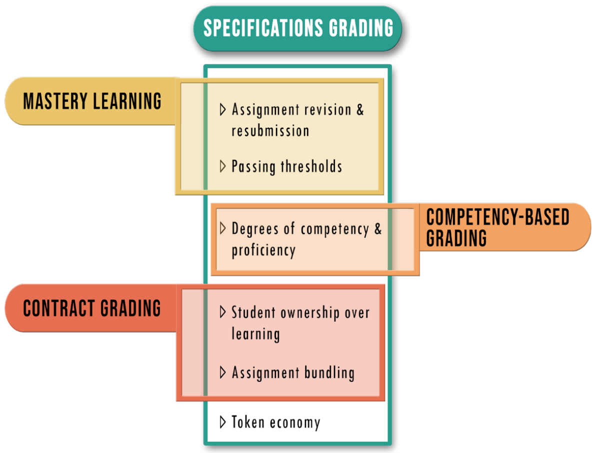
William J. Howitz,‡ Kate J. McKnelly,‡ & Renée D. Link*

Department of Chemistry, University of California—Irvine, Irvine, CA 92697, United States

ABSTRACT

Large, multi-section laboratory courses are particularly challenging when managing grading with as many as 35 teaching assistants (TAs). Traditional grading systems using point-based rubrics lead to significant variations in how individual TAs grade, which necessitates the use of curving across laboratory sections. Final grade uncertainty perpetuates student anxieties and disincentivizes a collaborative learning environment, so we adopted an alternative grading system, called specifications grading. In this system each student knows exactly what level of proficiency they must demonstrate to earn their desired course grade. Higher grades require demonstrating mastery of skills and content at defined higher levels. Each students' grade is solely dependent on the work they produce rather than the performance of other students. We piloted specifications grading in the smaller, third quarter course of the lower division organic chemistry laboratory series held during a summer term. Open-ended questions were chosen to gather student and TA perceptions of the new grading system. TAs felt that the new grading system reduced the weekly grading time because it was less ambiguous. Responses from students about the nature of the grading system were mixed. Their perceptions indicate that initial buy-in and multiple reminders about the bigger picture of the grading system will be essential to the success of this grading system on a larger scale.

GRAPHICAL ABSTRACT



KEYWORDS

First-Year Undergraduate, Second-Year Undergraduate, Upper-Division Undergraduate, Organic Chemistry, Laboratory Instruction, Curriculum, Collaborative/Cooperative Learning, Testing/Assessment

INTRODUCTION

Grading, a fundamental component of assessment in higher education, is intended to reflect student achievement of course learning outcomes. Finding an objective way to assess qualitative work is challenging and has depended traditionally on points-based grading systems.[1,2] This approach to grading in the college classroom is not ideal as it places emphasis on the extrinsic motivational factor of accumulating points rather than the intrinsic motivation of learning and meeting course learning outcomes.[3–5] The education community has demonstrated awareness of flaws in the traditional, points-based, grading systems it employs, as evidenced by the continuous development of methods to improve the grading process.[1,2,6–9] Specifications grading — popularized by Linda Nilson in 2014 — represents a new grading system that moves away from a reliance on points and has the potential to make substantial positive changes in student learning.[10] In this paper we discuss the origins of the specifications grading system, outline the potential benefits of adopting it for large university science, technology, engineering, and mathematics (STEM) programs, and describe what we believe to be the first implementation and the outcomes of this grading system in an organic chemistry laboratory course.

Evolution of Alternative Grading Systems

Specifications grading evolved from three previous grading systems: mastery learning, competency-based grading, and contract grading, and has been adopted in a variety of college-level courses.[10–33] Mastery learning approaches, such as Bloom's Learning for Mastery and Fred Keller's Personalized System of Instruction, require that students meet an instructor's established performance standards on one course topic before advancing to subsequent topics.[34,35] These approaches advocate for variation in teaching methods and flexibility in time allotted for students to complete course topics to better meet students' individual learning needs. While mastery learning approaches are effective, the challenges of providing individualized instructional strategies and having sufficient time to ensure all students achieve the same level of learning make the mastery learning approach daunting for instructors to implement.[36]

Technological advancements, such as the internet, enabled the development of competency-based grading — an extension of mastery learning. Competency-based grading similarly uses instructor-defined passing thresholds on assessments, but these thresholds are differentiated into multiple categories based on a student's level of competency. This approach empowers students to take greater control over their own learning by providing them the option to demonstrate proficiency on an assessment above the minimum level.[11,37–40] The use of technology in this grading approach allows for personalized and immediate feedback as students move through course material at their own pace. However, technology is not required in a competency-based grading approach, and in a technology-driven version of this approach, care must be taken to design the course without unintentionally de-emphasizing student engagement with instructors and peers as this is important for student retention of course material.[41–44]

Contract-based grading gives even greater control to students over their learning than competency-based grading. In this system, students negotiate a contract with the instructor to define which assignments they want to complete for a predetermined grade in the course.[45,46] If students meet the level of performance expected, the instructor awards the student the predetermined letter grade. This system retains the student-instructor and student-student engagement that may be lost in competency-based approaches, while also giving students more ownership over their learning. This ownership is valuable because it increases student motivation to learn the course material.[47,48] This system also has the added benefit of eliminating competition between students as each student's grade is independent of their peers' grades. However, this method of grading has been criticized for potentially allowing students to easily earn higher grades while putting in less effort than a traditional grading system.[10,49] Another drawback — similar to the original version of mastery learning — is that most contract-based grading requires extensive amounts of instructor time because each student has to develop their own contract which is then instructor-approved.

The benefits and drawbacks of each of these three grading systems informed the design and development of the specifications grading system.[10] To keep course workload manageable for instructors in specifications grading, the instructor, rather than the student, defines the contract options that are tied to specific letter grades. Students still retain a degree of ownership over their learning by choosing to complete the bundle of assignments — that is, the contract — for the letter grade they want to earn in the course. The instructor defines passing thresholds for each assignment in the contract that students must meet to achieve proficiency, which ties back to the core idea of competency-based education. Because each bundle of assignments is developed with the course learning outcomes in mind, the learning outcomes a student has met will be evident based on their satisfactory completion of the associated bundle. In addition, the specifications grading system includes a token system, which provides students with limited options to revise and resubmit work that does not meet the criteria set to reach a satisfactory level. Limiting options for resubmitting work is necessary to keep the time needed for grading manageable. The token system incorporates a mastery learning element and gives students increased ownership over their learning in the specifications grading system, as students can choose which work they will revise and resubmit in exchange for using a token.

Grading Challenges in Large, Multi-section Laboratory Courses

The high-enrollment, multi-section laboratory courses that predominate in most large college and university STEM programs present particular grading challenges. The traditional, points-based systems typically used in these courses do not always accurately reflect student achievement of course learning outcomes. In addition, the necessity for multiple graders, grade standardization, and curving leaves students confused about their standing in the course and in competition with their peers. Specifications grading can minimize, and potentially resolve, these issues.

Grading should evaluate student success based on achievement of competency in one or more course learning outcomes. Under traditional, points-based, grading systems, this is not always the case. If points are not clearly allotted for specific course learning outcomes, students may earn enough cumulative points to pass the course without clearly meeting any of the course learning outcomes.[10,12,13,23] The structure of specifications grading resolves this issue because each bundle of assignments tied to the final course grade is developed with the course learning outcomes in mind. Therefore, a student demonstrates competency in course learning outcomes by satisfactorily completing the associated assignment bundle. Another drawback of the points-based system is the focus students place on receiving points rather than meeting course learning outcomes. Students may view points as a transaction for effort put into work submitted rather than credit to be earned for demonstration of understanding course material. By removing the distribution of points on assignments and using a binary satisfactory/unsatisfactory approach, specifications grading shifts student focus to understanding course concepts and demonstrating skills.[10,13,25]

In a points-based system, students are generally unsure of their final course grade because they are unable to anticipate how the score standardization and course curve will change their unstandardized scores. Variation in TA grading requires final grades to be normalized and curved because each laboratory section can have drastically different section averages.[25,50] Without standardization, students with less critical TAs would be rewarded with higher course letter grades while students with more critical TAs would be punished with lower course letter grades. This uncertainty in grade standing not only contributes to student anxiety, but is also contrary to a cooperative and collaborative learning environment because this grading system perpetuates a student culture of competition.[51–53] Each student feels that they are competing against other students for each point so they can have a higher point total at the end of the course. The higher their point total, the better their chance of benefiting from the curve when final letter grades are determined. Under a specifications grading system, the need for standardization and a curve is eliminated. Each student knows exactly what they must accomplish to earn their desired course grade, and each student's grade is solely dependent on the work they produce, rather than being dependent on the performance of other students.

At the University of California, Irvine, our courses include the issues described above. As many as 35 teaching assistants (TAs) are responsible for grading assignments from over 1,000 students spread across 60 or more organic chemistry laboratory sections for a single course. Because grading systems using point-based rubrics can lead to significant variations in how individual TAs grade students' work, the scale of our courses at UCI requires final grade standardizations to account for the large number of TAs and associated laboratory sections. In our experience we have found that TAs generally agree on the quality of student work, but not the point values they assign as partial credit. Students also disagree with TAs about the number of partial credit points they are awarded per assignment. Removing points from the grading system could reduce these grading inconsistencies because specifications grading uses a binary satisfactory/unsatisfactory approach. TAs only need to identify one threshold per rubric item as opposed to the spectrum of thresholds contained within point-based rubrics. The specifications grading system could also have the advantage of reducing time spent on grading because less grader energy is put into deciding between a satisfactory or unsatisfactory assessment compared to having to select a score along a spectrum.

DESIGNING A SCALABLE SPECIFICATIONS GRADING SYSTEM FOR A LABORATORY COURSE

Specifications grading has been used in various STEM courses, including chemistry lecture courses, but has not yet been reported in a chemistry laboratory-only course.[15–22,24–26,28,30,33] With an end-goal of scaling up specifications grading to our larger, 1,000 plus student, on-sequence courses, we chose to pilot a specifications grading system in the final course in the organic chemistry laboratory sequence. We specifically chose to pilot specifications grading in the accelerated summer session course because it has the smallest enrollment — about 40 students. Each week in this course, students attend two 50-minute laboratory lectures taught by the instructor and two four-hour laboratory sections taught by a graduate student TA.

To transition the organic chemistry laboratory course grading system, we began by defining criteria students must meet to achieve specific grade levels: A, B, C, D, or F (Table 1).[54] These criteria were designed to reflect the Student Learning Outcomes (SLO's) for the course and encompassed all previously graded components of the course: online pre-laboratory homework, pre-laboratory video quizzes, laboratory notebook assignments, post-laboratory assignments, laboratory lecture participation, and practical exams. Students were given a Student Grade Tracker as a checklist tool to track progress towards earning their desired grade (shown in the Supporting Information). Rubrics for course laboratory notebook and post-laboratory assignments were adjusted to a binary satisfactory/unsatisfactory form, consistent with the specifications grading system. To incorporate a mastery learning aspect into the system, we also instituted a token system where students could redeem a token for the opportunity to resubmit an assignment that was assessed as unsatisfactory. We also divided the practical exam into components and specified which components students needed to complete to earn their desired letter grade.

Course letter grade bundles were defined and included on the Student Grade Tracker, and students earned the highest grade for which they met all of the criteria. For example, to earn a C-level grade, a student must have achieved at least 70% of the points for the online pre-laboratory homework, 75% of the points for the pre-laboratory video quizzes, five or more satisfactory laboratory notebook assignments, three or more satisfactory post-laboratory assignments, have attended four or more laboratory lectures, and have passed the required practical exam components. The course letter grade bundles were designed to align with the following course SLO's:

1. Perform fundamental organic chemistry techniques in the context of laboratory experiments.
2. Demonstrate understanding of concepts underlying fundamental techniques by proposing solutions to actual or potential problems encountered during an experiment.
3. Accurately draw reaction mechanisms for reactions conducted in laboratory sessions.
4. Use spectroscopy data to determine structures of unknown molecules.
5. Use data collected from an experiment to make claims supported by evidence.
6. Identify safe and unsafe practices related to techniques used in laboratory sessions.

Students could earn higher grades by achieving requirements for higher grade bundles, and ultimately would earn the highest grade for which they met all of the criteria in a given bundle. Higher grade bundles required higher levels of performance as demonstrated through higher percentages on homework/video quizzes, completing more laboratory assignments as satisfactory, and passing additional exam components. For example, to earn an A-level grade a student would need to: earn 90-

100% on online pre-laboratory homework, earn 90-100% on pre-laboratory quizzes, earn a satisfactory assessment for seven laboratory notebook assignments, earn a satisfactory assessment for five post-laboratory assignments and one full written laboratory report, attend seven laboratory lectures, and pass the practical exam components at the appropriate levels.

Table 1. Comparison of letter grade requirements under the previous, points-based grading system and the specifications grading system.

	Criteria from Points-Based Grading System		Criteria from Specifications Grading System	
Course Requirements	Items Students Must Complete	Final Grade Weight	Course Grade Level*	Set of Criteria Completed
Online Pre-laboratory Homework Assignments	1 every week	28 points	A B C D	90 - 100 % correct 80 - 100 % correct 70 - 100% correct < 70% correct
Pre-laboratory Video Quizzes	1 every week	18 points	A B C D	85 - 100 % correct 80 - 100 % correct 75 - 100% correct < 75% correct
Laboratory Notebook Assignments	8	15 points/ day	A B C D	7 Satisfactory 6 - 7 Satisfactory 5 - 6 Satisfactory 4 Satisfactory
Post-laboratory Assignments	4	20-110 points	A B C D	5 Satisfactory + 1 full written laboratory report 4 Satisfactory 3 Satisfactory 2 Satisfactory
Lab Lecture Participation	Must participate	18 points	A B C D	7 required 6 required 4 - 5 required < 4 required
Practical Exam	1 final exam	205	A	Pass Mastery Final

points		Pass Knowledge Check w/S Passed 3 Lab Techniques Passed 4/6 safety questions
	B	Pass Mastery Final Pass Knowledge Check w/S Passed 2 Lab Techniques Passed 4/6 safety questions
	C	Pass Knowledge Check w/S Passed 1 Lab Technique Passed 4/6 safety questions
	D	< above criteria

*Students who do not meet the minimum criteria for D grade earn an F in the course.

Specifications Grading Assignment Rubrics

Under the specifications grading system, expectations for satisfactory work on assignments must be provided clearly. To communicate these expectations, we adjusted the assignment rubrics from the points-based rubrics — which allowed for partial credit in addition to full credit — to binary satisfactory/unsatisfactory-based rubrics. In this new rubric design, students either earned credit for a rubric item or they did not; no partial credit was awarded. This redesign necessitated revision of the points-based rubrics to better separate elements that had been grouped together into defined, separate, rubric items. For example, we parsed the singular theory rubric item of an experiment's post-laboratory assignment under the old grading system into four individual rubric items under the specifications grading system (Table 2, see the Supporting Information for a more detailed example). Satisfactory thresholds for assignments were set to approximately 80% of the total rubric items. These thresholds were chosen to ensure that students who earned credit for an assignment achieved proficiency.

Table 2. Comparison of a section of a points-based rubric and a specifications rubric for a post-laboratory assignment.

Criteria from Points-Based Rubric	Points	Criteria from Specifications Rubric	Satisfactory	Unsatisfactory
Theory (Full Credit): Student discusses fundamentals of column chromatography and relates the technique to TLC, noting similarities and differences and how a successful separation is achieved.	7	Theory 1a: Clearly describes the chemical principle(s) that govern how compounds are separated using column chromatography. Note: Be sure to include the importance of solvent choice.	☐	☐
		Theory 1b: Clearly compares and contrasts column chromatography to TLC.	☐	☐
		Theory 1c: Clearly describes what procedural steps must be taken to achieve a successful separation using column chromatography.	☐	☐
		Theory 1d: Clearly explains how separation is monitored in real time, and how this allows the determination of whether the separation was successful or not.	☐	☐

Restructuring the rubrics may provide the added benefit of simplifying grading for the TAs. While grading, TAs only need to view one criterion, or rubric item, at a time and decide whether the student's work meets the criterion or not. This system is intended to reduce the time TAs need to spend deciding what score — on the spectrum of each rubric criterion from the original rubrics — a student's report should earn.

For a competency-based approach to function under the specifications grading framework, students need to be given opportunities to learn from their mistakes and to be reassessed. Any students whose work does not meet the satisfactory threshold established for an assignment does not earn any credit for that assignment. The token system provides students with a limited number of opportunities to revise and resubmit work for credit that would overwrite their previous grade. This structure not only allows students to incorporate feedback to pass assignments they initially did not, but it also permits students to choose if and when to resubmit work. The token system provides the additional benefit of acting as a safety net for students when unexpected events temporarily hinder their ability to complete coursework.

We were inspired by Blackstone et al.'s token system, and used it as a model for our own.[23] To earn an initial four tokens, students completed a short, self-regulatory learning assignment at the beginning of the course.[55] Students could earn an additional, limited number of tokens throughout the course for completing additional tasks such as participating in midterm and end-of-term course feedback surveys. In addition to using tokens for assignment revisions, students could also redeem tokens in other course contexts, such as extending assignment deadlines or attending a make-up laboratory if a laboratory section was missed. This flexibility eliminated the need for students to provide explanations and to request exemptions for late work and absences. Token redemptions were tracked through the use of a Google form and a placeholder assignment in the course learning management system (LMS) that listed each student's current token count.[56]

Specifications Grading Exams

Converting our course to a specifications grading system also necessitated a restructuring of the laboratory practical exam. Under the previous points-based grading system, students completed a laboratory practical during the last week of the term; this exam consisted of a wet laboratory portion, where students performed an organic chemistry laboratory technique (e.g. thin-layer chromatography, recrystallization, extraction, or melting point), and a dry laboratory portion. For the dry laboratory portion, students answered a critical thinking question, performed experiment-based calculations (e.g. theoretical yield, unit conversions, etc.), drew an accurate reaction mechanism for a reaction covered during the course, used provided spectra to identify an unknown organic compound, and answered multiple-choice laboratory safety questions.

Under the specifications grading system, we defined four components of the laboratory practical exam. The first three components — a safety final, a knowledge check final, and a technique final — represent the core competencies a student needed to demonstrate to pass the course and were required to earn a C-level or higher grade (Figure 1). Students who aimed for a higher letter grade were required to complete additional laboratory techniques and to complete the fourth component of the practical exam — the mastery final.

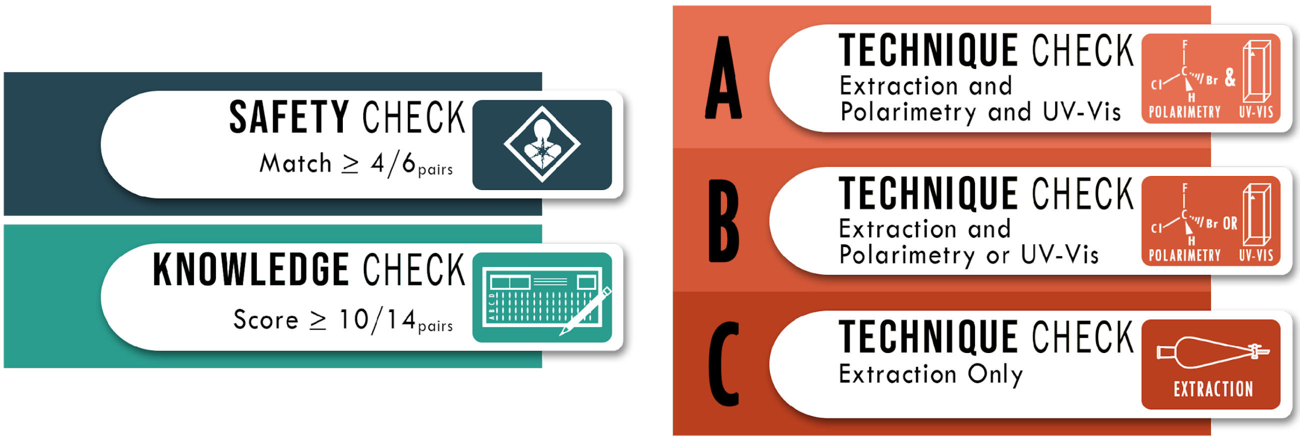


Figure 1. Required components of the laboratory practical exam under the specifications grading system. Students must perform the safety check, knowledge check, and the technique check. Depending on the final grade the student aims for, they can choose to complete only the liquid-liquid extraction technique for a C-level grade, both the liquid-liquid extraction technique and polarimetry or UV-visible spectroscopy technique for a B-level grade, or all three techniques for the technique check portion of the exam.

The safety final component was included to determine if students had achieved competency in determining best safety practices in the lab; it consisted of a collection of six images illustrating unsafe laboratory practices (e.g. glass waste in the trash can, an open chemical container sitting out on the bench top, etc.). Students matched each image to the appropriate unsafe practice chosen from an answer bank. To pass the safety component, four of the six unsafe practices had to be matched correctly (Figure 1).

The knowledge check component was included as a multiple-choice exam with fourteen questions to assess if students achieved competency in fundamental course concepts and skills. The exam included conceptual questions on each laboratory technique taught that term, stoichiometry and limiting reagent calculations, identification of GHS hazard symbols, matching a 1HNMR spectrum to a molecular structure, and recognizing the correctly drawn reaction mechanism for a reaction conducted that term.[57] These questions tied directly to course SLOs 2, 5, 6, 4 and 3. To pass the knowledge check component, ten of the fourteen questions must have been answered correctly (Figure 1). If students did not pass this exam component on their first attempt, they were given a second chance to pass by taking a different version of the exam on a specified date within the campus' final exam window. This final retake option was included to incorporate a mastery learning component to the final exam structure, where students are given an opportunity to learn from their mistakes and be reassessed.

The technique final exam components, designed to test students' ability to perform fundamental laboratory skills (SLO 1), retained the format from the wet laboratory portion of the previous version of the laboratory practical exam. The three techniques assessed (liquid-liquid extraction, absorbance spectroscopy, and polarimetry) were chosen because other common organic chemistry laboratory techniques were assessed in previous courses and were not used often in this course (e.g. thin-layer chromatography, melting point, recrystallization). At the beginning of this pilot course, the instructor informed the students which laboratory techniques would be included in the technique components of the final exam. Liquid-liquid extraction was selected as the required technique for a C-level grade because it was the laboratory technique used most frequently throughout the course (Figure 1). Students aiming for a B-level grade had to perform and pass one additional technique; they could choose between polarimetry and absorbance spectroscopy because these techniques were not included in prior organic chemistry laboratory courses. Those students aiming for an A-level grade had to perform and pass both of the additional techniques.

The mastery final component provided students an opportunity to demonstrate a level of proficiency of the course content greater than the competency demonstrated on the knowledge check component. The mastery exam component consisted of three main question categories: conceptual critical thinking (SLO 2), experimental calculation critical thinking (SLO 5), and spectroscopy (SLO 4) (Figure 2). Two open-ended questions requiring responses with thorough explanations were provided in each category, and students were given the option to complete one or both questions in each category. Because points were additive, not deductive, students were encouraged to attempt as many questions as they felt they could answer. Each question was given a defined partial pass threshold (0.5) and a full pass threshold (1), where the difference between the two hinged on the depth of the student explanation. TAs were given detailed rubrics that defined assignment of partial and full pass thresholds. The following cumulative pass thresholds were needed to achieve the corresponding letter grade: 3 for an A, 2.5 for an A-, 2 for a B+, 1.5 for a B, and 1 for a B- (Table 1). The students only earned these final letter grades if they also met all other criteria specified for those letter grades, including score thresholds for online pre-lab homework, video quizzes, and lab lecture participation as listed on the Student Grade Tracker (Supporting Information).

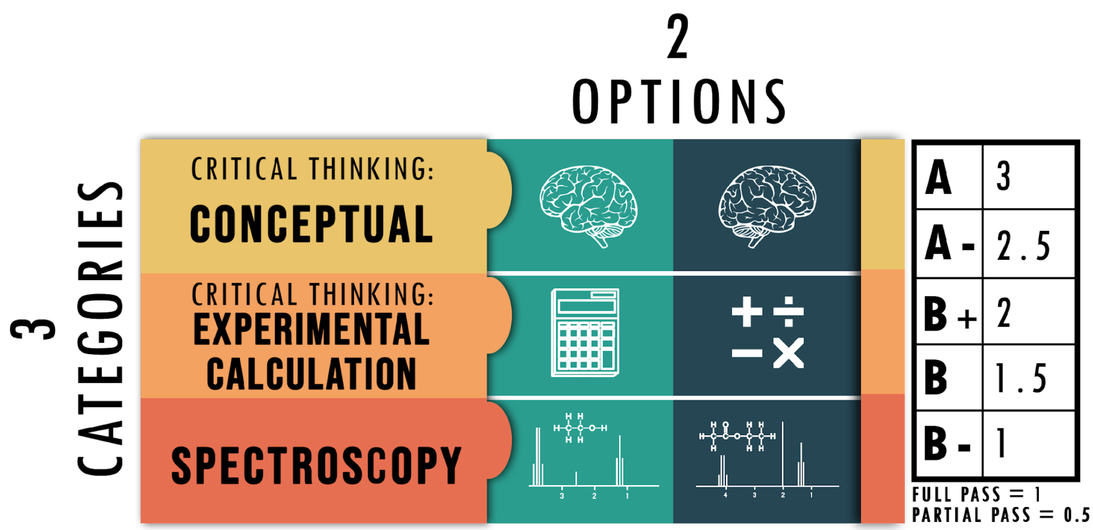


Figure 2. Categories of the mastery final portion of the laboratory practical exam under the specifications grading system. Students must complete the mastery final component if they wish to aim for an A-level or B-level final grade in the course. There are two questions per category, with a total of six questions on the mastery final. Students can attempt any of the six questions, and earn the depicted number of full or partial passes to be eligible to earn the corresponding letter grades.

PILOT IMPLEMENTATION STUDY OUTCOMES

In this specifications grading pilot implementation, we endeavored to trial the new specifications grading system and to determine how students and TAs perceived it. The implementation of this system in the small organic chemistry laboratory course allowed us to assess whether the grading system could be viable in a large laboratory course setting. Students were introduced to the new grading system at the first class meeting where we attempted to establish buy-in by emphasizing how specifications grading eliminates competition and provides students with greater opportunities to exercise agency in their learning. Because this system was new, students were also given specific opportunities to ask questions about grading at every laboratory lecture meeting. We surveyed students midway through the course, and we asked for both student and TA feedback at the conclusion of the course to determine their perceptions of what worked well and what needed improvement. Survey questions and responses for both groups are provided in the Supporting

Information. We also compared students' course letter grades in this course to a previous course offering which did not use a specifications grading system.

Teaching Assistant Perceptions

The two course TAs — who have taught two or more organic chemistry laboratory courses in previous terms — were asked for feedback on the specifications grading system at the conclusion of the course. Their perceptions of the grading system were strongly positive, and both described grading student work with the new rubrics as simpler and faster compared to using traditional rubrics with partial credit options:

"I think this grading and overall system is a lot easier to use and it makes the workload for TAs less intensive and time consuming."

"I liked that it was in binary."

"I think it makes it a lot easier to grade and gets rid of the uncertainty about meeting the rubric criteria."

In addition to efficiency, the TAs also reported spending more time discussing student understanding of course material, over email and in person, than discussing complaints over assignment grading. This report contrasts with anecdotes from previous TAs, who taught in iterations of the course where traditional points-based rubrics were used. The TAs stated that students generally contacted them in an attempt to negotiate for more points.

Student Perceptions

Students were surveyed twice in this course to determine their perceptions of specifications grading. Anonymous surveys were administered midway through the course and at the end of the course. Student attitudes toward the specifications grading system were mixed, and changed from more negative during the course to more positive after the course concluded.

Of the 37 students enrolled in the course, five responded to the midterm survey, and four to the post-course survey. Although the response rate was low, as is typical when incentives such as extra credit are not offered, recorded perceptions matched what students reported anecdotally through in-person interactions with the instructor. In the midterm survey, students commented that the "all-or-nothing" aspect of the assignment grading made the class more stressful for them (Table 3). Although students praised the token system, they did not like that a token was necessary to revise and resubmit an assignment that had missed the “satisfactory” cutoff by only one rubric item. Several students commented that they felt it was unfair that they did not receive any credit for turning in work even though it did not meet the “satisfactory” criteria. Students also commented that the new rubrics were far less detailed than previous versions. However, TAs had the opposite perception of the rubrics, describing the new rubrics as more detailed and clearer.

Table 3. Student feedback during and after the course.

Midterm Feedback (n = 5)	Post-Course Feedback (n = 4)
Grading is stressful because of the “all-or-nothing” approach	Grading is less stressful because students could track their grade
Liked tokens in general	Grading is less stressful because of the option to revise for credit
Did not like that a token was required to revise if only short one rubric item	Perceived TA grading as more standardized
Perceived rubrics as less detailed	Wanted partial credit
	Satisfactory thresholds set too high

The student feedback from the post-course survey was more positive than the midterm feedback. Three students commented that the grading was less stressful because they always knew where they stood and because the system allowed them to try again when needed. Two of these three students also felt that the grading between TAs was more standardized with the all-or-nothing rubric items. Of the remaining two students who provided feedback, one had a more negative view of the specifications grading system. The student commented that partial credit from the old grading system was better because at least they could get some credit for an assignment, whereas in the specifications grading system, missing a requirement for a grade in any one category could ruin their chances of earning that grade. The other student felt that the cutoffs for earning a satisfactory on assignments was too high for undergraduates and that the cutoff should be set at a C-level, requiring only 70% of rubric items.

Comparison of Grade Distributions

Although students voiced concerns that the lack of partial credit opportunities would hurt their grades, students in the specifications graded course earned higher letter grades than students in a previous course offering with points-based grading. Final letter grades for students in the specifications graded course (n = 37) were compared to those from a traditionally graded version of the same course taught by the same instructor in a prior year (n = 68). In the specifications graded course 43% of students earned A-level grades, 46% of students earned B-level grades, and 11% of students earned C-level grades (Figure 3).[58–60] These grades represent a shift toward higher overall grades when compared to the traditionally graded version of the course where students earned 34% A-level grades, 43% B-level grades, 22% C-level grades and 1% D-level grades. No F grades were recorded for either course, and there were no withdrawals because university rules do not permit students to withdraw after the conclusion of the second week of the term. Similar increases in final course grade distributions between the specifications and points-based grading systems have been reported in other implementations of specifications grading.[17,24] The increase in course grades has been postulated to result from an emphasis on the mastery learning aspects of the grading system.[61–63] The higher final grades in our course using a specifications grading system could indicate that students were better able to meet course outcomes because they were provided opportunities for mastery, but we are not able to determine the cause of higher student grades because assessments in the points-based version of the course were less clearly tied to SLOs.

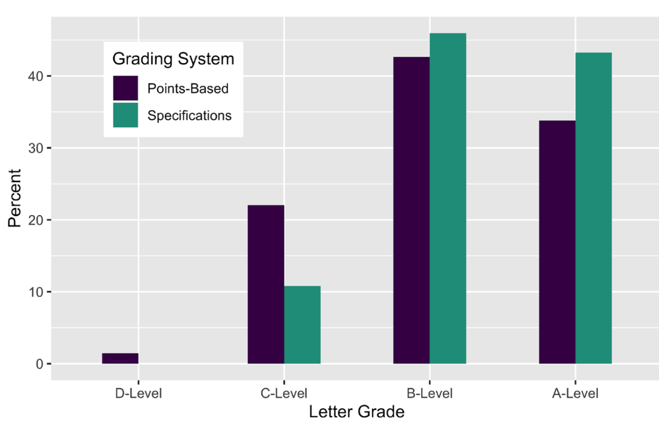


Figure 3. Grade distributions of a previous iteration of the course using a points-based system and the current course with the specifications grading system. See Figure S1 in the Supporting Information for the distribution including +/- grades.

CONSIDERATIONS FOR FUTURE SCALED-UP COURSE IMPLEMENTATIONS

The primary goal of implementing a specifications grading system in this course was to test how the new system would work on a small scale. By identifying and resolving concerns in this smaller course, we would be better prepared to implement specifications grading on a larger scale in on-sequence lab course offerings. We were concerned that the time spent on managing the token system and grading revised assignment submissions would prove laborious, but we found that this aspect of the grading system should indeed be scalable. We also learned that establishing student buy-in to the new grading system was especially important to prevent student misconceptions about their course grade standing.

Contrary to our initial concerns, the time required to implement the token system and grade revised student work was not onerous. The instructor checked the Google form and updated students' token balances by changing the "score" in the placeholder assignment in the course LMS — a process that required approximately 10 minutes per day, on average. Most token trade requests were for assignment revisions, and by viewing the marked assignment rubric in the course LMS, students could identify items for which they did not earn a satisfactory assessment. If students chose to use a token to revise and resubmit an assignment, they only had to revise the unsatisfactory sections and resubmit to the same assignment in the course LMS by a specified deadline. TAs graded the revised submissions, focusing on the rubric items where students did not earn credit on the first attempt. Even with assignment revision requests, TAs reported that the time commitment was not burdensome. These considerations suggest that the specifications grading system should also be manageable in other STEM laboratory courses.

To prevent student misconceptions about their course grade standing, which can result in an overwhelming number of complaints in a larger course, it will be essential to establish buy-in and consistently provide reminders about the big picture of the grading system. Throughout the first half of the course, students were focused on the perceived higher stakes for individual assignments. Students had access to a Student Grade Tracker, but they seemed unaware of how individual assignments related to the requirements for each letter grade. Students indicated they were stressed about not earning satisfactory scores for post-laboratory assignments, but they did not realize that they could earn an unsatisfactory score on one post-laboratory assignment and still earn an A-level grade in the course. After the midterm survey, we realized students were misinterpreting the Student Grade Tracker, so we devoted a small amount of class time to reviewing the tracker.

To address student misconceptions, we will provide more information at the beginning of the next course offering to establish greater student buy-in. Students will explicitly be told to shift their focus from individual category achievement to their overall grade standing in the course. Our goal is to ensure students realize that missing one category from the Student Grade Tracker will not cause them to fail the course. We will also reframe the binary satisfactory or unsatisfactory grading system as satisfactory or needs revision, and emphasize that students can resubmit assignments using the token system. These adjustments should result in less student concerns about stress related to earning a satisfactory score on all assignments.

Our pilot implementation of a specifications grading system in an organic chemistry laboratory course was successful, but questions still remain. Despite students' concerns about "all-or-nothing" grading, final course grades were higher overall for the course when offered with specifications grading than with points-based grading. We do not know whether students were more successful in meeting course objectives or our grade requirements need to be adjusted in the future. Furthermore, although our specifications grading system is designed to be scalable, we will need to verify its utility in a large course. We will explore these lingering questions in future studies.

ASSOCIATED CONTENT

Supporting Information

Supporting Information (pdf)

Specifications Grading Rubric Example (xlsx)

AUTHOR INFORMATION

Corresponding Author

*E-mail: rlink@uci.edu

Author Contributions

‡W.J.H. and K.J.M contributed equally.

NOTES

The authors declare no competing financial interests.

ACKNOWLEDGMENTS

We thank Katelyn Haduong, Selina Sanchez, Emma Wilfong, Jasmine Yuen, Elizabeth Wright, Natalie Samaan, and Michelle Diab for their feedback on rubric design. Taylor Thane and Simon Lam provided integral support in course implementation. We also thank Professor Kieron Burke for the inspiration to create a new, scalable grading approach. Finally, we thank Denise Bui for designing the graphical abstract and the first two figures in the paper.

REFERENCES

(1) Rom, M. C. Grading More Accurately. *Journal of Political Science Education* **2011**, *7* (2), 208–223.

(2) Guskey, T. R.; Bailey, J. M. *Developing Grading and Reporting Systems for Student Learning*; Guskey, T. R., Marzano, R. J., Eds.; Experts in Assessment; Corwin Press: Thousand Oaks, CA, 2001.

(3) Butler, R.; Nisan, M. Effects of No Feedback, Task-Related Comments, and Grades on Intrinsic Motivation and Performance. *J. Educ. Psychol.* **1986**, *78* (3), 210–216.

(4) Butler, R. Enhancing and Undermining Intrinsic Motivation: The Effects of Task-Involving and Ego-Involving Evaluation on Interest and Performance. *Br. J. Educ. Psychol.* **1988**, *58* (1), 1–14.

(5) Kitchen, E.; King, S. H.; Robison, D. F.; Sudweeks, R. R.; Bradshaw, W. S.; Bell, J. D. Rethinking Exams and Letter Grades: How Much Can Teachers Delegate to Students? *CBE Life Sci. Educ.* **2006**, *5* (3), 270–280.

(6) Reddy, Y. M.; Andrade, H. A Review of Rubric Use in Higher Education. *Assessment & Evaluation in Higher Education* **2010**, *35* (4), 435–448.

(7) Allen, D.; Tanner, K. Rubrics: Tools for Making Learning Goals and Evaluation Criteria Explicit for Both Teachers and Learners. *CBE Life Sci. Educ.* **2006**, *5* (3), 197–203.

(8) Jonsson, A.; Svingby, G. The Use of Scoring Rubrics: Reliability, Validity and Educational Consequences. *Educational Research Review* **2007**, *2* (2), 130–144.

(9) Schinske, J.; Tanner, K. Teaching More by Grading Less (or Differently). *CBE Life Sci. Educ.* **2014**, *13* (2), 159–166.

(10) Nilson, L. B.; Stanny, C. J. *Specifications Grading: Restoring Rigor, Motivating Students, and Saving Faculty Time*, Reprint edition.; Stylus Publishing: Sterling, VA, 2014.

(11) Bonner, M. W. Grading Rigor in Counselor Education: A Specifications Grading Framework. *Educational Research Quarterly* **2016**, *39* (4), 21.

(12) Elkins, D. M. Grading to Learn: An Analysis of the Importance and Application of Specifications Grading in a Communication Course. *Kentucky Journal of Communication* **2016**, *35* (2), 2016.

(13) Toledo, S.; Dubas, J. M. A Learner-Centered Grading Method Focused on Reaching Proficiency with Course Learning Outcomes. *J. Chem. Educ.* **2017**, *94* (8), 1043–1050.

(14) Blodgett, B. J. Grading Matters in Theological Education. *Teach Theol Relig* **2017**, *20* (4), 314–326.

(15) Ring, J. ConfChem Conference on Select 2016 BCCE Presentations: Specifications Grading in the Flipped Organic Classroom. *J. Chem. Educ.* **2017**, *94* (12), 2005–2006.

(16) Boesdorfer, S. B.; Baldwin, E.; Lieberum, K. A. Emphasizing Learning: Using Standards-Based Grading in a Large Nonmajors' General Chemistry Survey Course. *J. Chem. Educ.* **2018**, *95* (8), 1291–1300.

(17) Hollinsed, W. C. Applying Innovations in Teaching to General Chemistry. In *Increasing Retention of Under-Represented Students in STEM through Affective and Cognitive Interventions*; ACS Symposium Series; American Chemical Society: Washington, DC, 2018; Vol. 1301, pp 145–152.

(18) Mendez, J. Standards-Based Specifications Grading in a Hybrid Course. In *2018 ASEE Annual Conference & Exposition*; 2018.

(19) Mendez, J. Standards-Based Specifications Grading in Thermodynamics. In *2018 ASEE Annual Conference & Exposition*; 2018.

(20) Mirsky, G. M. Effectiveness of Specifications Grading in Teaching Technical Writing to Computer Science Students. *J. Comput. Sci. Coll.* **2018**, *34* (1), 104–110.

(21) Roberson, C. Techniques for Using Specifications Grading in Computer Science. *J. Comput. Sci. Coll.* **2018**, *33* (6), 192–193.

(22) Williams, K. Specifications-Based Grading in an Introduction to Proofs Course. *PRIMUS* **2018**, *28* (2), 128–142.

(23) Blackstone, B.; Oldmixon, E. Specifications Grading in Political Science. *Journal of Political Science Education* **2019**, *15* (2), 191–205.

(24) Carlisle, S. Simple Specifications Grading. *PRIMUS* **2019**, 1–26.

(25) Martin, L. J. Introducing Components of Specifications Grading to a General Chemistry I Course. In *Enhancing Retention in Introductory Chemistry Courses: Teaching Practices and Assessments*; ACS Symposium Series; American Chemical Society, 2019; Vol. 1330, pp 105–119.

(26) McGuffee, J. W.; Largent, D. L.; Roberson, C. Transform Your Computer Science Course with Specifications Grading. In *Proceedings of the 50th ACM Technical Symposium on Computer Science Education*; SIGCSE '19; Association for Computing Machinery: New York, NY, USA, 2019; p 1234.

(27) Shields, K.; Denlinger, K.; Webb, M. Not Missing the Point (s): Applying Specifications Grading to Credit-Bearing Information Literacy Classes. In *The Grounded Instruction Librarian: Participating in The Scholarship of Teaching and Learning*; Association of College and Research Libraries: Chicago, IL, 2019.

(28) Tsoi, M. Y.; Anzovino, M. E.; Erickson, A. H.; Forringer, E. R.; Henary, E. Variations in Implementation of Specifications Grading in STEM Courses. *Georgia Journal of Science* **2019**, *77* (2).

(29) Pope, L.; Parker, H. B.; Ultsch, S. Assessment of Specifications Grading in an Undergraduate Dietetics Course. *J. Nutr. Educ. Behav.* **2020**, *52* (4), 439–446.

(30) Gratwick, R.; Kinnear, G.; Wood, A. K. An Online Course Promoting Wider Access to University Mathematics. *Proceedings of the British Society for Research into Learning Mathematics* **2020**, *40* (1).

(31) Jones, P. A. Implementing Specifications Grading in MPA Courses: A Potential Strategy for Better Work-Life Balance. *Journal of Public Affairs Education* **2020**, 1–17.

(32) Leslie, P.; Lundblom, E. Specifications Grading: What It Is and Lessons Learned. *Semin. Speech Lang.* **2020**.

(33) Prasad, P. V. Using Revision and Specifications Grading to Develop Students' Mathematical Habits of Mind. *PRIMUS* **2020**, 1–18.

(34) Bloom, B. S. Learning for Mastery. Instruction and Curriculum. Regional Education Laboratory for the Carolinas and Virginia, Topical Papers and Reprints, Number 1. *Evaluation Comment* **1968**, *1* (2).

(35) Keller, F. S. Engineering Personalized Instruction in the Classroom. *RIP/IJP* **1967**, *1* (3), 198–197.

(36) Bangert-Drowns, R. L.; Kulik, C.-L. C.; Kulik, J. A.; Morgan, M. The Instructional Effect of Feedback in Test-Like Events. *Rev. Educ. Res.* **1991**, *61* (2), 213–238.

(37) Diegelman-Parente, A. The Use of Mastery Learning with Competency-Based Grading in an Organic Chemistry Course. *J. Coll. Sci. Teach.* **2011**, *40* (5), 50–58.

(38) Docan, T. N. Positive and Negative Incentives in the Classroom: An Analysis of Grading Systems and Student Motivation. *Journal of Scholarship of Teaching and Learning* **2006**, *6* (2), 21–40.

(39) Voorhees, R. A. Competency-Based Learning Models: A Necessary Future. *New Directions for Institutional Research* **2001**, *2001* (110), 5–13.

(40) Competency refers to a student's ability to achieve course student learning outcomes successfully or efficiently. In contrast, student proficiency is achieved with a high level of competence.

(41) Joksimović, S.; Gašević, D.; Kovanović, V.; Riecke, B. E.; Hatala, M. Social Presence in Online Discussions as a Process Predictor of Academic Performance. *Journal of Computer Assisted Learning* **2015**, *31* (6), 638–654.

(42) Leong, P. Role of Social Presence and Cognitive Absorption in Online Learning Environments. *Distance Education* **2011**, *32* (1), 5–28.

(43) Shea, P.; Sau Li, C.; Pickett, A. A Study of Teaching Presence and Student Sense of Learning Community in Fully Online and Web-Enhanced College Courses. *The Internet and Higher Education* **2006**, *9* (3), 175–190.

(44) Swan, K. Virtual Interaction: Design Factors Affecting Student Satisfaction and Perceived Learning in Asynchronous Online Courses. *Distance Education* **2001**, *22* (2), 306–331.

(45) Taylor, H. *Contract Grading*; EPIC-TM-75; EPIC Clearinghouse on Tests, Measurement, and Evaluation, 1980.

(46) Beare, P. G. The Contract--An Individualized Approach to Competency-Based Learning and Evaluation. In *Thinking across the Disciplines*; 1986.

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

(48) Hiller, T. B.; Hietapelto, A. B. Contract Grading: Encouraging Commitment to the Learning Process through Voice in the Evaluation Process. *Journal of Management Education* **2001**, *25* (6), 660–684.

(49) Boe, J. What the F! *Writing from the Edge* **2010**, *20* (2), 5–7.

(50) Carter, C. S.; Brickhouse, N. W. What Makes Chemistry Difficult? Alternate Perceptions. *J. Chem. Educ.* **1989**, *66* (3), 223.

(51) Seymour, E. *Talking About Leaving: Why Undergraduates Leave The Sciences*, 1 edition.; Westview Press: Boulder, CO, 1997.

(52) Humphreys, B.; Johnson, R. T.; Johnson, D. W. Effects of Cooperative, Competitive, and Individualistic Learning on Students' Achievement in Science Class. *J. Res. Sci. Teach.* **1982**, *19* (5), 351–356.

(53) Tobias, S. *They're Not Dumb, They're Different*; Research Corporation: Tucson, AZ, 1990.

(54) UCI does include plus and minus grades, and our course does have set criteria for students to achieve +/- grades. We have left them out here for clarity. The Supporting Information contains full criteria for all letter grades.

(55) Nilson, L. *Creating Self-Regulated Learners: Strategies to Strengthen Students? Self-Awareness and Learning Skills*; Stylus Publishing: Sterling, VA, 2013.

(56) The LMS does not have a straightforward method for students to keep track of tokens. We used the assignment feature in the LMS to give each student four tokens as four points for their "score." as the students use tokens, points are subsequently deducted from the assignment "score" total.

(57) The GHS symbol questions were included in the knowledge check exam component to emphasize that recognizing these symbols is a fundamental aspect of training in a laboratory class. By contrast, the safety exam component focused on identifying hazardous laboratory situations, not chemicals.

(58) R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2019.

(59) Garnier, S. viridis: Default Color Maps from “matplotlib” https://CRAN.R-project.org/package=viridis (accessed Sep 12, 2019).

(60) Wickham, H.; Averick, M.; Bryan, J.; Chang, W.; McGowan, L.; François, R.; Grolemund, G.; Hayes, A.; Henry, L.; Hester, J.; Kuhn, M.; Pedersen, T.; Miller, E.; Bache, S.; Müller, K.; Ooms, J.; Robinson, D.; Seidel, D.; Spinu, V.; Takahashi, K.; Vaughan, D.; Wilke, C.; Woo, K.; Yutani, H. Welcome to the Tidyverse. *JOSS* **2019**, *4* (43), 1686.

(61) Diegelman-Parente, A. The Use of Mastery Learning with Competency-Based Grading in an Organic Chemistry Course. *J. Coll. Sci. Teach.* **2011**, *40* (5), 50–58.

(62) Garver, K. A Computerized Approach to Mastery Learning: Helping Community College Students Make the Grade in Anatomy. *J. Coll. Sci. Teach.* **1998**, *28* (2), 94–96.

(63) Livingston, J. A.; Gentile, J. R. Mastery Learning and the Decreasing Variability Hypothesis. *J. Educ. Res.* **1996**, *90* (2), 67–74.