Georgia Journal of Science

Volume 77 No. 2 Scholarly Contributions from the Membership and Others

Article 10

2019

Variations in Implementation of Specifications Grading in STEM Courses

Mai Yin Tsoi
Georgia Gwinnett College, mtsoi@ggc.edu

Mary E. Anzovino
Georgia Gwinnett College, manzovino@ggc.edu

Amy H. Lin Erickson
Georgia Gwinnett College, aerickso@ggc.edu

Edward R. Forringer
Georgia Gwinnett College, eforring@ggc.edu

Emily Henary
Georgia Gwinnett College, ehenary@ggc.edu

See next page for additional authors

Variations in Implementation of Specifications Grading in STEM Courses

Acknowledgements
We wish to acknowledge Georgia Gwinnett College School of Science and Technology, Dr. Thomas Mundie (Dean of SST), Dr. Sebastien Siva, the National Science Foundation and University System of Georgia for their funding of our faculty learning community, and our family and friends for their support of this work.

Authors
Mai Yin Tsoi, Mary E. Anzovino, Amy H. Lin Erickson, Edward R. Forringer, Emily Henary, Angela Lively, Michael S. Morton, Karen Perell-Gerson, Stan Perrine, Omar Villanueva, MaryGeorge Whitney, and Cynthia M. Woodbridge

VARIATIONS IN IMPLEMENTATION OF SPECIFICATIONS GRADING IN STEM COURSES

Mai Yin Tsoi
Mary E. Anzovino
Amy H. Lin Erickson
Edward R. Forringer
Emily Henary
Angela Lively
Michael S. Morton
Karen Perell-Gerson
Stan Perrine
Omar Villanueva
MaryGeorge Whitney
Cynthia M. Woodbridge
Georgia Gwinnett College
Lawrenceville, Georgia, 30043
mtsoi@ggc.edu, manzovino@ggc.edu, aerickso@ggc.edu, eforring@ggc.edu, ehenary@ggc.edu, alively@ggc.edu, mmorton1@ggc.edu, kperellg@ggc.edu, sperrine@ggc.edu, ovillanueva@ggc.edu, mwhitney@ggc.edu, cwoodbridge@ggc.edu

ABSTRACT

Specifications grading is an assessment strategy based on mastery learning, clear learning objectives, and frequent evaluations and feedback. Twelve instructors at a southeastern four-year public college implemented the specifications grading method across eight discrete courses in four STEM areas. In this modified assessment strategy, the students controlled their grades through multiple attempts, with limitations, on assessments of course objectives. The instructors designed and executed specifications grading in unique ways that aligned with their content areas, teaching beliefs, and individual teaching styles.

Preliminary observations suggest that, regardless of subject area, specifications grading can be used as an alternative to traditional assessment methodologies in STEM courses, regardless of the content area. In general, three major variations of implementation arose from this initial trial. Major differences and commonalities among these types are discussed as they relate to the course subject area in which they are used. The results of this work add a unique set of assessment practices to the current body of knowledge in that other practitioners may gain insight on variations of the specifications grading method that may be practical and applicable in their own classrooms.

Keywords: assessment, grades, course design, college math, college chemistry, college physics, college anatomy and physiology, mastery learning, specifications grading, STEM education

INTRODUCTION

Historically, assessment has been an integral part of the instructional process. Coming into the educational system in the 1700s, assessment became a formal part of education in the 1800s. Practices have had a long history of being individualized, and a lack of standardization, spawned the idea of common assessments where a standard set of knowledge and skills are tested. Common practices such as curving grades and grade inflation, along with individual differences in the interpretations assigned to grades, have further complicated the assessment milieu. Over the last half-century, we have seen many educational institutions forced to institute standardized measurements, such as the SAT and ACT, in an attempt to better make sense of what information a student's grades are meant to convey.

In addition to these areas of competitive evaluation, how others interpret student grades can also lead to a range of understandings about the quality of learning, amount of content learned, and a host of other learning factors that occur in the classroom. These interpretations can lead to beliefs that impact students as well as instructors, with far-reaching consequences such as influencing self-efficacy and motivation. Because of the variety and importance of the proper assessment of student learning and performance, educators have dedicated a great deal of time to the research and practice of assessment.

In 2015, Linda B. Nilson published *Specifications Grading: Restoring Rigor, Motivating Students, and Saving Faculty Time*, a book that describes a novel approach to assessing students. She describes a method that merges three areas of assessment that have never been combined: mastery learning, repeated attempts, and student control of grades. Nilson's proposed method aims to simplify the process of grading while restoring curricular rigor through the connection of grades to specific course outcomes. The work presented herein was implemented based on Nilson's ideology. This novel assessment method was implemented by 12 instructors across several disciplines in multiple class sections at a four-year college. Each instructor had freedom to design and develop his or her own variation of implementation. This manuscript attempts to describe and explain the evolution of the three major variations in the implementation of specifications grading so as to help inform other educators who may be interested in exploring this method within their own STEM courses.

CONTEXT

The institution at which this work has been conducted is unique. Georgia Gwinnett College is the first four-year public college established in the 21st century in the United States and the first new college established in the state in 100 years. Charged with upending conventional practices that may not necessarily target the needs of today's students, the faculty are encouraged to explore novel methods to best meet the learning needs of students as they prepare to enter the workforce. This institution accepts students without regard to high school or transfer GPA. Thus, the student population served by this college encompasses many backgrounds to include first-generation students, traditional high school graduates, international students, and nontraditional students who are returning to school after an extended absence or even another career in their adult years. The culture is one of innovative practices and of targeted instruction that aim towards institutional objectives and goals. This culture serves as the cornerstone of the growth and development of this college. For example, these guiding principles led the college to cap class section enrollment in STEM courses at 30 students, as research has indicated that a lower teacher-to-student ratio can have a positive impact on student learning. It is within this context that the authors implemented specifications grading into their STEM courses.

BACKGROUND

The principal investigator of this work first read about the concept of specifications grading in an article by Joshua Ring (2017) in which the implementation of this method in a first-semester organic chemistry course was described. During spring 2017, the principal investigator implemented specifications grading in two sections of first-semester organic chemistry, with each section enrolling 24 students.

The experiences of that first semester were shared at an in-house teaching and learning conference. Several faculty members expressed interest in the philosophy of mastery learning and how specifications grading is implemented within the assessment portion of the course, whereas most classroom interventions occur within the lecture portion. In the following semesters, several more faculty adopted this novel assessment method. At the time of press of this work, even more instructors have joined this work and are now implementing specifications grading in several upper-level STEM courses, including Biochemistry, Organometallic Chemistry, Bioinorganic Chemistry, Human Biomechanics, and Cell Biology.

***Table I*. Summary of instructors, content areas, and semester of specifications grading implementation**

Semester	Course name
Spring 2017	Organic Chemistry I
Summer 2017	Organic Chemistry I
	Principles of Chemistry I
	Principles of Chemistry II
Fall 2017	Organic Chemistry I
	Principles of Chemistry I
	Principles of Chemistry II
	Access Algebra*
	Calculus II
	Intro to Physics II
	Anatomy & Physiology I
Spring 2018	Organic Chemistry I
	Principles of Chemistry I
	Principles of Chemistry II
	Survey of Chemistry I
	Introduction to Physics I
	Introduction to Physics II

***Table I* (continued)**

Semester	Course name
Spring 2018	Calculus I
	Calculus II
	Access Algebra*
	College Algebra
	Anatomy & Physiology I
Summer 2018	Organic Chemistry II
	Principles of Chemistry I
	Principles of Chemistry II

One commonality among the 12 instructors is that each had previously attempted to increase the efficacy of student learning by implementing different teaching approaches within their classrooms, but none had extensively explored any alternative assessment practices such as specifications grading. Interested in maximizing student understanding and retention of skills, these instructors were open to trying novel approaches to teaching and learning. For example, two instructors had tried the flipped classroom technique (Makice 2012) to increase student buy-in and responsibility. Another instructor had been investigating the impact of case studies as a way to improve student retention of course content (Hamilton and Corbett-Whittier 2013). The teaching experience ranged novice to expert, with 15+ years in the classroom. The instructors involved in this work included both males and females with a breath of ethnic diversity and educational expertise. All instructors were self-selected, choosing to participate in the specifications grading project. They exhibited a strong commitment to improving student learning and to increasing rigor within their classrooms. Each instructor engaged in many hours of extra work beyond their regular teaching duties to redesign their courses as they implemented specifications grading.

Nilson (2015) has stated that specifications grading, as an assessment strategy, is truly content independent and can be applied to a multitude of situations, learning environments, and subject areas. Therefore, the instructors had the freedom to infuse this assessment method into their courses as they saw fit and in ways that supported their individual teaching styles. Several instructors collaborated and consulted with colleagues who were also involved with this work. There were many discussions and much reflection on methods of adaptation, implementation, and course practices. Many instructors have noted that these discussions were very useful in thinking about how to improve student learning.

SPECIFICATIONS GRADING

In traditional courses, students are usually assessed with several high-stakes graded events where the instructor uses different styles of questioning to cover content deemed essential to the instructor. Students may feel that they do not have a clear idea as to the exact content on their large assessments in the course. Typical grading schemes depend on the grades of these exams to determine the overall course grade for students. Many courses end with a high-stakes final assessment that usually covers a comprehensive overview of the salient topics that the instructor has chosen. Research indicates that high-stakes testing can increase stress in students (Horn 2003).

When large amounts of material are assessed at one time, it is difficult to ascertain whether students have mastered the learning outcomes for the course. The infamous cramming of material is a common result of these high-stakes tests, and studies have shown that retention and transfer of content knowledge can be adversely affected by these study strategies (Madaus 1991). At times, students may not be aware of what skills and knowledge they will need to demonstrate on an assessment knowledge base. Commonly, students are left to guess as to what the instructor defines as essential and necessary for learning in that course (Nilson 2015). If a student determines inaccurately what is going to be assessed, then it is possible that the student will earn a bad grade—not because the student had a poor comprehension of the material, but because the assessment did not accurately measure the student's mastery of the material that was actually learned. The stress of assessment gives rise to practices like grade-grubbing, where students protest grades and pressure instructors for extra points and even cheating, where students try to game the system and submit work that is not representative of their comprehension (Nilson 2015).

Nilson (2015) discusses the importance of assessment and how it needs to serve as an instrument of information, both for the student and for the faculty member. She states that some of the goals of a grading system should be to uphold high standards, reflect student learning outcomes, motivate students to learn and excel, discourage cheating, reduce student stress, put responsibility of grades on students, and minimize conflicts between students and instructors. She also discusses how course objectives communicate to the student what the instructor finds valuable and important in a course. Termed "a contract" between the student and the teacher, course objectives help both parties understand and agree to a predetermined set of skills and knowledge that are required for mastery of the course content.

Within Nilson's parameters, the instructors determined how the major areas of the specifications grading system might be implemented within specific courses: retakes, token management, mixtures of specifications versus traditional grading, requirements for demonstrating mastery of the objectives, and methods for determining the overall course grade. All instructors incorporated retakes into their course in some way—students received more than one opportunity to demonstrate mastery of each learning outcome for more than one time. For many instructors, tokens are one way of giving students more control of their success by providing a way for them to earn additional retakes until they are able to demonstrate mastery. Some instructors placed some limitations on the number of retakes each student received via some kind of token system.

During the first week of school, each instructor dedicated a substantial amount of class time to describing the specifications grading system, explaining the method of implementation for that particular course, answering questions, and providing clarifying documents or video explanations to help familiarize students with this alternative and novel method of grade determination. Students were informed that the intent of the system was to increase rigor and content retention while rewarding students with control over their own learning and course grade. The students were also notified of other implications of enrolling in a specifications grading-based course, namely that the locus of control for learning and course grades is within the student and that students were expected to depend heavily on that locus throughout the course.

ASSESSMENTS AND GRADING

All instructors started with the intention to set the cut score, or minimum level of proficiency, at 80% for each assessment. This is considered mastery of the skill. If a student met or exceeded this minimum level for an objective, an assessment PASS was recorded in the gradebook. No numerical or letter grade was assigned for the individual specifications grading assessments. At the end of the course, the number of assessments passed determined the numerical or letter grade for the specifications grading portion of the overall course grade. However, if a student did not meet the minimum level for an objective, an assessment NO-PASS was recorded. Students were directed to retake a new assessment version addressing the same objective until they achieved a mastery score. Once mastery was achieved, an assessment PASS was recorded. No grade penalty occurred with multiple retakes for the same objective. Some instructors adjusted this benchmark in their courses with alternate requirements. This benchmark was used regardless of the number of items on any assessment. Each instructor independently determined the number of items on any given assessment, as well as the styles and types of assessment questions to be asked. The instructors based these choices on their content area and individual style of teaching. Students were informed that, to be successful, they would need to study for the course continually and to develop ways of remediating misconceptions. In the initial class meeting, instructors emphasized that, with this specifications grading system, students had full control of their course grades in that they could choose how many assessments to master and which ones.

The other key aspect of specifications grading that most of the instructors tried to incorporate was to remove partial credit that had direct impact on a student's grade in the course. As instructors became comfortable with implementation, some did bring back this aspect in alternate ways, for example with extra tokens. But very few of these instructors allowed students to round their way to a higher course grade without demonstrating mastery of course content. Because the system is based on the theory of mastery learning (Block and Burns 1976), problems on the assessments were graded as completely correct or completely incorrect without giving any partial credit. For example, on a math assessment, a student was able to calculate the derivative correctly. However, because a negative sign was forgotten on the third step of the calculation, the whole item would be marked as incorrect.

The instructors placed their assessments in a repository housed on the school's learning management system so that other instructors could access this bank of assessments and then tailor them further to best fit their teaching style and particular emphases in their own lectures. Some instructors even adjusted the course objectives so they were more in line with what was taught in their own particular class and with what they believed was critical to the learning success of the students. As more instructors joined the project, more assessments were created and placed into the repository. Then, with each new iteration of the course, instructors used the repository assessments as templates for more versions and questions. This allowed the repository to grow exponentially and provide all instructors with a rich library of assessments from which to draw.

Several instructors brainstormed ways to manage record-keeping for scores earned on student retakes of assessments. Some faculty had folders for each student and stored all retakes for that student in the folder. Others chose to store together all students' retakes for a particular outcome. Another method of data management involved technology. One instructor created an Excel spreadsheet in which students' retake scores could be input and the workbook was formatted so that the instructor could quickly ascertain, for each student, which assessments were passed, and which versions of assessments had been completed at any given time. Another instructor housed students' scores in the institutional online learning management system. To track which versions of assessment each student had completed, these instructors had to set up the gradebook such that there was a grade item for every version of every assessment. For example, for Assessment 2.6, the grade book would list "2.6 Version A, 2.6 Version B" and so on. To start the semester, the instructors set up each assessment for five versions. However, if a at least one student needed more than five attempts to pass an assessment, then more grade items were added to accommodate these scores. This practice did result in a very bloated grade book, but students and instructors alike were able to quickly view their scores at any time. While each assessment had a point value assigned with it, the assessments were recorded as pass or fail, and it only takes one passing for a student to receive credit for "mastery" of that particular course assessment.

TOKENS

Many instructors developed unique variations of the token system, whereby students could earn a token that was redeemable for an additional attempt at an assessment of a course objective. Students were informed that after the initial administration of an assessment for each objective, they could redeem tokens for additional tries. The instructors kept track of how many tokens each student earned and redeemed. The tokens were credited opportunities that were recorded on the school's online learning management system, or in a system that the instructor chose, such as a gradebook.

Most instructors devised and offered a variety of ways for students to earn tokens for retake opportunities, including course-based activities, text-based homework problems, online homework problems, correcting past quiz errors, evidence of study groups, evidence of attendance at the college's tutoring center, evidence of attendance at peer supplemental tutoring sessions, attendance at workshops or symposia, volunteering at school functions, and so on. Several instructors offered free tokens at the start of the course so students would be encouraged to attempt retakes early in the semester. As a group, the instructors agreed to make the token currency easy to understand and relatively low-effort to attain. This arose from a general observation that students tended to hoard their tokens until the end of the semester. As this practice resulted students arriving at the end of the semester with a build-up of unused tokens, or an unreasonable number of retakes needed to pass a course, the instructors made sure they repeatedly encouraged students to redeem tokens early and often throughout the semester.

COURSE OBJECTIVES

One of the core tenets of Nilson's (2015) work is the arrangement and listing of course objectives to communicate clearly to the students what content knowledge and skills are critical to success. The small, chunked pieces that make up the objectives give the student a clear map towards mastering the course curriculum in discrete, well-defined steps. Instructors wrote each objective with the intent that it was measurable, discrete, and clearly understood. Some objectives incorporated higher-order thinking skills such as synthesis, analysis, and prediction so students would have to demonstrate mastery in incorporating fundamental skills or knowledge into these higher-order objectives. Additionally, faculty designed the course objectives to align with goals and objectives set out by their disciplines. This ensures that students are assessed over the same of set of objectives regardless of instructor or assessment method. Following these tenets, the instructors formulated their own manifestations of course objectives and listed them clearly in student handouts. These objectives were used as the guiding benchmarks throughout the semester, and each assessment had its aligned objective listed clearly, in bold, at the top of the page in order to focus the students' attention on what skills and knowledge were being assessed. These objectives were also listed in the course syllabi.

To be clear, the process of selecting course objectives is not unique to specifications grading. In fact, courses taught at the college level typically have a list of course objectives. What is unique about the objectives under this pedagogical framework, however, is that each objective is tied to a single assessment on which the student must demonstrate mastery to earn credit for learning said objective.

THREE VARIATIONS OF SPECIFICATIONS GRADING IMPLEMENTATION

Three distinct variations of course objective organization arose from our first round of specifications grading. All three variations allowed instructors to clearly communicate course expectations to their students. This upfront, streamlined communication between teachers and learners eliminated the common practice of interjecting subjectivity into the business of assigning course grades to students (Nilson 2015). The variations differed in two main areas: categorization of course objectives and retake policies. Some instructors divided objectives into two categories: essential objectives that had to be passed for a student to even be eligible for a passing course grade, and general objectives, a subset of which had to be passed in addition to the essentials for a student to actually earn a passing course grade. In other courses, all objectives were treated equally and there was no single objective that would prevent a student from earning a passing course grade. Finally, a few courses used module system, implementing a thematic way of grouping the objective assessments. Students needed to pass a fixed number of objectives in each module to consider the module mastered.

Each instructor implemented this new assessment scheme in a way that best met the general expectations of their specific discipline while aligning with their own beliefs as to what they valued in their classroom. The instructors understood the basic tenets of the specifications grading system, and each instructor designed practices that supported those tenets. As a result, there were three adaptations of the specifications grading assessment system that arose during this first year of implementation.

Variation A: Core Objectives and Additional Objectives

For many of the instructors, one motivating factor for determining the list of course objectives and, for those who subdivide their objectives, choosing which objectives were deemed essential, is the level of dependence subsequent coursework has on the current course. For example, Principles of Chemistry I and II, both courses required of STEM majors, were analyzed for major skills and knowledge deemed central to success in both courses. Calculus I, Calculus II, Organic Chemistry I, Organic Chemistry II, and Physics I were the other courses that were also redesigned to ensure the core student

Table II. Different variations in implementation of specifications grading

Course	Course objective organization
Survey of Chemistry I	Essential objectives and general objectives
Principles of Chemistry I	Essential objectives and general objectives
Principles of Chemistry II	General objectives only
Organic Chemistry I	Essential objectives and general objectives
Organic Chemistry II	Modules
Anatomy & Physiology	General objectives only
Algebra I	Essential objectives and general objectives
Physics I	Essential objectives and general objectives
Physics II	Essential objectives and general objectives
Calculus I	Essential objectives and general objectives
Calculus II	Modules

learning outcomes were those that were most important to success in subsequent courses. Instructors collaborated and crafted a set of statements that clearly explained what was expected of the student in terms of their performance. The intent was to eliminate the extraneous information and skills that were not fundamentally critical to a student's success in the course. This task was difficult, as it forced instructors to let go of favorite topics and be honest about what were the fundamental milestones that are considered the hallmarks of success in chemistry. Detractors of specifications grading may contend that all the content and skills in the textbook could be considered vital and critical to student learning; however, adopting the specifications grading system gave the faculty members a reason to critically question the actual amount of material that students gained (and could retain) in their courses using a traditionally assessed system.

Each course's objectives were then organized into two groups: those that were essential to the student's success in the course and subsequent courses and those that were important but not as essential to the core foundations. A listing of the course objectives for Organic Chemistry I in Table III in Appendix A illustrates the grouping. To make the course manageable, these objectives were then pared down to an even smaller number so that students and the instructors would not be overwhelmed. Each objective was crafted so that there was an action that could be measured, with enough detail and descriptive language so all parties were clear as to what was expected.

All instructors who implemented the essential and general variation of objectives taught courses that serve as the prerequisite to at least one subsequent course. The main focus of these courses is to develop the foundational knowledge and skills needed by students to be successful in the next course in the sequence. Therefore, these courses are not standalone and hence, there is a desire to ensure that students who earn the required grade for the next course have indeed mastered the core content. These instructors also recognized that there are other topics that should be covered in the course, but this information might not be deemed essential to the student's performance in the next course.

In setting up the course objectives, instructors critically evaluated the course's topics and skills to determine what a student would find necessary for the next course. For example, in chemistry, balancing a chemical equation is a skill that is assumed to have been mastered when a student starts the course Principles of Chemistry II. Likewise, in physics, manipulating vectors is a requirement for success in Introduction to Physics II.

Once these essential topics and skills were identified, they were listed clearly and succinctly in one to two sentences or phrases such that the objective contained active verbs and simply-defined objects and were measurable in an assessment. Note in Table III how there are fewer essential objectives than there are general objectives. This is because of the nature of the essential skills—students must demonstrate mastery of these objectives to even pass the course overall. A smaller number of these required objectives made it more likely that students could be successful in the course.

Within the assessment structure of the courses, each instructor committed to making these essential objectives nonnegotiable factors in their course. In other words, students must master these essential objectives to successfully exit the course. If the student is unable to successfully demonstrate mastery of one or several of the essential objectives, then the overall course grade for the lecture portion of the course is an automatic D or F. According to Nilson (2015), this practice accomplishes several things. First, it is clear to the student that the instructor holds these essential objectives as the most important aspects of the course. The extreme consequence aligned with these essential objectives communicates that these are nonnegotiable and are requirements. Unlike the traditional grading scheme where students can earn partial credit on any of the course objectives and eventually cobble together a passing grade, this formulation of specifications grading assessment eliminates the possibility of students barely passing, and therefore moving on to the next course, with subpar mastery of these essential skills and content.

A passing grade in the course does not result from simply mastering these essential objectives, however. The rest of the course material is divided into general objectives. With these objectives, there is a minimum number that must be mastered in order to earn a D, C, B, or A. Some of the instructors were reluctant to require general objectives in order to pass the course successfully, as it was believed that the essential objectives should equate to a passing grade. However, after much deliberation and reflection, the instructors noted that in a traditionally assessed course, students would usually have to demonstrate 70% mastery of the material in order to earn a C. The small number of essential objectives did not equate to 70% of the total number of objectives; therefore, the instructors agreed to include some of the general objectives into the course requirements for a C. For Organic Chemistry I, for example, the seven essential objectives plus seven of the general objectives adds up to approximately 70% of the total number of objectives.

After the first iteration of this variation in Organic Chemistry I, students determined that they could demonstrate mastery of almost all the objectives, not complete the course's final exam (worth 20% of the overall lecture grade), and still earn an overall A in the course. Because course and program assessment at the institution are tied to student performance on the final exam, this was not a tenable arrangement for the course. In the second iteration of this variation, the cut score for each grade was raised by two objectives and the final exam allowed students to add or subtract general objectives from their total based on the overall final exam score. Earning a B on the final exam would earn the student back one general objective. The rationale behind this adjustment was that, if a student could show that they had mastered the material on the final exam, then the total number of essential and general objectives mastered should also reflect this. In the third iteration of specifications grading at our institution, several instructors have been aligning mastery of certain sections of the final exam to the reward of directly aligned essential and general outcomes.

Variation B: All Equal Objectives

In this version of specifications grading assessment, instructors taught courses that do not necessarily have a subsequent course that directly depends on the student's mastery of key concepts and skills. This is not to say that the course material learned in the course are not used or important to other courses in the student's program of study. The key difference in these courses is that instructors for later courses do not necessarily assume that the student walks in with these objectives mastered, but usually do presuppose that the student has had exposure to these topics and do not need a lot of time to refresh their memory on these topics, if needed.

In discussions with these instructors, it was clear that there was difficulty in isolating key concepts or skills that were critical to the course. Many would state that all course content is important. In fact, many of the instructors stated this at the start of their onboarding experience with specifications grading. However, several did eventually acknowledge that certain objectives were nonnegotiable. Those that were not able to differentiate any of the objectives chose to align themselves to variation B of specifications grading.

For these courses, discrete, succinct course objectives were developed in a manner not unlike the way that the essential and general objectives were formulated. Some instructors chose a course objective numbering system based on the chapter in which they were discussed. Others simply numbered the objectives in the order in which they were covered in lecture. The students were introduced to the system at the start of the course and told that their overall course grade was determined by the number of course objectives in which they demonstrated mastery. There was no preference to which objectives were to be mastered for a given course grade. The listing of course objectives for Principles of Chemistry II in Table IV in Appendix A illustrates how all the objectives were ranked equally.

In reflecting upon the success of this system, several instructors noted that this arrangement approximated, but improved upon, what they had experienced in past traditionally assessed courses. It was common for students to have mastered some course content but not others. Indeed, on course assessments, instructors noted that there have been grades assigned that did not truly reflect the student's mastery; partial credit assigned from every problem or credit for other activities such as homework would then add up to an overall passing assessment score. The student did not demonstrate true mastery of any one concept, but rather partial competence of many concepts.

In this formulation of equally weighted objectives, students and instructors could clearly point to the exact knowledge and skills that were mastered and isolate the ones that needed remediation. Due to the absence of partial credit or homework credit, the instructor could communicate through grades the level of required mastery necessary to pass an objective. Because of the equal ranking of all course objectives, students had all the power to determine their course grades in that they chose which objectives to master and which ones to leave behind. In this specifications grading formulation, instructors noted that students sometimes strategized as to which objectives they should focus their energies. Many of the students used the first administration of each objective assessment as a way to determine the type of questions and level of difficulty required by the instructor for mastery of that objective. Based on that first assessment, the student then either chose to retake it for a higher grade or moved on to another assessment, corralling their efforts in learning the course objectives for which they believed they had a higher chance of success.

For example, Anatomy & Physiology I represents an allied health course for pre-nursing and exercise science students. Anatomy & Physiology I covers four major body systems which, while influencing the other systems covered in Anatomy & Physiology II when considering bodily functions, can be studied in isolation. However, in discussions with nursing schools, all systems are important, so the development of critical and general objectives was not possible. Further, there is a limited number of questions that can be asked about each system, so retakes were limited. In the first iteration of the specifications grading for Anatomy & Physiology I, 30 quizzes were implemented across lecture and lab material. The instructor discovered that 30 quizzes was too many, and in the second iteration, the number of quizzes was dropped to 20 across lecture and lab material. Quizzes represented a small chunk of material for each system covered. Quizzes consisted of 20 multiple-choice questions for each specific objective. Retakes occurred within class time similar to a traditional testing time frame except that students had control over how many and which assessments to take within the retake session. In the second iteration, an additional retake session was provided to give students one more time to take any quizzes not mastered.

Variation C: Modules

The third variation of specifications grading assessment that was developed was the *module* version. Three courses adopted this formation: Introduction to Physics I, Calculus II, and Organic Chemistry II. Instructors for these courses were of the mind that there were core skills and knowledge central to the course, but they could not truly be distilled into discrete objective statements without negatively impacting the student learning in the course. Therefore, they devised a system whereby the course objectives were arranged into modules, or large bins that organized the course objectives. Within each module, there were a number of course objectives that supported the overarching theme of that module. To demonstrate mastery of that module, students must show success on a predetermined number of the objectives within that module.

There were variations in the ways in which the modules were counted. In some courses, there were modules that were set as essential. This meant that a student must demonstrate mastery of that module to pass the course. Within these essential modules were the course objectives considered critical to the foundations of that course. In this way, the teacher effectively communicated to the students that the skills and knowledge covered in these mandatory modules were critical to the students' success in the course. It was explained to the class that, without the material included in this module, the students could not successfully learn and effectively demonstrate mastery in other modules and in future subsequent courses. An example of this version is tabulated in Table V in Appendix A, where the modules and course objectives are listed for Calculus II.

In other module-based courses, all the course modules were ranked equally and there was no communicated preference for one module over the other. In this latter case, the module system was used simply as an organizational tool to help students keep track of the course objectives and make sense of how they related to each other.

The benefit of the module system was that the instructor could then be sure that students who earned a high grade in the course would have demonstrated a minimum level of competency in each of the modules of the course. However, the way in which each student chose to show that competency was up to the student because, within the module's objectives, students had the freedom to choose in which objectives they wanted to show mastery.

Regardless of the ranking of the modules, the act of dividing up the course content into the large chunks of material, organized by overarching themes, helped students and instructors compartmentalize the large body of course content into clear, understandable blocks. It was the intention that this helped both the instructor and students understand exactly where they were in the progression of the course at any given time.

This variation of specifications grading implementation also displayed an unexpected benefit: flexibility in the numbering and frequency of objective assessments. In Calculus II, the modules were not only very different from each other in terms of content coverage, but also in terms of the specific requirements in each module and the standards to be met by the student to show mastery of that module. Content blocks that inherently contained a high number of topics or skills would have a larger number of assessments, as well as a higher number of course objectives.

DIFFERENCES AMONG INSTRUCTORS—OTHER IMPLEMENTATION MODIFICATIONS

Many other variations arose among instructors as they designed their specifications grading courses. These differences highlight some of the instructor beliefs and practices that make each teacher unique.

Assessment Grading

The way in which the instructors graded assessments varied. As is usually the case in traditionally assessed courses, instructor bias as to what constitutes an error in an assessment could be subjective. One of the major tenets of specifications grading is that student work should be judged as proficient or not proficient. Therefore, there is generally no partial credit awarded in student work that could have a major impact on the student's course grade. This tenet lends itself better to STEM courses than to non-STEM courses, as in the former there is more likely to be a correct answer. However, instructors found themselves struggling during the implementation of specifications grading in their courses because of this rigor of assessing student learning. For example, the instructor of Calculus II questioned whether the whole problem should be marked incorrect simply because the student neglected a negative sign. The derivation of the answer was completely correct, and the student demonstrated proficiency in applying the mathematical concept to the problem. Should a simple oversight be considered indication of a lack of mastery of the course objective? In these cases, instructors called upon each other to discuss and reflect upon the meaning of mastery and demonstrated proficiency. The range with which instructors allowed minor errors to be counted as incorrect or correct was wide. It also was highly dependent on the individual philosophy of the instructor.

Retakes

To address the issue of students postponing their retakes of assessments, instructors employed different strategies. One instructor required that the critical objective assessments had to be successfully passed by the midterm of the course. If students could not demonstrate mastery of the critical objectives by this date, then the student earned a failing grade for the course. This was recorded and communicated to the student with enough time for the student to withdraw from the course with a grade of W, if so desired. To support the students in meeting this deadline, the instructor offered opportunities to retake assessments at every class meeting. This was possible because the lecture and lab portions of the course were offered at the same times and in the same room. Therefore, the instructor had a lot of freedom with which to apportion instructional time to lecture, complete lab activities, or retake assessments. The instructor kept a full set of assessments at all times to accommodate any last-minute assessment retake requests by students.

Several instructors offered retake opportunities during scheduled, in-class retake sessions. These sessions would replace the traditional exam hours that commonly entail high-stakes graded events that cover numerous chapters of course content. During these retake sessions, students could choose a certain number of assessments to retake and attempt those assessments.

Instructors worked hard at generating different versions of each assessment, so students were dissuaded from simply memorizing answers from previous administrations of the assessments. Especially for courses where the course content required extensive drawings, calculations, or customized formulas, generating these assessment versions was extremely time-consuming. Therefore, almost all the instructors in this project collected and kept all retake attempts but allowed students to refer to them when requested.

Student Consultations and Interventions

Two instructors required attendance at a mandatory in-person consultation for students who had not passed all of their critical objective assessments. These instructors would discuss studying strategies and offer encouragement to the students, in the hopes that the students would avail themselves to more retake opportunities and reflect on more effective ways of mastering the course skills.

Final Exams

A majority of the instructors administered a cumulative final assessment in their courses. Usually, these graded events contained items that were common with all instructors of a subject area and level. These items are used to assess student performance across several sections and instructors. The authors are currently analyzing student performance on these assessment items between students who were enrolled in specifications grading sections and those enrolled in traditionally assessed sections.

However, some instructors incorporated the final assessment into the specifications grading scheme; some did not. For those who did incorporate the final assessment into the specifications grading scheme, a student's level of performance on the final exam could result in a student qualifying for a higher—or lower—course grade. In the situation described above, final exams, regardless of content area or section, were graded in a traditional fashion whereby students could earn partial credit for certain pieces of information or skills exhibited in their work on the exam. For instructors who did not incorporate the final exam into the specifications grading scheme, they counted the final exam score for a percentage of the student's overall grade in the course. For example, the final exam counted for 20% and the specifications grading portion counted for 80% of the overall grade.

Overall Course Grade Determination

Another example of grading variation was in how the final grade was calculated. In Anatomy & Physiology I, the instructor found that many students were consistently scoring 75% on assessments. In a traditional assessment course, this would represent a C final grade. In a specifications grading course, this would represent a final grade of F because the student would not have passed any assessments at the 80% level. As a result, the final grade was determined by the higher grade of either the number of assessments passed or the raw scores of the highest grade on the assessments. In most cases, students who were passing the majority of assessments at the 70–75% level earned a C and did so by mastering the material at that level without any additional credit derived from homework or extra credit. The specifications grading format benefitted students who were scoring 80–90% on assessments. These students who had mastered all of the assessments would earn an A but in a traditional graded course would have earned a B. Examples of grading schemes for all three variations of specifications grading are presented in Appendix B, where varying degrees of mastery translate to course grades.

CONCLUSION

In this paper, we presented an initial implementation of the specifications grading assessment method in a range of STEM courses at a 4-year public institution. Intrinsically motivated to positively impact student learning, the authors of this paper learned about the method and developed their own ways to implement the specifications grading tenets into their classes. By allowing all instructors to make their own choices about organization, policies, and practices, the primary investigator was able to bring onboard a large number of instructors in a number of different subject areas. This freedom resulted in three major variations of specifications grading implementation that spoke to each discipline's traditions and needs, to the instructors' beliefs and values, and to our student body's unique demographics and learning needs. Future directions include investigating the withdrawal and failure rates, as well as student performance on common assessment items, between courses and sections that implement specifications grading and sections that are traditionally assessed. We believe that changing the focus of college course assessment towards mastery learning and enabling students to have multiple opportunities to demonstrate that mastery, while clearly communicating to students what is exactly expected, are positive modifications that may have an impact on student learning in STEM subjects. Additionally, by forcing instructors to critically re-examine their own beliefs about assessment, grades, and course objectives, specifications grading implementation is, at the very least, a worthy exercise in professional development as educators reflect on the reasons behind traditional assessment practices that have been used for decades in classrooms.

ACKNOWLEDGEMENTS

We wish to acknowledge Georgia Gwinnett College School of Science and Technology, Dr. Thomas Mundie (Dean, GGC School of Science and Technology), Dr. Sebastien Siva, the National Science Foundation and University System of Georgia for their funding of our faculty learning community, and our family and friends for their support of this work.

REFERENCES

Block, J.H. and R.B. Burns. 1976. Mastery learning. Review of Research in Education, 4(1), 3–49.

Hamilton, L. and C. Corbett-Whittier. 2013. Using Case Study in Education Research. Sage.

Horn, C. 2003. High-stakes testing and students: stopping or perpetuating a cycle of failure? Theory Into Practice, 42(1), 31–41.

Madaus, G.F. 1991. The effects of important tests on students: implications for a national examination system. The Phi Delta Kappan, 73(3), 226–231.

Makice, K. 2012. Flipping the classroom requires more than video. Wired. Retrieved from http://www.wired.com/geekdad/2012/04/flipping-the-classroom/.

Nilson, L. 2015. Specifications grading: restoring rigor, motivating students, and saving faculty time. Stylus.

Ring, J. 2017. ConfChem conference on select 2016 BCCE presentations: specifications grading in the flipped organic classroom. Journal of Chemical Education, 94(12), 2005–2006.

Appendix A

Examples of Course Objectives by Variation Type of Specifications Grading

Table III. List of essential and general objectives for Organic Chemistry I. This is an example of variation A of our specifications grading implementations.

Type	Description of the objective
Essential	Draw Lewis structures and predicting molecular polarity
Essential	Interconvert between Lewis structures, condensed formulas, and bond-line structures with functional groups
Essential	Identify and explain charge stability (including formal charge and resonance)
Essential	Use basic nomenclature (alkanes, alkynes, and cycloalkanes)
Essential	Predict chemical structure through spectroscopic analysis
Essential	Draw and identify patterns of arrow pushing in reaction mechanisms (acid-base, substitution, addition, elimination)
Essential	Apply synthetic strategies to a one or two-step synthesis
General	Draw stable configurations of substituted cycloalkanes and alkanes in Newman projections
General	Identify and designate chirality (including molecules represented in Fisher diagrams)
General	Classify isomeric relationships
General	Predict relative acidity and basicity
General	Describe and predict reaction mechanisms and products for substitution reactions
General	Describe and Predict Reaction Mechanisms and Products for Elimination Reactions
General	Predict whether reactions occur via SN1/SN2/E1/E2 mechanisms and provide reaction products
General	Describe reaction mechanisms of addition reactions to alkenes
General	Predict reactants, reagents, and products of addition reactions to alkenes
General	Describe reaction mechanisms of addition reactions to alkynes
General	Predict reactants, reagents, and products of addition reactions to alkynes
General	Predict reactants, reagents, and products of carbon-carbon bond forming and cleaving reactions
General	Apply advanced nomenclature to complex molecules
General	Apply synthetic strategies to a multistep synthesis

Table IV. List of essential and general objectives for Principles of Chemistry II. This is an example of variation B of our specifications grading implementations.

Type	Description of the objective
General	Draw the BEST Lewis structure of any molecule or polyatomic ion.
General	Use Lewis structures to describe bond polarity
General	Use Lewis structures to determine the VESPR shapes and hybridization of central atoms in small molecules and interior atoms of larger molecules.
General	Use molecular orbital theory to determine bonding properties of diatomic molecules
General	Based on molecular shape and polarity, determine IMFs present, and relate to physical properties
General	Interpret phase diagrams of pure substances to determine phase changes
General	Describe factors affecting solubility.
General	Use macroscopic rate data to do calculations including interpreting graphical representations of reactions.
General	Write expressions for *K* and Q, use data to calculate equilibrium composition, use data to calculate *K*, and use LeChatelier's principle to predict the direction of reaction shifts.
General	Apply equilibrium concepts to acids and bases to calculate the *p*H of an aqueous solution of an acid, a base, or a salt.
General	Discuss trends in strength and structure of monoprotic and polyprotic acids.
General	Identify a good buffer, how to prepare a buffer, and calculate its *p*H (including after the addition of a strong acid or base).
General	Represent and interpret a titration graphically and calculate the *p*H at any point in the titration of a strong-strong or strong-weak system.
General	Calculate the solubility of any salt, describe what happens when a common ion or uncommon ion is added, and determine whether a salt will precipitate.
General	Calculate the macroscopic or microscopic entropy of any system or reaction and use entropy to determine whether a reaction is spontaneous.
General	Calculate the free energy of any system and use this to determine whether the process is spontaneous. Also, be able to relate free energy to equilibrium.
General	For any redox cell, balance the overall reaction, calculate the standard reduction potential, identify the species oxidized and reduced, and identify the oxidizing and reducing agents.
General	Calculate energy of a system under nonstandard conditions and relate energy (or E_o) to free energy, Q, and K.
General	Use stoichiometry to calculate the amount of reactants/products used in a redox process

Table V: List of modules and objectives for Calculus II. This is an example of variation C of our specifications grading implementations. The term GO stands for general objective.

Module number	Description of the objective
0	GO 5: Integrate using antiderivatives, the fundamental theorem of calculus, and substitution.
1	GO 6.1-6.2: Find net change given a rate of change, including finding distance and displacement given acceleration or velocity, and find the (net) area between two curves. GO 6.3–6.4: Find the volume of a solid of revolution by appropriately choosing the disk, washer, or shell method. GO 6.7: Solve physical application problems and models, to include finding ▫ mass from density, ▫ work done in moving an object or by a spring, ▫ work done to lift liquids or ropes/chains, and ▫ hydrostatic force on a surface.
2	GO 7.1: Integrate expressions by splitting up fractions, rewriting rational expressions using division, completing the square, and multiplying by 1. GO 7.2: Integrate expressions using parts, possibly multiple times, by choosing suitable expressions for u and dv. GO 7.3: Evaluate integrals of the form $\int \sin^m x \cos^n x \, dx$ and $\int \tan^m x \sec^n x \, dx$ by properly rewriting and substituting. GO 7.4: Evaluate integrals involving $a^2 - x^2$, $a^2 + x^2$, and $x^2 - a^2$ by choosing the correct trigonometric substitution. GO 7.5: Integrate rational functions by decomposing them into rational functions with irreducible denominators. GO 7.1–7.5: Evaluate integrals by choosing an applicable strategy. GO 7.8: Determine the value of integrals with infinite integrals or unbounded integrands.
3	GO 4.7: Identify and calculate indeterminate limits using algebraic manipulation and L'Hôpital's rule. GO 8.1 and 8.2: Determine whether sequences converge or diverge, and the values to which they converge (if possible). GO 8.1 and 8.3: Recognize geometric series and telescoping series, determine whether they converge or diverge, and find the values to which they converge (if applicable).

Table V (continued)

Module number	Description of the objective
3	GO 8.4.Div: Determine if a series diverges using the divergence test. GO 8.4.p: Identify and determine convergence of p-series. GO 8.4.Int: Apply the integral test to series that satisfy its conditions to determine convergence. GO 8.5.RatioRoot: Utilize the ratio and root tests to determine whether series converge or diverge. GO 8.5.Comp: Utilize the comparison and limit comparison tests to determine whether series converge or diverge. GO 8.6: Use the alternating series test to alternating series that meets its requirements to determine convergence. GO 8.3–8.6: Test series to determine whether they converge absolutely, converge conditionally, or diverge by selecting an applicable test.
4	GO 9.1: Create approximating polynomials for functions, use these polynomials to approximate function values, and estimate the error of the approximation. GO 9.2: Make power series for functions using algebra, differentiation, and integration on other power series, and find the intervals and radii of convergence of these new series. GO 9.3: Construct and manipulate Maclaurin and Taylor series for functions. GO 9.4: Find limits of, differentiate, and integrate functions using Taylor series.
5	GO 10.1: Plot parametric curves, rewrite parametric equations as equations in x and y by eliminating the parameter, formulate parametric equations for equations in x and y, and differentiate parametric equations. GO 10.2: Plot polar curves, convert points and equations between rectangular and polar coordinates.

Appendix B

Examples of Course Grading Schemes

Table VI. Description of student mastery requirements for grades in a specifications grading course implementing variation A

Course grade	Requirement
A	Pass 7 essential objectives and 16–17 general objectives
B	Pass 7 essential objectives and 14–15 general objectives
C	Pass 7 essential objectives and 12–13 general objectives
D	Pass 7 essential objectives and 10–11 general objectives
F	[Pass less than 7 essential objectives] or [pass 7 essential objectives and less than 10 general objectives]

Table VII. Description of student mastery requirements for grades in a specifications grading course implementing variation B

Course grade	Requirement
A	Pass 18 essential objectives and 16–17 general objectives
B	Pass 16 essential objectives and 14–15 general objectives
C	Pass 14 essential objectives and 12–13 general objectives
D	Pass 12 essential objectives and 10–11 general objectives
F	Pass 11 or less essential objectives

Table VIII. Description of student mastery requirements for grades in a specifications grading course implementing variation C

Course grade	Requirement
A	Pass essential module 1 and 5 general modules, student choice
B	Pass essential module 1 and 4 general modules, student choice
C	Pass essential module 1 and 3 general modules, student choice
D	Pass essential module 1 and 2 general modules, student choice
F	[Do not pass essential module 1] or [pass essential module 1 and (1 or none of the general modules, student choice]