

Review

Delving into the Design and Implementation of Specifications Grading Systems in Higher Education

William J. Howitz ¹, Kate J. McKnelly ² and Renée D. Link ^{3,*}¹ Department of Chemistry, University of Minnesota, Minneapolis, MN 55455, USA; whowitz@umn.edu² Department of Chemistry, Emory University, Atlanta, GA 30322, USA; kate.joy.mcknelly@emory.edu³ Department of Chemistry, University of California—Irvine, Irvine, CA 92697, USA

* Correspondence: rlink@uci.edu

Abstract: Specifications grading is an alternative grading system that has been used with increasing frequency in higher education. Since first introduced by Linda Nilson in 2014, more than 90 publications on the design and implementation of specifications grading systems have been published. This work presents a systematic review of the current literature to analyze the variety of ways specifications grading systems are executed, including the diverse design and implementation considerations, as well as to present and discuss emergent themes. We analyzed 90 publications and present their relevant findings in the results. The following databases were last searched on 5 October 2024 for publications: IEEE Xplore, ACS Publications, ASEE PEER, PER, Scopus, ERIC, ACM, ScienceDirect, and Web of Science. All peer-reviewed journal articles, conference proceedings, and book chapters that implemented at least two structural features of specifications grading in an undergraduate or graduate course were included in this review. Theses, dissertations, conference abstracts, posters, workshops, blogs, opinion pieces, social media exchanges, and content provided on websites were not included. Additionally, reports of specifications grading systems in K-12 courses or those that only presented the design and/or implementation of less than two structural features of the grading system were similarly excluded. Our findings from the literature reveal that the following themes emerge from educators who use specifications grading: instructor commentary on time investment, academic performance, and student reactions to specifications grading. This review provides a resource for those interested in exploring this alternative grading system, and the emergent themes indicate that there are ripe opportunities for future study.



Academic Editors: Marius Boboc and Xinqiao Liu

Received: 21 October 2024

Revised: 20 December 2024

Accepted: 27 December 2024

Published: 14 January 2025

Citation: Howitz, W. J., McKnelly, K. J., & Link, R. D. (2025). Delving into the Design and Implementation of Specifications Grading Systems in Higher Education. *Education Sciences*, 15(1), 83. <https://doi.org/10.3390/educsci15010083>

Copyright: © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: specifications grading; alternative grading; assessment

1. Introduction

What is a course grade meant to represent? In U.S. higher education, course letter grades are determined frequently by the number of overall points a student accumulates throughout a term culminating with a final letter grade assigned on an A–F scale (Schinske & Tanner, 2014; Brookhart et al., 2016). The A–F scale has been the dominant grading scheme in the U.S. since the mid-1900s (Durm, 1993). The first documented use of the A–F scale defined each letter grade as follows: A is excellent, B is good, C is fair, D is passed, and F is failed. However, what each letter grade means at each institution, course, etc. is not standardized.

The importance of letter grades to earn degrees, maintain scholarships, and gain access to graduate and professional programs incentivizes students to focus on accumulating points (an extrinsic motivation) rather than on learning (an intrinsic motivation) (Schinske

& Tanner, 2014). Furthermore, when we assign points and partial credit to student work and sum the points, there is often no direct connection between the grade students earn and what course learning outcomes (LOs) they have achieved (Brookhart et al., 2016; Nilson & Stanny, 2014). Grades become a ranking of students against each other rather than an indication of the achievement of the LOs by each individual student (Brookhart et al., 2016). This ranking issue is exacerbated in courses that employ curves to determine how many points correspond to a letter grade (Seymour, 1997; Bowen & Cooper, 2022).

The problematic use of points to assign grades causes further challenges. The focus on accumulating points sets up an antagonistic relationship between the instructor and students, instills competition between students, and amplifies student and faculty stress, anxiety, and mental health issues (Nilson & Stanny, 2014; Eyster, 2024; Hammoudi Halat et al., 2023). A feature of points-based grading systems is the inclusion of partial credit allotment. This inclusion increases the time-consuming faculty activity of meeting with students who argue for partial credit. Another feature of points-based grading systems is often including high-stakes assignments with no opportunities for showing proficiency or competency with feedback and opportunities to try again. This feature benefits students who come from more privileged backgrounds and penalizes students from minoritized groups (Smeding et al., 2013).

Alternative grading practices, including mastery grading, standards-based grading (SBG), contract grading, and specifications grading, have been developed to address these challenges with traditional, points-based grading systems (Clark & Talbert, 2023).

In a mastery grading system, a student's work on an assessment of a topic or learning outcome must meet the instructor's established performance threshold before the student can move on to a new topic or learning outcome (Bloom, 1968; Kulik et al., 1990). Because the assessment does not count toward the student's final course grade until the work meets the defined threshold, students must be provided multiple attempts without penalty. These reassessment opportunities lower the stakes of assignments and provide an incentive for students to use instructor-provided feedback to demonstrate learning over time.

In a standards-based grading (SBG) system, a student's performance on an assessment is evaluated with respect to one or more of the course learning outcomes, and the number and/or type of learning outcomes a student meets determines their final letter grade (Knight & Cooper, 2019). Partial credit and points are removed entirely in an SBG system. The incentive for students to focus on accumulating points is eliminated, minimizing antagonistic interactions between the instructor and students over the points and partial credit allocated to an assignment. Similarly to mastery grading, SBG also permits multiple attempts without penalty, encouraging students to leverage instructor feedback provided on assignments to address the gaps in their knowledge to meet the course learning outcomes.

Contract grading establishes a collaborative relationship between the student and the instructor, returning agency to students by giving them more control over the way in which their course grade is determined. In a contract grading system, each student negotiates a contract with the instructor that specifies what assessments they need to complete and how the assessments must be completed to earn a predetermined final letter grade chosen by the student (Hassencahl, 1979; Taylor, 1980; Beare, 1986). The transparency inherent in contract grading provides students with clear expectations and a roadmap to achieve their desired final letter grade.

The alternative grading systems described above employ a variety of frameworks, including mindset theory (Dweck, 2013; Harsy et al., 2021; Lewis, 2022), achievement goal orientation theory (Elliot & McGregor, 2001; Elsinger & Lewis, 2020), and self-efficacy theory (Bandura, 1986; Carberry et al., 2012). However, the extent to which these frameworks and others are discussed in the literature on alternative grading systems in U.S. higher

education is limited (Hackerson et al., 2024). Clark and Talbert describe the “four pillars of alternative grading” as a practical framework that can support instructors in designing effective alternative grading systems (Clark & Talbert, 2023). Their framework identifies the features of alternative grading systems that focus on student learning over time: (1) clearly defined standards describing acceptable evidence of learning, (2) actionable feedback used by students to improve their learning, (3) marks on student work that indicate progress toward meeting standards, and (4) reassessment of student work without penalty to provide students with opportunities to use the actionable feedback.

Specifications grading, first reported by Nilson in 2014 (Nilson & Stanny, 2014), combines aspects of the previous grading systems and incorporates the four pillars described by Clark and Talbert (Elkins, 2016; Howitz et al., 2021). Leslie and Lundblom provide the following summary of the principles underlying specifications grading (Leslie & Lundblom, 2020).

The core principles of specifications grading are:

- Course assignments are aligned with course learning objectives.
- Expectations (“specifications”) are clear.
- Students decide what grade they aim for (self-imposing learning demands consistent with the grade).
- Feedback relates expectations to performance.
- Defined (and limited) options are provided for revisions.
- Assignments are completed at a clearly defined level of performance (e.g., corresponding to a grade of B or C) to demonstrate competency.
- Advanced learning options in breadth and/or depth are offered for self-motivated students.

While the principles delineated above describe the underpinnings of specifications grading, these principles must be operationalized. In practice, the following structural components comprise specifications grading systems (Nilson & Stanny, 2014):

- At the course level, students are provided with defined grade bundles that clearly delineate what assignments they need to complete and at what level to earn their chosen letter grade.
- At the assignment level, where assignments encompass all work submitted by a student, including homework, quizzes, papers, exams, etc., students are provided clear rubrics that contain the specifications required for the assignments and the set threshold they must achieve to demonstrate competency. Partial credit is not available.
- Students are provided with one or more mechanisms by which they are able to revise work that does not meet the required specifications.
- Optionally, a token system that can be used to limit opportunities for revisions and provide flexibility to students in how they navigate the course is provided.

Numerous examples of specifications grading systems have been published in the primary teaching and education literature of many individual disciplines. A recent scoping review by Hackerson et al. highlights alternative grading systems in science, technology, engineering, and mathematics (STEM) courses (Hackerson et al., 2024). Harrington et al. examined the body of literature on contract grading and specifications grading in computer science courses (Harrington et al., 2024). However, reviews of the existing literature focused solely on describing implementations of specifications grading across all disciplines do not exist. Here we provide a review of publications describing the design and implementation of specifications grading systems in all disciplines across higher education. Our aims in this review are (1) to provide a resource for instructors designing their own specifications grading systems and (2) to build a roadmap for education researchers to facilitate collabo-

ration with practitioners to study outcomes and impacts of specifications grading in higher education based on emergent themes in the current literature.

2. Methods

2.1. Research Questions

Our research was guided by the following questions regarding specifications grading systems described in the literature:

1. What is the current landscape of peer-reviewed literature describing implementations of specifications grading in higher education?
2. What are the structures of the specifications grading system implementations currently described in the literature?
3. What themes have emerged in the literature on specifications grading that have not yet been studied systematically?

2.2. Article Selection and Analysis

We conducted a literature search following the PRISMA guidelines using the keywords “specifications grading,” “specs grading,” “specifications-based grading,” and “alternative grading” to find relevant publications ranging from October 2014 through September 2024 (Figure 1). October 2014 was chosen as the start date because it was when Linda Nilson’s book on the subject was published. The databases that were searched included the following: Web of Science, Scopus, the Education Resources Information Center (ERIC), Institute of Electrical and Electronics Engineers (IEEE) Xplore, ScienceDirect, the American Chemical Society (ACS) Publications, the Association for Computing Machinery (ACM) Digital Library, the American Society for Engineering Education (ASEE) Papers on Engineering Education Repository (PEER), and Physics Education Research (PER) Central. The number of publications identified using the search terms from each database is specified in Table 1. Several publications were identified through more than one database. Citation searching was used to locate any publications not found through the aforementioned databases. A total of 21 additional publications that met the inclusion criteria described below were found by this method.

Table 1. Total number of publications found and included in this review, separated by the search database in which the publications were found. The total number of included publications below is greater than the number of publications in the corpus because multiple publications were found in more than one database.

Database	Publications Identified Using Search Terms After Removing Duplicates	Publications Included in Review Corpus After Applying Inclusion and Exclusion Criteria
IEEE Xplore	1	1
ACS Publications	67	17
ASEE PEER	123	17
PER	2	0
Scopus	96	56
ERIC	22	14
ACM	18	3
ScienceDirect	3	1
Web of Science	40	27

Because this review was intended to characterize the landscape of empirical research on specifications grading in the higher education (undergraduate- or graduate-level) setting,

we focused on peer-reviewed publications (Table 2). Theses, dissertations, and informal means of communication such as blogs, opinion pieces, social media exchanges, and websites were excluded from the corpus as they are not peer-reviewed. We limited the scope of this review to publications that describe the design and implementation of specifications grading systems. Publications that discussed the design of a specifications grading system but did not implement it or that referenced the implementation of specifications grading without additional detail or expansion were excluded from the analysis. Specifically, each publication was only included if commentary on the implementation of at least two of the four structural features of specifications grading—grade bundles, rubrics with specifications and defined passing thresholds, opportunities to revise and resubmit work, and a token system—was included. As such, journal articles, conference papers, and book chapters were included, but conference abstracts, posters, and workshops were excluded as they lacked sufficient detail to meet the inclusion criteria. The total number of publications included in the corpus was 90. The full corpus of manuscripts included in this review is available in Table S1.

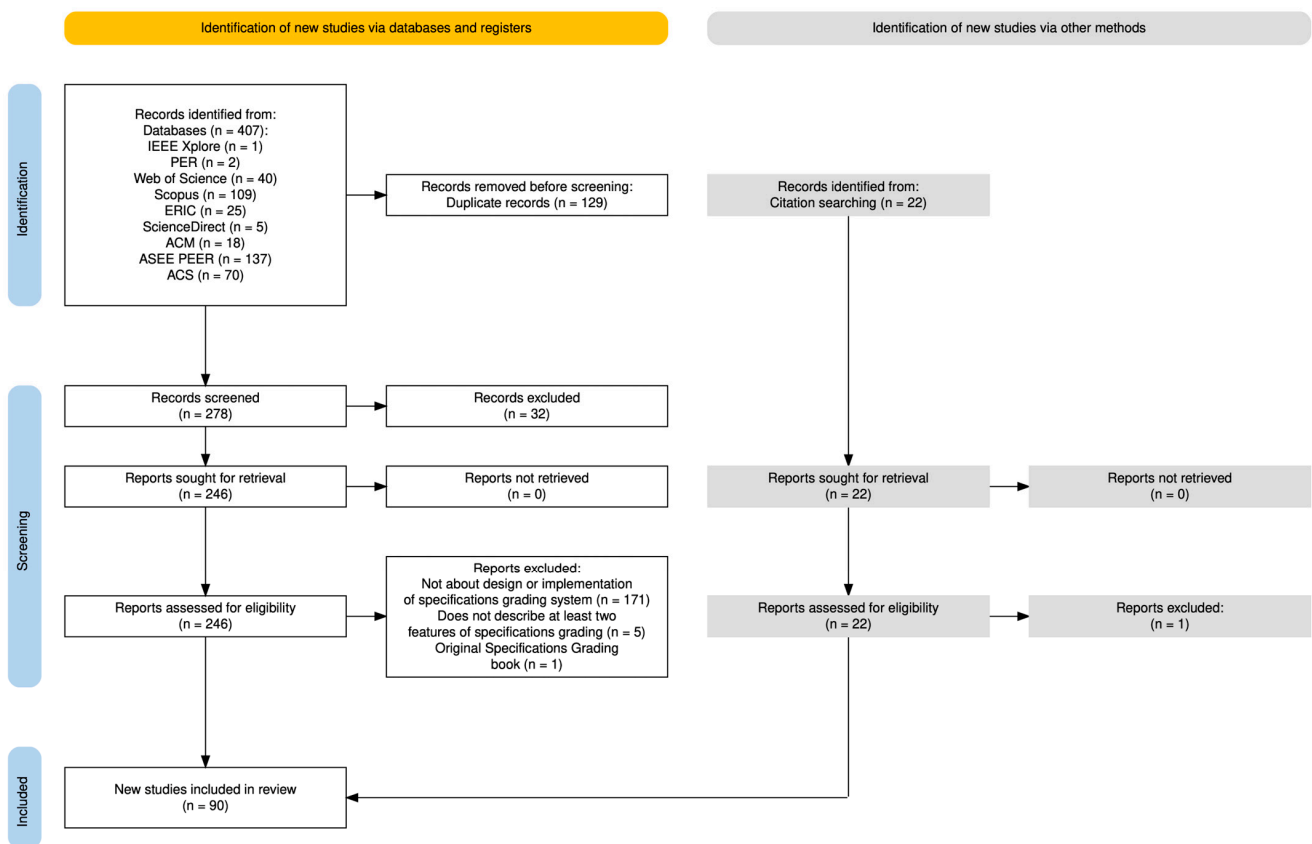


Figure 1. Diagram of this study corpus selection process generated with the PRISMA flow diagram Shiny app (Haddaway et al., 2022).

Table 2. Inclusion and exclusion criteria used in the analysis conducted for this review.

Inclusion Criteria	Exclusion Criteria
Peer-reviewed empirical research, including journal articles, conference proceedings, and book chapters	Theses, dissertations, conference abstracts, posters, workshops, blogs, opinion pieces, social media exchanges, websites
Implementation of at least two structural features of specification grading	Design only or implementation of less than two structural features of specifications grading
Undergraduate and graduate populations	K-12 populations

Two reviewers screened each publication for inclusion independently. Data from each publication were extracted by one reviewer and verified by a second reviewer. Because most publications included in the review were descriptive studies conducted by instructors on their own courses, the risk of bias in individual studies was not assessed. Characteristics of courses were taken directly from the descriptions included in the publications, and disciplines were assigned based on the name or description of the course(s).

3. Results

3.1. Publication Trends

Following Nilson's book in 2014, the first publications on the design and implementation of specifications grading appeared in 2016. In every year after 2016, there have been at least five publications, with 2023 as the year with the most publications at 22 (Figure 2). From January to 30 September 2024, a total of 12 publications were released.

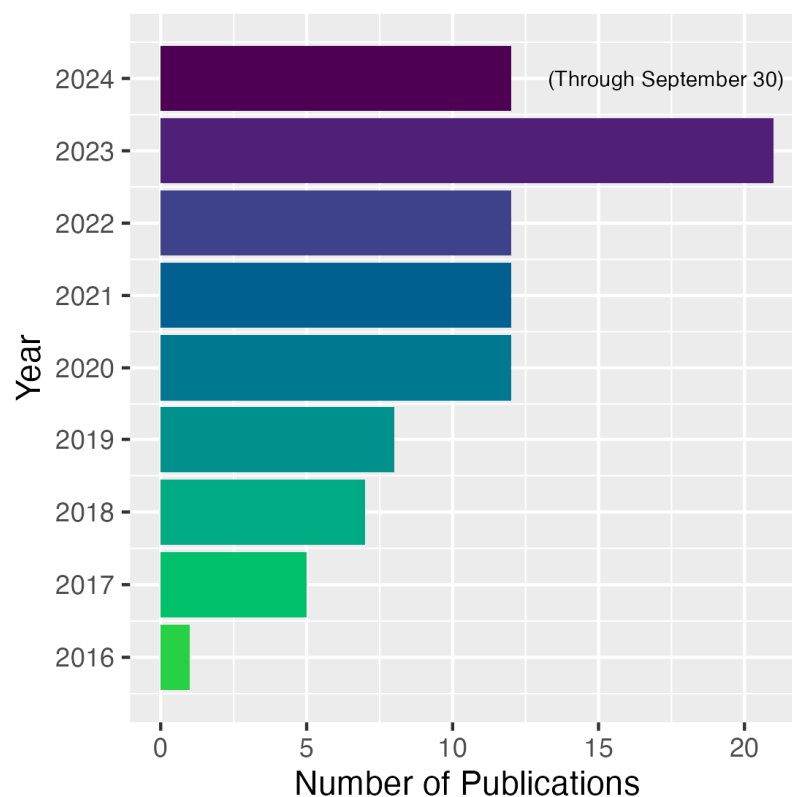


Figure 2. Number of peer-reviewed publications describing specifications grading systems published annually, January 2016–September 2024. There were no publications in 2014 or 2015 aside from Nilson's book.

STEM disciplines are most represented in publications describing specifications grading systems (Figure 3). Chemistry represents the greatest number of publications with 20, followed by engineering with 19, computer science with 10, and mathematics with 8. All other disciplines have four or fewer publications, and disciplines within the humanities are least represented. The type of publications also varies by discipline. While journal articles are most common amongst nearly all disciplines, conference papers are the only type of publication from engineering. Book chapters are almost exclusive to chemistry, with the only exceptions being one publication from information literacy, one from French studies, and one from computer science. These discrepancies in types of dissemination are tied to disciplinary norms. In engineering and computer science, peer-reviewed conference papers

are more common than journal articles, and the American Chemical Society specifically publishes peer-reviewed books as an alternative form of dissemination.

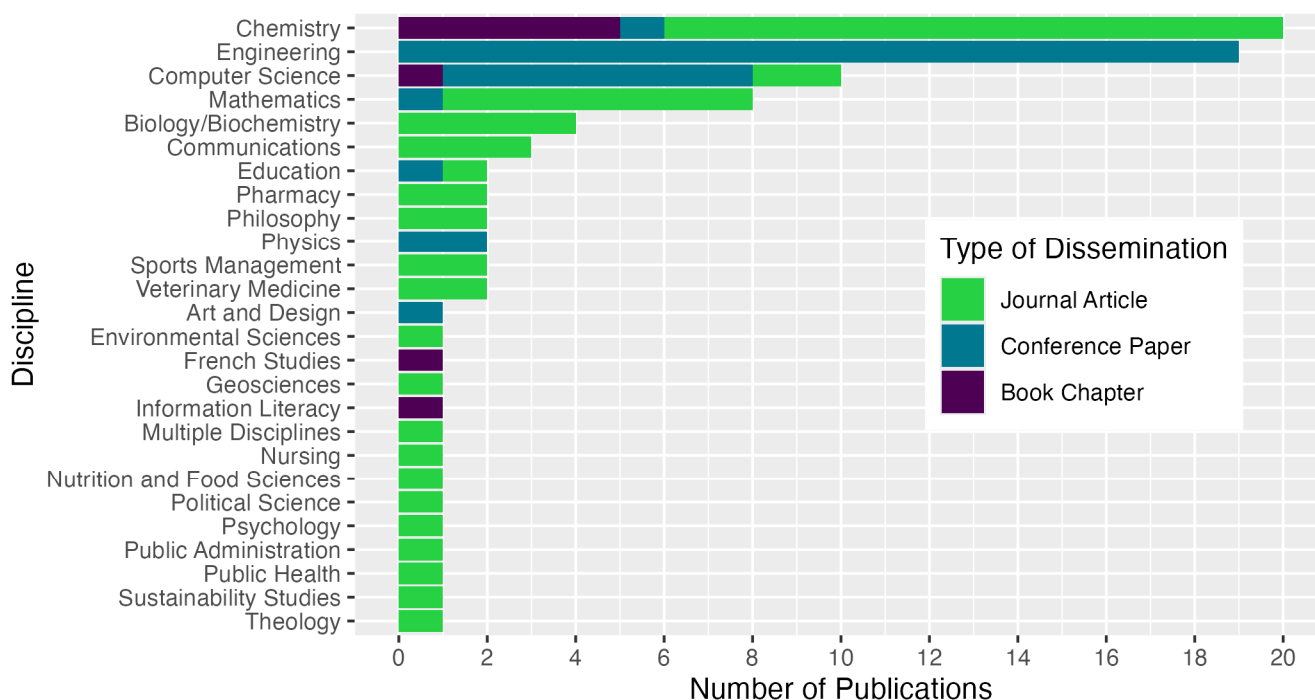


Figure 3. Peer-reviewed publications describing specifications grading systems, organized by discipline and type of dissemination.

Descriptions of course sizes vary from publication to publication. Some include descriptions of a single course or multiple courses with no information about size, a single course with information about size, or multiple courses with information about size. Course sizes were extracted from the descriptions provided in the publications and were grouped into five categories: 35 or fewer, 36–60, 61–100, 101–999, and 1000 or greater. Of the courses described that include information about course size, 63% (56 courses) are small courses with total enrollments of 35 or fewer students (Figure 4). The commonality of this course size is unsurprising, as the majority of these publications describe courses taught at smaller colleges and universities. The majority of published implementations of specifications grading in courses with enrollments of 100 students or greater describe introductory-level courses, mostly at large, four-year institutions. Although most published examples of specifications grading describe courses at the undergraduate level, including independent study or capstone courses (Earl, 2021; Mendez, 2024; Martin et al., 2021; Gargac, 2022a; Fernandez et al., 2020), a small number of graduate and professional courses do appear (Blodgett, 2017; Dennen & Bagdy, 2020; Gay & Poproski, 2023; Hofmeister et al., 2022a, 2022b; Jones, 2020; Quintana & Quintana, 2020; Walden, 2022; Joseph et al., 2023; Joshi, 2023; Dupree et al., 2024; Moster & Zingales, 2024; Santucci & Golas, 2023). Of the 90 publications included in the review corpus, 20 publications describe implementations of specifications grading in online or hybrid modality courses (Elkins, 2016; Dennen & Bagdy, 2020; Gay & Poproski, 2023; Quintana & Quintana, 2020; Moster & Zingales, 2024; Santucci & Golas, 2023; Mendez, 2018a, 2019; Houseknecht & Bates, 2020; Shields et al., 2019; Gestwicki, 2021; Gratwick et al., 2020; Wasniewski et al., 2021; Evensen, 2022; Gargac, 2022b; Dabney & VanDerWoude, 2023; Suresh, 2023; Closser et al., 2024; Kinnear et al., 2022; Johanesen et al., 2024).

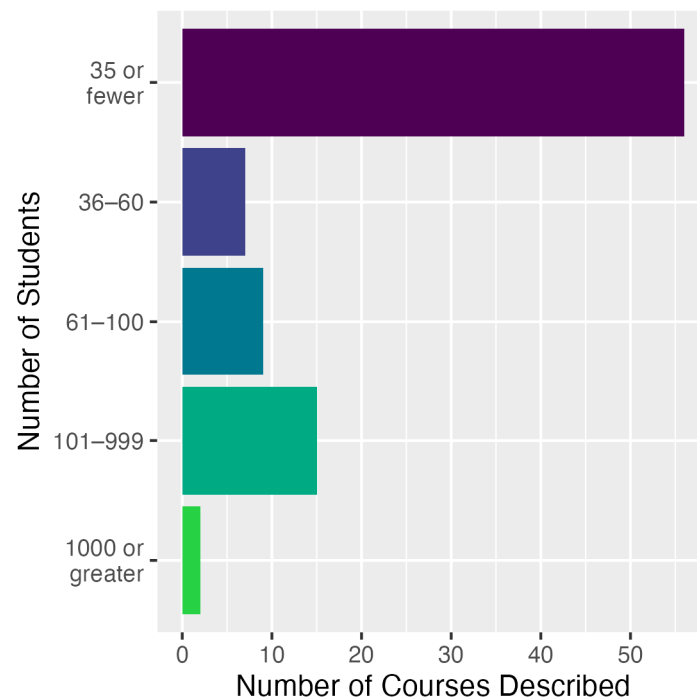


Figure 4. Peer-reviewed publications describing specifications grading systems organized by course size. The total number of courses described does not match the total number of publications included in the corpus, as 11 publications describe the design and implementation of specifications grading in more than one course.

3.2. Structures of Specifications Grading Systems

The four structural components of specifications grading—grade bundles, rubrics with specifications and defined passing thresholds, opportunities to revise and resubmit work, and a token system—are incorporated into courses in a variety of ways and to different degrees. In this section we discuss three of the four structural components independently, giving context for how they have been implemented. The extent to which retakes and revisions are allowed varies, depending on how the other structural components were designed. As such, a conversation about this component will be threaded throughout the next three subsections.

3.2.1. Grade Bundles

A core principle of specifications grading systems is to align assignments with course LOs, and a key structure of these systems is bundling these assignments together to determine course letter grades (Nilson & Stanny, 2014). Four main methods of bundling, which we call configurations, emerged in the analysis of peer-reviewed descriptions of specifications grading. Tsoi et al. originally described three main configurations (called “implementations”) in the context of lecture courses: core and additional LOs, all equal LOs, and modules (Tsoi et al., 2019). A fourth configuration best described as all equal LOs with repetition and/or complexity (ELORC) arose in the context of laboratory and writing courses.

In the core and additional LOs configuration, all course LOs are sorted into “core” and “additional” categories. “Core” LOs are those deemed by the instructor to be essential for earning a grade of C or above. All other LOs are categorized as “additional,” and meeting a larger subset of these LOs is required to earn B or A grades. Frequently this configuration is used in introductory lecture courses that serve as prerequisites, where having fundamental knowledge and skills is essential for success in subsequent courses.

While the terms “core” and “additional” arose from Tsoi et al., a number of publications use different terminology, such as core and advanced LOs or essential and general LOs, to refer to this same configuration (Carlisle, 2020; Ring, 2017). An example of a course with the core and additional LOs configuration is Carlisle’s discrete and combinatorial algebra course, in which each of the six major subject areas is divided into a core topic and an advanced topic (Carlisle, 2020). To earn at least a C grade, students must pass all six core topics and one advanced topic. If a student wishes to earn a B grade, they need to pass an additional two advanced topics, and if a student wishes to earn an A grade, they need to pass an additional four advanced topics. Similarly, LOs in Biers’s first-year French course are classified as basic or advanced and are grouped into three main categories: proficiency, cultural knowledge, and metalinguistic awareness (Biers, 2022). To earn at least a C grade, students must pass all of the basic LOs in all three categories. If a student wishes to earn a B grade, they also need to pass advanced LOs from one of the three categories (or two of the categories if they wish to earn an A).

In an all equal LOs configuration, all course LOs are given equal priority. Students earn their letter grade based on the total number of LOs met. Frequently this configuration is used in advanced or elective courses. For example, in Carlisle’s differential equations course, the content is split into 20 “problem topics.” Students must pass 14 of the topics to earn a C, 17 to earn a B, and all 20 to earn an A (Carlisle, 2020). Similarly, in Mendez’s sophomore-level thermodynamics course, there are 15 LOs, and each is assessed by one quiz (Mendez, 2018a). Students must pass 11 quizzes to earn a C, 13 for a B, and all 15 for an A.

In the modules configuration, thematically related course LOs are binned together into modules. To pass a module, students must meet a predetermined number of the LOs in that module. According to Tsoi et al., the module configuration is used when the “. . .skills and knowledge central to the course. . .(cannot). . .be distilled into discrete objective statements without negatively impacting the student learning in the course (Tsoi et al., 2019).” While Tsoi et al. list modules as a third, standalone configuration, the modules configuration could be considered a subcategory of either the “core and additional LOs” or the “all equal LOs” configurations, depending on how the instructor chooses to bundle the module for letter grades. Tsoi et al. indicate that modules may be classified as “essential” (consistent with the core and additional LOs configuration), while in other cases, modules may be ranked equally (consistent with the all equal LOs configuration).

The three configurations described by Tsoi et al. all emerged in the context of lecture courses, whereas the fourth ELORC configuration emerged in the context of laboratory and writing courses. Lecture courses generally have a large number of LOs to cover all of the content knowledge students must learn, especially at the introductory level. In contrast to lecture courses, laboratory and writing courses tend to be more focused on the practical application of knowledge and the development of technical skills. In these types of courses, there are often fewer course LOs, and they are often bundled in ways that require students to meet the same LO(s) multiple times and potentially at varying levels of complexity to earn higher grades. In the context of a laboratory course, LOs align with students both applying knowledge they have learned and developing practical technical skills to obtain and analyze data. In the context of a writing course, LOs focus on students practicing the processes of brainstorming, outlining, drafting, editing, and polishing various pieces of writing. An example of a course with this configuration is McKnelly et al.’s Writing for Chemists course (McKnelly et al., 2021). In this course, there are four large writing assignments that all assess the same course LO, “Students will be able to create professional papers, proposals, reports, and other forms of scientific writing.” These four assignments also assess different amounts of additional LOs. Students must earn a low pass on three of

the assignments to earn a C, a low pass on all four assignments to earn a B, and a high pass on three and a low pass on one to earn an A. (Passing threshold levels will be discussed in the next section).

Although Nilson's original introduction of specifications grading specifically called for designing assignments and grade bundles that are aligned with course LOs, this alignment is not always clearly demonstrated in publications describing specifications grading systems. Some instructors described grade bundles (and thus implied configurations) based on assignments but did not clearly describe how these assignments were aligned with course LOs. Tsoi et al. proposed the configurations with the assumption that individual assignments are mapped to individual course LOs. However, in some of the courses described in the specifications grading literature, one assignment may correspond to one LO, one assignment may correspond to multiple LOs or multiple assignments may correspond to the same LO. These discrepancies are reflected in Yik et al.'s analysis of specifications grading in chemistry courses in which the grade bundles observed for some courses reflect a focus on specific assignments rather than specific LOs (Yik et al., 2024). Without clear descriptions of alignment between assignments and LOs in some publications included in our review, it was not always possible to determine what type of LO-focused configuration was being employed.

3.2.2. Rubrics with Specifications and Passing Thresholds

Within a specifications grading system, the specifications are embedded as the rubric criteria for assignments. In general, student work is evaluated against each rubric criterion or specification, and then the assignment outcome is determined by performance across all of the specifications set for the assignment. The most common way to determine if students have met a specification is using a binary system, which in practice generally appears as a student earning credit for a rubric criterion (specification) or not. In contrast, a student's overall assignment may be evaluated using a 2-level, 3-level, or 4-level system. Other variations for both specification and assignment evaluations exist, but these are more complex (Gargac, 2022a, 2022b; Gestwicki, 2021; Toledo & Dubas, 2017; Mirsky, 2018; Henriksen et al., 2020; Cosoroaba, 2020; Donato & Marsh, 2023; Rupakheti et al., 2018). It is not uncommon for different assignments in a course to have different evaluation outcome types as needed. Descriptions and examples of binary specifications rubrics with 2-level, 3-level, and 4-level assignment outcomes follow.

The majority of publications included in this review use assignments with binary specifications and 2-level assignment outcomes. In this approach, a student's work either does or does not meet an individual specification. The instructor sets a threshold, i.e., a number of specifications that must be met for the overall assignment to earn credit. Instructors may also set some specifications as "required" so that the assignment does not earn credit if those "required" specifications are not met, regardless of how many others are met. Wording for the assignment outcomes varies, but some of the common phrases that have been used are pass/fail, satisfactory/unsatisfactory, satisfactory/needs revision, meets specifications/does not meet specifications, accept/revise, or complete/incomplete (Howitz et al., 2021; Blodgett, 2017; Jones, 2020; Santucci & Golas, 2023; Blackstone & Oldmixon, 2019; Lillard & Taggart, 2022).

A closely related, but less common approach is to use binary specifications with 3-level assignment outcomes, in which two different thresholds are set. Depending on how the instructor sets up their specifications grading system, students may only earn credit for an assignment if they meet the higher of the two thresholds (Williams, 2018), or they may earn credit as long as they meet at least one of the two thresholds (McKnelly et al., 2021). McKnelly et al.'s Writing for Chemists course, described previously, provides an example

of the latter case. Grades are bundled based not only on the number of times a student is able to pass a large writing assignment but also at what passing threshold (a high pass, low pass, or needs revision). For example, students must pass all four writing assignments to earn an A or a B, but for the B grade, students must earn low passes or better on all four assignments, whereas for the A grade, students need to earn at least three high passes and may earn only one low pass.

Binary specifications rubrics may also be combined with 4-level assignment outcomes based on the EMRF rubric that allows for two levels of passing work (E: excellent, M: meets expectations) and two levels of work that does not pass (R: needs revision, F: fragmentary) (Stutzman & Race, 2004). An advantage of this 4-level system is the ability to differentiate the quality of student work within the passing and not passing categories. In most cases, the F designation has been replaced with N (not assessable) to overcome students' association of the F with failing, resulting in the more frequently used acronym, EMRN (Talbert (n.d.)). In Mendez's sophomore-level thermodynamics course, described previously, each quiz is assessed using the EMRN rubric (Mendez, 2018a). While students need to pass 11 quizzes to earn a C, 13 for a B, and 15 for an A, the threshold at which students pass the quiz also matters. For example, to earn a C, none of the quizzes need to be assessed as excellent (E), but to earn a B, 7 of the 13 quizzes need to be assessed as excellent (E), and to earn an A, 11 of the 15 quizzes must be assessed as excellent (E).

3.2.3. Token Systems

Of the 90 publications selected for this review, 46 specifically commented on the inclusion of a token system, and three commented specifically on choosing not to include a token system. The remaining 41 publications do not include any specific information about token systems. Instructors chose to provide tokens in their specifications grading systems by providing a set number of tokens to each student at the beginning of the course (seven examples) (Fernandez et al., 2020; Joshi, 2023; Dabney & VanDerWoude, 2023; Lillard & Taggart, 2022; Vitale & Concepción, 2021; Johnson, 2023; Tamés, 2021), by providing opportunities for students to earn tokens throughout the term of a course (fourteen examples) (Howitz et al., 2021; Joseph et al., 2023; Evensen, 2022; Hunter et al., 2022; Ludwigsen, 2017; Helmke, 2019; Martin, 2019; Prasad, 2020; McKnelly et al., 2023; Kelz et al., 2023; Rojas & Quan, 2023; Saluga et al., 2023; Howitz et al., 2023; Cerkez, 2024), or by combining both aforementioned approaches (seventeen examples) (Fernandez et al., 2020; Dennen & Bagdy, 2020; Moster & Zingales, 2024; Suresh, 2023; Closser et al., 2024; Johanesen et al., 2024; Tsoi et al., 2019; Carlisle, 2020; McKnelly et al., 2021; Cosoroaba, 2020; Blackstone & Oldmixon, 2019; Williams, 2018; Brown & Kennedy, 2022; Fierke, 2024; Yang & Korsnack, 2024; Mio, 2024; Copp, 2024). In the earn-only or combination approaches, students were given opportunities to earn tokens by a variety of means, such as completing metacognitive reflection assignments or completing low-stakes course activities such as readings and homework.

In all token systems described, students were given the option to use tokens for additional attempts at assignments (either through revision or attempting a new version of an assignment such as a quiz), for flexibility on assignment deadlines, or both. In six examples, students could also use tokens to earn back credit on a low-stakes assignment that they missed originally or to replace attendance credit lost for missing class when attendance was required (Dennen & Bagdy, 2020; McKnelly et al., 2021; Vitale & Concepción, 2021; McKnelly et al., 2023; Mio, 2024; Kiefer & Earle, 2023). In all cases, the choice of how to use tokens was left with the students, although instructors did provide encouragement to use tokens as needed and reminders of how to do so.

Few authors provided details on how many tokens they chose to make available, the rates at which students used tokens or the specific methods by which instructors tracked tokens. However, the few examples provided do give useful guidance for instructors considering implementing a token system in their courses. Hunter et al. suggest determining the number of tokens to provide by counting the number of high-stakes assignments and adding one (Hunter et al., 2022). Vitale and Concepción suggest a similar approach—providing tokens that correspond to the number of high-stakes assignments plus or minus one (Vitale & Concepción, 2021).

Based on the limited examples provided, students do not appear to run out of tokens. In a first-year engineering course, 69% of students used at least one token in the course, and a grand total of 41% of available tokens in the course were used by the end (Fernandez et al., 2020). On any given assignment in this engineering course, at least one student chose to use a token and the number of students who used a token on an assignment increased as the term progressed. Kelz et al. found that only four of ninety-nine students used all of their tokens (Kelz et al., 2023). Dennen and Bagdy indicated that few students used all of the provided tokens, and those who did chose specifically to do so (Dennen & Bagdy, 2020). With one exception in which students can trade tokens for candy (Williams, 2018), authors did not report rewarding students for tokens left over at the end of a course. Despite this lack of reward, two publications reported that students display token-hoarding behaviors (Tsoi et al., 2019; Martin, 2019). One instance of “gaming the system” was described, in which a team of students working on a group project opted to submit work that did not display a good-faith effort at completion and then replace the missing credit for that work with a token (Fernandez et al., 2020). The authors stated that this “gaming the system” was performed to gain more time to work on another aspect of the course.

Current learning management system (LMS) options are not designed to support specifications grading and do not provide tools to support a token system. However, a placeholder assignment (Martin, 2019) or ungraded quizzes in the course LMS (Dennen & Bagdy, 2020) can be used to track students’ tokens. Alternatively, token usage can be tracked using an online form with a spreadsheet alone (Fernandez et al., 2020; Blackstone & Oldmixon, 2019; Kelz et al., 2023) or in combination with an LMS placeholder assignment (Howitz et al., 2021; McKnelly et al., 2023).

3.3. Themes and Opportunities

Many publications on the design and implementation of specifications grading discuss outcomes, impacts, and lessons learned. A review of the current literature reveals there are common themes that emerge from these discussions. These themes include instructor time investment, comparisons of academic performance, and student reactions to the grading scheme.

3.3.1. Instructor Commentary on Time Investment

One concern about adopting specifications grading is an instructor time commitment. While some publications do indicate an increase in the time spent grading (Hofmeister et al., 2022b; Joseph et al., 2023; Closser et al., 2024; Ring, 2017; Henriksen et al., 2020; Tamés, 2021; Hunter et al., 2022; Martin, 2019; Rojas & Quan, 2023; Cerkez, 2024; Kiefer & Earle, 2023; Hollinsed, 2018), the majority of adopters of specifications grading reported that they spent about the same amount of time (Elkins, 2016; Earl, 2021; Moster & Zingales, 2024; Suresh, 2023; Carlisle, 2020; McKnelly et al., 2021; Blackstone & Oldmixon, 2019; Lillard & Taggart, 2022; Vitale & Concepción, 2021; McKnelly et al., 2023; Lovell, 2018; Largent, 2024), or less time (Howitz et al., 2021; Jones, 2020; Walden, 2022; Dupree et al., 2024; Toledo & Dubas, 2017; Mirsky, 2018; Williams, 2018; Kelz et al., 2023; Fierke, 2024; Copp, 2024; Tuson &

Hickey, 2022; Sanft et al., 2021; Mendez, 2018b, 2023; Pascal et al., 2020; Trachsler et al., 2023; Noell et al., 2023; Tuson & Hickey, 2023) grading under the new system than they spent using a traditional points-based grading system (Table 3). The most common reason cited for saving time was the removal of partial credit (Earl, 2021; Fernandez et al., 2020; Walden, 2022; Dabney & VanDerWoude, 2023; Toledo & Dubas, 2017; Williams, 2018; Rojas & Quan, 2023; Copp, 2024; Largent, 2024; Tuson & Hickey, 2022; Tuson & Hickey, 2023). Without partial credit, the cognitive load associated with choosing the appropriate allocation of points is reduced. Instructors who did not experience time savings noted that the time they spent on grading was allocated differently; they could spend more time providing feedback because they spent less time deciding how much partial credit to award. Aside from removing partial credit, reducing the number of LOs being assessed (Mendez, 2018a; Toledo & Dubas, 2017) and removing assignments from the course that did not map to the LOs (Jones, 2020) were cited as contributing to the time saved when grading. Although not explicitly stated, it can be inferred that time may have also been saved from students choosing to not submit work for assignments that were not required for their target letter grade (Jones, 2020; Blackstone & Oldmixon, 2019).

Table 3. Number of publications that reported the time commitment required to design and implement a specifications grading system in a course.

Reported Time Commitments for Designing and Implementing Specifications Grading	Number of References
Increased time grading	12
No change in time grading	12
Decreased time grading	18
Removing partial credit saved time when grading regardless of changes in overall time spent grading	11
Significant time investment to design specifications grading system	20
Significant time investment to generate multiple versions of quizzes or exams	11

Although most publications on specifications grading indicate no change in, or a reduction in, grading time, several comments on the time investment required to design and build the system. The consensus is that the time investment associated with designing the specifications grading system is substantial (Elkins, 2016; Earl, 2021; Blodgett, 2017; Jones, 2020; Joseph et al., 2023; Dupree et al., 2024; Moster & Zingales, 2024; Shields et al., 2019; Henriksen et al., 2020; Blackstone & Oldmixon, 2019; Vitale & Concepción, 2021; McKnelly et al., 2023; Rojas & Quan, 2023; Mio, 2024; Copp, 2024; Trachsler et al., 2023; Noell et al., 2023; Mendez, 2023; LeHew, 2019; Anzovino et al., 2023). Time-consuming aspects of designing a specifications grading system that were mentioned include developing quiz questions and building question banks (Suresh, 2023; Henriksen et al., 2020; Rojas & Quan, 2023; Copp, 2024; Kiefer & Earle, 2023; Lovell, 2018; Tuson & Hickey, 2022; Noell et al., 2023; Mendez, 2023; Anzovino et al., 2023; Mattfeld, 2023), constructing new rubrics and assignment guidelines (Elkins, 2016; Earl, 2021; Dupree et al., 2024), and crafting grade bundles in which assignments are appropriately mapped to final letter grades (Elkins, 2016; Dupree et al., 2024; Rojas & Quan, 2023; Anzovino et al., 2023).

Additionally, some publications indicated that extra time was required for the specifications grading implementation that did not involve grading student work. One report of additional time focused on the need to meet with students outside of class because time for additional assignment attempts was not built into the course schedule (Lovell, 2018). It was also reported that additional time was required to address student concerns around

the removal of partial credit, to normalize the experience of needing to try again, and to achieve student buy-in for an unfamiliar grading system (LeHew, 2019). While the majority of publications do not describe the time commitment required to implement a token system, only two described the time commitment as onerous (Hofmeister et al., 2022b; Martin, 2019) and five describe it as minimal (Howitz et al., 2021; Gestwicki, 2021; McKnelly et al., 2021; Kelz et al., 2023; Copp, 2024). Management of the token system through the course LMS was cited as a way to keep the workload manageable (Copp, 2024).

3.3.2. Impacts on Academic Performance

One of the ways improvement in student learning is evaluated following the implementation of novel pedagogy is to examine whether the distribution of final letter grades in the course changed or not. For the majority of those who have published their implementations of specifications grading, the distribution of final letter grades either shifts in the direction of a larger percentage of collective A and B grades (Howitz et al., 2021; Hofmeister et al., 2022b; Jones, 2020; Moster & Zingales, 2024; Houseknecht & Bates, 2020; Evensen, 2022; Carlisle, 2020; Toledo & Dubas, 2017; Lillard & Taggart, 2022; Vitale & Concepción, 2021; Helmke, 2019; McKnelly et al., 2023; Kelz et al., 2023; Rojas & Quan, 2023; Mio, 2024; Kiefer & Earle, 2023; Hollinsed, 2018; Lovell, 2018; Noell et al., 2023; Anzovino et al., 2023; Katzman et al., 2021; Bunnell et al., 2023) or remains the same (Earl, 2021; Dennen & Bagdy, 2020; Closser et al., 2024; Blackstone & Oldmixon, 2019; Copp, 2024; Largent, 2024; Ahlberg, 2021) (Table 4). The reported impacts on course drop, fail, and withdrawal (DFW) rates are mixed. Some adopters of specifications grading report a decrease in their overall DFW rate (Moster & Zingales, 2024; Evensen, 2022; Toledo & Dubas, 2017; McKnelly et al., 2023; Kiefer & Earle, 2023; Anzovino et al., 2023), while others report no change (Earl, 2021; Evensen, 2022; Hollinsed, 2018; Bunnell et al., 2023) or an increase (Gargac, 2022a; Lillard & Taggart, 2022; Noell et al., 2023; Anzovino et al., 2023). Noell et al. observed an increase in the DFW rate, noting that the Ds and Fs decreased, but the Ws increased (Noell et al., 2023). The increase in withdrawals was attributed to the transparency of the specifications grading system, allowing students to know if they would be able to pass the course before the withdrawal deadline.

Table 4. Number of publications that reference how implementing a specifications grading system impacted course letter grades and DFW rates.

Specifications Grading Impact on Course Letter Grades and DFW Rates	Number of References
Larger percentage of collective A and B grades	22
No change in course letter grade distribution	7
Course GPA decreased	4
DFW rate increased	4
DFW rate decreased	6
No change in DFW rate	4

Only four publications did not report a positive shift in the final letter grade distribution (Gargac, 2022a, 2022b; Jones, 2020; Wasniewski et al., 2021). One of these publications specifically reported that more students chose to complete the assignments associated with a B grade than with an A grade, which likely accounted for the decrease in the course GPA (Jones, 2020). No commentary was provided in the other three publications to explain what may account for the decrease in the final letter grade distribution (Gargac, 2022a, 2022b; Wasniewski et al., 2021).

It is reasonable to expect that the positive shift in final letter grade distributions observed by many adopters of specifications grading could be the result of an improve-

ment in the quality of student work. However, the majority of publications only address comparisons of student work anecdotally between versions of the course taught using a points-based grading system and the version taught using a specifications grading system. In these cases, the perception reported was that the quality of student work submitted under the specifications grading system was similar (Quintana & Quintana, 2020; Walden, 2022; Donato & Marsh, 2023; Blackstone & Oldmixon, 2019; Noell et al., 2023) or better (Earl, 2021; Fernandez et al., 2020; Dennen & Bagdy, 2020; Jones, 2020; Joshi, 2023; Wasniewski et al., 2021; McKnelly et al., 2021; Cosoroaba, 2020; Kelz et al., 2023; Fierke, 2024; Sanft et al., 2021; Mendez, 2018b) than work submitted in points-based systems.

Five publications included a quantitative comparison of the quality of work between the two grading systems. Two of these publications compared final exams between the points-based and specifications grading versions of their courses. Ring compared student work by grading final exams from their previous points-based version of their course with the same rubric as the specifications grading course (i.e., without partial credit) and found that the students from the specifications grading version of the course passed questions at a higher rate (Ring, 2017). However, Ring did not comment on whether there was a change in the final letter grade distribution between the two courses. In Martin's courses, the same final exam composed of 60 multiple-choice questions was used for both a points-based and a specifications grading version of their general chemistry course (Martin, 2019). The average score in four previous course iterations that used points-based grading ranged between 30 and 38 (50–63%), whereas in the three terms following the adoption of a specifications grading system, the averages ranged between 38 and 42 (63–70%). Changes to the final letter grade distribution by adopting specifications grading were not discussed.

In addition to the two studies analyzing student performance on final exams, three other publications reported quantitative comparisons of the quality of work unrelated to final exams in different grading versions of the courses. Helmke set the passing threshold for assignments and exams in their specifications grading course to a B (85%) and compared the number of students in the specifications grading course that met the threshold to the number of students that earned at least an 85% on assignments and exams in the points-based course (Helmke, 2019). Helmke found that fewer students from the specifications grading course passed on the first attempts of both the homework assignments and unit exams; however, more students passed on the final attempts for homework assignments, and a similar number passed on the final attempts for the unit exams. Under Helmke's specifications grading system, the percentage of students earning As remained the same, but the percentage of Bs increased and the percentage of Cs decreased. Katzman took a different approach by administering an end-of-semester survey to students who completed a version of their course under a points-based grading system and to students who completed a version of the course under a specifications grading system (Katzman et al., 2021). Students who completed the course using a specifications grading system earned higher median and maximum scores on the content assignment questions in the survey than those who completed the course using a points-based grading system. Katzman observed no difference in the percentage of students who did pass and who did not pass the course between the two grading systems. However, amongst the passing students, the percentage of As increased, and the percentage of Bs and Cs decreased. Amongst the non-passing students, the percentage of Ds and Ws increased, but the percentage of Fs decreased. Finally, McKnelly et al. compared student performance on a laboratory report (McKnelly et al., 2023). They graded the same number of reports for a single assignment from the points-based grading and specifications grading versions of the course using the original points-based rubric and found no statistically significant difference in the average scores on the reports. Despite the similar scores, McKnelly et al. found an increase in the percentage of students earning

A and B grades, but a decrease in students earning C and D grades. One explanation they offer for the discrepancy is the limited sample size and comparison of quality on a single assignment. They propose that assignment quality may have increased for some assignments and not others.

One publication did not directly compare the quality of work between the two grading systems but did compare the impact of taking a course graded using a points-based system versus a specifications grading system on the passing rate in that course and the one following it. In both a general and an organic chemistry course sequence, Anzovino et al. found no statistically significant difference amongst students continuing to the second course in the sequence regardless of whether they took the first course in each sequence under a points-based system or a specifications grading system (Anzovino et al., 2023). Additionally, they found no statistically significant differences in the passing rates (C or better) in the second course in each sequence, regardless of whether students took the first course in each sequence graded using a points-based system or a specifications grading system. This may suggest that the quality of work submitted by students under a specifications grading system is not any lower than that submitted by students under a points-based grading system.

Finally, one publication directly addresses potential grade inflation in specifications grading—a common concern. In a specification grading graduate-level organic chemistry course using a modules configuration (six essential and seven general), Moster and Zingales observed higher final letter grades than in the prior points-based iteration (Moster & Zingales, 2024). Students earned more As and fewer Bs and Cs, with 10% more students passing the course overall. Students who earned As earned higher average scores on a 50-question final exam than students who earned Bs. Moster and Zingales connect this higher exam performance to the fact that students who earned A grades were required to complete an additional essential module and four general modules as compared to students who earned B grades. Because students had to complete 54% of the course content (seven weekly modules; five essential and two general) before taking the final exam, and the passing threshold on each module was set to 80%, it follows that students passing the course should be scoring at least 43% ($0.54 \times 0.80 = 0.43$) on the final exam. Because 90% of the students met this criterion, with many earning scores far greater than 43%, Moster and Zingales argue that the final exam outcomes indicate that students have met the course SLOs—and retained the knowledge—at levels that are clearly commensurate with the letter grades they earned.

3.3.3. Student Reactions to Specifications Grading

Within the theme of student reactions to specifications grading, three primary sub-themes emerged as follows: focus on learning, the transparency of the grading system, and stress and anxiety. These subthemes are derived from 1) anecdotal data from instructor observations, conversations with students, and quotes pulled from final evaluations and 2) solicited feedback from students through surveys that were not validated instruments. These subthemes are tied to goals of specifications grading according to Nilson: to shift student focus from points to learning, to provide a transparent grading scheme, and to reduce student stress and anxiety.

Focus on Learning. Many educators who have implemented specifications grading observed a shift from students focusing on accumulating points to focusing on learning the course material (Howitz et al., 2021; Joseph et al., 2023; Santucci & Golas, 2023; Wasniewski et al., 2021; Suresh, 2023; Johanesen et al., 2024; Mirsky, 2018; Lillard & Taggart, 2022; Johnson, 2023; Tamés, 2021; Hunter et al., 2022; Ludwigsen, 2017; Helmke, 2019; McKnelly et al., 2023; Kelz et al., 2023; Cerkez, 2024; Mio, 2024; Copp, 2024; Mattfeld, 2023). For

example, Hunter et al. state, “Students clearly grasp why their assignment earned the grade that it did, and I no longer field questions about point allocation. Instead, student questions after an exam focus on concepts and improvement, and the process of revision absolutely improves the students’ understanding of the material (Hunter et al., 2022)”. Similarly, Henriksen et al. and Lovell observed that students took greater advantage of office hours to ask questions about concepts covered earlier in the course because the incentive for students to seek and apply feedback was greater under the specifications grading system where revision is encouraged (Henriksen et al., 2020; Lovell, 2018). Part of the reduction in students’ focus on points, aside from revision opportunities, was attributed to the greater transparency and clarity of the specifications grading system, specifically the expectations detailed in assignment rubrics (Fernandez et al., 2020; Vitale & Concepción, 2021; Pascal et al., 2020).

Despite increased student focus on learning, some students opposed the removal of partial credit. In four publications, authors noted that students in their specifications grading courses felt the passing thresholds were set too high (Gratwick et al., 2020; Kinnear et al., 2022; McKnelly et al., 2021; McKnelly et al., 2023). Kinnear et al., McKnelly et al., and Williams observed student frustration when they just missed meeting the passing threshold (Kinnear et al., 2022; Williams, 2018; McKnelly et al., 2023). In these cases, students felt the effort they put into the work they submitted was not being taken into account, as no partial credit was given for assignment submissions that did not pass. Reports from Toledo and Dubas, from Rojas, and from Blodgett also acknowledge student displeasure with not earning partial credit on work that did not meet the passing threshold (Blodgett, 2017; Toledo & Dubas, 2017; Rojas & Quan, 2023).

Grading Transparency. Many adopters of specifications grading report that students find the grading system to be more transparent than points-based courses. For some students, this transparency refers to the clarity of assignment expectations and the associated rubrics (Henriksen et al., 2020; Cosoroaba, 2020; Vitale & Concepción, 2021; Lovell, 2018). For others, transparency refers to knowing what assignments need to be completed to earn their desired final letter grade and what their standing is in the course at any point throughout the term (Howitz et al., 2021; Dennen & Bagdy, 2020; Hofmeister et al., 2022b; Kelz et al., 2023; Copp, 2024; Pascal et al., 2020). In two publications, authors commented that their students appreciated that the transparency of the specifications grading system allowed them to choose what assignments they needed to complete to achieve the grade they aimed for (LeHew, 2019; Mirth, 2017). Jones indicated that their students found that the choices afforded by the specifications grading system led to a greater ability to balance coursework in other classes and their life obligations (Jones, 2020). In three cases, authors noted that their students felt that the transparency allowed them to direct their efforts in the course because the grading system provided guidance about which topics they understood well and which they needed to continue working on (Toledo & Dubas, 2017; Hunter et al., 2022; Rojas & Quan, 2023).

Although many instructors reported that students found the grading system more transparent, there were several who reported student confusion about the new grading system (Earl, 2021; Martin et al., 2021; Hofmeister et al., 2022b; Joseph et al., 2023; Moster & Zingales, 2024; Santucci & Golas, 2023; Evensen, 2022; Closser et al., 2024; Biers, 2022; Lillard & Taggart, 2022; Hunter et al., 2022; Mio, 2024). Some students, at least initially, felt the new assignment expectations and rubrics were unclear (Howitz et al., 2021; Kelz et al., 2023; Cerkez, 2024). Others were unaware of how to determine their final letter grade according to the grade bundles (Howitz et al., 2021; Joseph et al., 2023; Toledo & Dubas, 2017). Some instructors adopted strategies in an attempt to achieve student buy-in and minimize confusion with this novel grading system. Early in the term, some instructors

included in class or video explanations of the grading system (Elkins, 2016; Howitz et al., 2021; Hofmeister et al., 2022b; Jones, 2020; Quintana & Quintana, 2020; Dupree et al., 2024; Dabney & VanDerWoude, 2023; Tsoi et al., 2019; Carlisle, 2020; Blackstone & Oldmixon, 2019; Hunter et al., 2022; Prasad, 2020; McKnelly et al., 2023; Cerkez, 2024; Fierke, 2024; Mio, 2024; Pascal et al., 2020; LeHew, 2019), some compared specifications grading to traditional grading to help students understand the purpose behind why it was adopted (Martin et al., 2021; Prasad, 2020), and some developed activities or tools for students to learn how to track and determine their final letter grade (Howitz et al., 2021; Shields et al., 2019; Carlisle, 2020; Toledo & Dubas, 2017; Lillard & Taggart, 2022; McKnelly et al., 2023; Pascal et al., 2020; Reck, 2022). Other instructors provided regular reminders throughout the term (Jones, 2020; Dupree et al., 2024; Santucci & Golas, 2023; Evensen, 2022; Dabney & VanDerWoude, 2023; Carlisle, 2020; Johnson, 2023; McKnelly et al., 2023; Fierke, 2024; Reck, 2022) and/or implemented metacognitive reflections or goal-setting exercises (Howitz et al., 2021; Biers, 2022; Johnson, 2023; Prasad, 2020; McKnelly et al., 2023; Yang & Korsnack, 2024; Copp, 2024), which prompted students to think through how they could earn their desired final letter grade under the specifications grading system. The extent to which these interventions mitigated or resolved student confusion varied.

Stress and Anxiety. Student perceptions of the impact of specifications grading on their stress and anxiety compared to points-based grading vary. Twenty-four articles report a reduction in student stress and anxiety and comment on possible reasons for this change. In some cases this reduction is attributed to multiple revision opportunities being provided, lowering the stakes on assignments (Earl, 2021; Fernandez et al., 2020; Walden, 2022; Joseph et al., 2023; Moster & Zingales, 2024; Evensen, 2022; Closser et al., 2024; Carlisle, 2020; Biers, 2022; Henriksen et al., 2020; Hunter et al., 2022; McKnelly et al., 2023; Cerkez, 2024; Fierke, 2024; Tuson & Hickey, 2022; Mattfeld, 2023; Ahlberg, 2021; Pope et al., 2020). In other instances, students reported feeling that the grading was lower stakes, which allowed them to put more of their focus into learning the course material. However, they did not elaborate upon what made the grading feel lower stakes (Jones, 2020; Helmke, 2019). Trachsler et al. reported that students felt reduced stress and anxiety stemmed from the flexibility afforded to them by being able to choose which assignments to complete to earn their desired grade (Trachsler et al., 2023). Additionally, students perceived greater transparency in assignment expectations and how to earn their desired final letter grade, which accounted for a reduction in stress and anxiety (Howitz et al., 2021; Earl, 2021; Quintana & Quintana, 2020; Biers, 2022; Cerkez, 2024; LeHew, 2019; Pope et al., 2020).

In contrast to a decrease in stress and anxiety, other students reported an increase with the specifications grading system. Many of the concerns raised by students were related to the pass/fail nature of specifications grading, and some students equated not passing an assignment on the first attempt to failing despite the fact that they were able to try again without a grade penalty. Some students felt that the expectations to pass an assignment were too high (Joshi, 2023; Shields et al., 2019; Kinnear et al., 2022; Carlisle, 2020; Helmke, 2019; Kiefer & Earle, 2023; Lovell, 2018; Pope et al., 2020). Many students did not like the absence of partial credit and felt that they should receive some credit for the work they submitted rather than not earning any credit for an assignment that did not meet the passing threshold (Howitz et al., 2021; Shields et al., 2019; Prasad, 2020; McKnelly et al., 2023; Copp, 2024; Noell et al., 2023). Students expressed stress and anxiety about how one mistake could make the difference between passing and not passing an assignment and consequently affect their grades in the course (McKnelly et al., 2021; Cosoroaba, 2020; McKnelly et al., 2023; LeHew, 2019). Elkins acknowledges that not earning credit for the work they submitted can be a “harsh reality for students who are used to earning at least partial credit no matter how low their level of work (Elkins, 2016)”. Other instructors

identified similar perceptions in which students appeared to believe that needing to revise an assignment was equivalent to “receiving an F (Johanesen et al., 2024),” “a demoralizing ‘0’ rather than an opportunity for improvement” (Rojas & Quan, 2023), or would “result in the assignment somehow being worth ‘less’ (LeHew, 2019).” Blodgett shared a student quote that highlights this demoralization: “Making a great deal of effort and then getting a 0 for an assignment made me wonder why bother at all (Blodgett, 2017)”. Noell et al. and Hunter et al. found that students are not accustomed to having opportunities to revise work, so it is useful in these situations to give reminders that not being assessed as passing on the first attempt is not a sign of failure and that revision opportunities are built into the course (Hunter et al., 2022; Noell et al., 2023).

Other sources of stress and anxiety unrelated to the pass/fail nature of specifications grading included the frequency of testing (Ring, 2017; Noell et al., 2023), the responsibility placed on the students to self-track their grades (Toledo & Dubas, 2017), and the tendency for increased procrastination by some students because they knew they had retake opportunities (Henriksen et al., 2020; Tuson & Hickey, 2022). Students who expressed having good grades going into a final assignment did not like that poor performance on a final assignment could negatively impact their final letter grade (Carlisle, 2020; Prasad, 2020; McKnelly et al., 2023).

4. Discussion

As the number of publications describing the design and implementation of specifications grading has grown, so has the breadth of disciplines from which instructors have reported their experiences. Despite specifications grading systems having four common features—grade bundles, rubrics with specifications and defined passing thresholds, opportunities to revise and resubmit work, and a token system—the details of each design and implementation vary substantially from instructor to instructor. Design choices likely impacted the outcomes of each implementation. Following, we discuss the themes that arose from instructors’ design choices and how they influenced implementation outcomes in specifications grading systems, focusing particularly on impacts on time, academic performance, and student reactions to the grading system.

4.1. Instructor Commentary on Time Investment

In general, switching to specifications grading required a substantial time commitment prior to the start of a course to design the system. One of the time-consuming aspects of the design was the construction of the grade bundles. As instructors worked to construct grade bundles, they were incentivized to reevaluate their course LOs and the assignments that mapped to those LOs. This backward course design in tandem with a specifications grading approach where the focus is put on the student achievement of LOs rather than the accumulation of points to earn a letter grade suggests that a grade a student earns in a course under specifications grading may more accurately reflect the knowledge and skills they have gained by completing the course (Earl, 2021; Blackstone & Oldmixon, 2019; McKnelly et al., 2023; Mio, 2024).

Although the time required to design a specifications grading system was substantial, most instructors who implemented specifications grading in their courses reported that the time required to implement the course was about the same or less than the amount of time spent implementing prior points-based courses. Instructors cite that the time they would have spent on allocating points on assignments in the points-based course was instead spent on giving students more feedback on assignments (Earl, 2021; Moster & Zingales, 2024; Suresh, 2023; Largent, 2024; Mendez, 2018b). Previous research has shown that when both scores and feedback are provided to students, students will not necessarily pay

attention to the feedback, in part due to the spatial separation of scores and feedback within the LMS (Winstone et al., 2021). However, when students have opportunities to revise their work, they show higher cognitive engagement with the feedback (Fredricks et al., 2016; Espasa et al., 2022). The opportunity for students to revise or retake assignments under a specifications grading system may incentivize students to more carefully review the feedback that was given rather than only checking their scores. In this way, the additional feedback instructors leave on student assignments in courses using specifications grading may be more likely to lead to students producing higher quality work and achieving higher final letter grades compared to points-based courses.

Although the majority of instructors found the time to implement specifications grading was not greater than for a points-based course, a small number reported spending more time implementing the new system. In analyzing how these instructors structured their grading systems, it became apparent that students were given many opportunities and/or unlimited time to retake or resubmit assignments (Henriksen et al., 2020; Martin, 2019; Prasad, 2020; Hollinsed, 2018). Allowing many reattempts can result in students submitting low-quality work on early attempts because the incentive to prepare appropriately for assignments is not present (Prasad, 2020; Hollinsed, 2018; Mendez, 2018b). If deadlines for revising and resubmitting or retaking assignments are not provided, it may increase the likelihood of procrastination (Closser et al., 2024; Prasad, 2020). This can result in a high instructor workload at the end of the term when many assignments are submitted, and it may also result in students being unable to earn the grade they are aiming for because there is insufficient time to complete all of the unfinished assignment revisions (Closser et al., 2024; Prasad, 2020). It appears that token systems were not used in many of these cases, which could have mitigated some of the challenges with implementing opportunities for students to try again. Instructors who proactively included limited revision opportunities, deadlines for revising and resubmitting or retaking assignments, and/or a token system in their design and implementation of specifications grading reported manageable time investments (Howitz et al., 2021; Dabney & VanDerWoude, 2023; McKnelly et al., 2023; Kelz et al., 2023; Copp, 2024; Mendez, 2023).

Several publications included discussions of how the specifications grading system evolved over time, specifically focusing on changes that were made to streamline the course and reduce the instructor's time commitment. Changes that saved faculty time included a reduction in the number of LOs assessed and/or the number of assignments necessary to earn a letter grade (Earl, 2021; Dupree et al., 2024; Mendez, 2018a; Toledo & Dubas, 2017), consolidation of course content coverage (Closser et al., 2024), limiting the number of retakes per assignment (Closser et al., 2024; Prasad, 2020; Hollinsed, 2018), and/or imposing deadlines on the time available to students to revise or retake an assignment (Closser et al., 2024; Fierke, 2024). This suggests that while the adoption of specifications grading may require a significant time investment initially, the time savings become more apparent during implementation, especially if appropriate revisions are made to the system over multiple iterations of the course. While not necessarily implemented, other instructors suggest that changing assignments to be auto-graded (Hofmeister et al., 2022b), developing methods to auto-generate test questions (Mattfeld, 2023), and building on the work of others, such as through a faculty learning community (Anzovino et al., 2023), could also save time in future implementations of their specifications grading course.

4.2. Impacts on Academic Performance

The majority of publications report a shift in the final letter grade distribution toward more A and B grades after adopting specifications grading. In general, it would be expected that as the final letter grade distribution shifts toward more A and B grades, the quality of

student work would show substantial improvement. However, this potential correlation has not been studied systematically. Most reported comparisons of the quality of student work between points-based and specifications grading systems are anecdotal. This lack of detailed comparisons of student work is unsurprising due to the drastic changes that are made to a course when converting from a points-based grading system to a specifications grading system. Assignments are typically modified to better align with the LOs. The removal of partial credit makes conducting a rigorous comparison of student work particularly challenging because the scales are not the same between the two grading systems (e.g., an assignment task may be graded out of five in a points-based grading system versus being graded using the binary pass/no pass rating in a specifications grading system).

Without rigorous studies indicating that the quality of student work is higher under specifications grading, an argument could be made that the positive shift in final letter grades is a form of grade inflation due to reduced rigor. However, this argument fails to take into account that students are held to a high standard to earn credit on assignments when no partial credit is included. For example, an instructor may set a threshold of B-level work or above for a student to earn credit on an assignment, as recommended by Nilson. Holding students to these high standards often results in students not passing all assignments on the first attempt. Without the ability to accumulate partial credit, students are held accountable for revisiting the material and taking advantage of revision or resubmission opportunities to demonstrate an improved understanding of LOs. Providing opportunities for students to try again to meet the high standard set by the instructor supports learning and results in grade elevation, not grade inflation (Streifer & Palmer, 2021). This sentiment is supported by work conducted by Moster and Zingales in their organic chemistry course described earlier in this review (Moster & Zingales, 2024). They suggest the higher final letter grades students earn under the specifications grading system are consistent with the achievement, of course, LOs based on their performance on a final exam. While their work presents some initial evidence that students' quality of work correlates with final letter grades, additional studies are necessary to quell the concerns of those who are hesitant to adopt alternative grading systems for fear of propagating grade inflation.

4.3. Student Reactions to Specifications Grading

Overall, student sentiment toward specifications grading systems tends to be positive and appears to improve as a course progresses. Negative reactions to specifications grading may be attributed to student unfamiliarity with the grading scheme—because many students are habituated to traditional points-based grading—and/or to specific design choices within courses. Reports of student confusion about the specifications grading system and final letter grade determination at the start of a course were common. The confusion was likely due to students encountering an alternative grading approach for the first time and having to adjust to changes in grading norms. Instructors observed improvements in students' understanding of the specifications grading system when the number of LOs was reduced, course content coverage was consolidated, grade bundles were streamlined, and/or when a tool was provided for students to track progress toward and determine their final letter grade. Many publications reported attempts to increase students' buy-in to the specifications grading system. However, these interventions seemed to have varied success in increasing students' comfort with the system. Future studies could investigate the efficacy of buy-in interventions, evaluate how student buy-in changes during a term, and explore whether student buy-in improves after multiple offerings of the course under the specifications grading system.

In addition to reports of student confusion, cases of increased stress and anxiety were also reported. This increase in stress and anxiety, much like the confusion experienced

by students taking a course taught using a specifications grading system for the first time, may stem primarily from students' habituation to traditional points-based grading systems. For example, students did not like that they could not earn any credit when they submitted work that was not of passing quality. In these cases, students wanted credit for effort expended and/or felt that the passing threshold was set too high. This desire for partial credit and/or credit awarded for effort regardless of performance may reflect a performance-focused goal orientation adopted by these students. The fact that these complaints come from only a small percentage of students in a given course with a specifications grading system may indicate that the majority of the students adopted a mastery-focused goal orientation (Elliot & McGregor, 2001; Elsinger & Lewis, 2020).

Design choices made by instructors when creating their specifications grading systems may also contribute to some of the increased student stress and anxiety about grades. For example, setting the passing threshold of an assignment to 100% would send the message to students that perfection is required and could easily lead to students feeling that they will never pass an assignment or the course. Additionally, providing unlimited attempts or unlimited time with no clear due dates may remove the structure that students need to stay on track and avoid procrastination. These observations suggest that setting an appropriate passing threshold for an assignment and providing structure around assignment attempts are critical design choices. Sorensen-Unruh argued that while self-regulated learning theory (SRL) is often cited as the underlying theory of alternative grading systems (captured under the umbrella term "ungrading"), SRL is inherently deficit-framed (Sorensen-Unruh, 2024). Instructors may unintentionally be adopting a deficit framing when designing their specifications grading systems, and this underlying assumption may be leading to design choices that undermine the goals of alternative grading.

While negative sentiments arose from students experiencing specifications grading systems, often for the first time, most students reacted positively to the structural components of these grading systems that are not often, if ever, present in traditional points-based grading systems. For example, having opportunities to revise and resubmit reduced student stress and anxiety by reducing the stakes of any single assignment. Additionally, the combination of not being expected to complete or pass every assignment and the pass/fail approach of specifications grading systems encouraged students to focus on the learning process rather than the accumulation of points to earn their final letter grade. From a theoretical perspective, the shift away from a focus on points may indicate a change in students' achievement goal orientation toward a mastery focus (Elliot & McGregor, 2001; Elsinger & Lewis, 2020). From a practical perspective, this shift away from a focus on points is logical because allocating partial credit on assignments creates many borderline cases, which in turn incentivizes students to fight for every point. Without partial credit, the number of borderline cases drops drastically as each specification is frequently graded on a binary scale (Joshi, 2023). The structural components discussed above, combined with students having the ability to choose the type of assignment to complete, provided students with greater agency and, in most cases, a reduction in stress and anxiety. This increased agency also allowed students to better balance their obligations outside of academics and to better allocate time across all of their academic courses.

The features of specifications grading systems discussed above may not only be improving student academic performance but may also be developing students' professional and social identities in their chosen disciplines. Observed gains in students' math identity (Villalobos et al., 2024) suggest that specifications grading systems can provide an avenue to improve inclusion and equity in STEM courses, and similar studies could be expanded to non-STEM disciplines. The implications of specifications grading systems to help address

equity concerns in higher education provide ripe opportunities for further research to investigate these avenues.

4.4. Limitations

In addition to those discussed above, there are other limitations present in our analysis. Less formal methods of dissemination, such as theses, posters, workshops, etc., were excluded, so there are some design and/or implementation considerations about how instructors are implementing specifications grading systems in their courses that we did not discuss. Additionally, there are instructors who are contingent faculty or who are not in positions where publication is necessary for their academic institution, so there are unpublished examples of specifications grading systems that we cannot analyze and include in this review. This limitation may be especially present in implementations of specifications grading in community colleges and other two-year institutions. As is discussed above, most of the instructors publishing in the specifications grading system literature have written descriptive papers, which do not necessarily include controlled studies, so conclusions about the effects of specifications grading systems on students are limited. Finally, in a finding that echoes work by Hackerson et al., we note that connections to theoretical frameworks were rarely discussed in the specifications grading literature, which limited our ability to address how theory is being incorporated into the design and implementation of specifications grading systems (Hackerson et al., 2024).

5. Conclusions

Alternative grading systems have emerged to address challenges associated with points-based grading systems. These challenges include but are not limited to, a misalignment between the LOs students achieve and their final course grade, a student's focus on achieving a grade due to the influence of external rather than internal motivators, and a rise in student stress, anxiety, and mental health challenges. Specifications grading systems have been gaining popularity as one of the alternative systems that have emerged in the hopes of mitigating some of the challenges associated with points-based grading systems. The findings in this review indicate that publications on the design and implementation of specifications grading continue to grow, with the majority of publications occurring in STEM fields (chemistry, engineering, computer science, and mathematics) and in small courses (≤ 35 students), although examples across all course sizes and a wide variety of disciplines do exist (research question 1). Grade bundles, rubrics with specifications and defined passing thresholds, opportunities to revise and resubmit work, and a token system are the four structural components of specifications grading systems and are found to be incorporated in a variety of ways (research question 2).

Analysis of the implementations of the specifications grading systems and emergent themes of instructor commentary on time investment, academic performance, and student reactions to specifications grading reveal (1) important considerations for the design and development of specifications grading systems and (2) a roadmap for education researchers to collaborate with practitioners to study outcomes and impacts of specifications grading in higher education (research question 3). Many of the current publications on specifications grading are descriptive studies on individual designs and implementations. While additional descriptive publications will be valuable resources for practitioners seeking inspiration for their courses, especially from disciplines or types of courses that are not yet represented, enough evidence now exists to point to the fact that we need future research to include systematic studies on topics such as impacts on student academic performance, implications of design choices, effective practices for securing student buy-in, and potential for increased equity of specifications grading systems. Hackerson et al. have called for

interdisciplinary studies on alternative grading systems across STEM fields (Hackerson et al., 2024), and we broaden that call to include the need for interdisciplinary studies on specifications grading systems across both STEM and non-STEM disciplines.

Supplementary Materials: The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/educsci15010083/s1>, Table S1: Full Corpus of Papers Analyzed.

Author Contributions: Methodology, conceptualization, research, writing, review, and editing. W.J.H., K.J.M. and R.D.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Data Availability Statement: The full corpus of publications analyzed for this review is available in the supplementary files.

Acknowledgments: The authors thank Clarissa Sorensen-Unruh and Katy Hosbein for helpful conversations on review structures.

Conflicts of Interest: The authors declare no conflicts of interest.

References

- Ahlberg, L. (2021). Organic chemistry core competencies: Helping students engage using specifications. In *Engaging students in organic chemistry* (Vol. 1378, pp. 25–36). American Chemical Society.
- Anzovino, M. E., Behmke, D., Villanueva, O., & Woodbridge, C. M. (2023). Specifications grading and COVID. In *ACS Symposium Series* (pp. 89–105). American Chemical Society.
- Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Prentice Hall.
- Beare, P. G. (1986, October 9–11). *The contract—an individualized approach to competency-based learning and evaluation*. Thinking across the Disciplines. Annual Conference of the International Society for Individualized Instruction, Atlanta, GA, USA.
- Biers, K. (2022). Decolonial and feminist course design and assessment in the first-year french curriculum. In S. Bouamer, & L. Bourdeau (Eds.), *Diversity and decolonization in French studies: New approaches to teaching* (pp. 245–257). Springer International Publishing.
- Blackstone, B., & Oldmixon, E. (2019). Specifications grading in political science. *Journal of Political Science Education*, *15*, 191–205. [CrossRef]
- Blodgett, B. J. (2017). Grading matters in theological education. *Teaching Theology & Religion*, *20*, 314–326. [CrossRef]
- Bloom, B. S. (1968). Learning for Mastery. Instruction and curriculum. Regional education laboratory for the carolinas and virginia, topical papers and reprints, number 1. *Evaluation Comment*, *1*, n2.
- Bowen, R. S., & Cooper, M. M. (2022). Grading on a curve as a systemic issue of equity in chemistry education. *Journal of Chemical Education*, *99*, 185–194. [CrossRef]
- Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., Stevens, M. T., & Welsh, M. E. (2016). A century of grading research: Meaning and value in the most common educational measure. *Review of Educational Research*, *86*, 803–848. [CrossRef]
- Brown, R., & Kennedy, T. (2022, June 26–29). *Work in progress: Differentiated learning in a specifications grading framework*. 2022 ASEE Annual Conference & Exposition, ASEE Conferences, Minneapolis, MN, USA.
- Bunnell, B., LeBourgeois, L., Doble, J., Gute, B., & Wainman, J. W. (2023). Specifications-based grading facilitates student–instructor interactions in a flipped-format general chemistry II course. *Journal of Chemical Education*, *100*, 4318–4326. [CrossRef]
- Carberry, A. R., Siniawski, M. T., & Dionisio, J. D. N. (2012, October 3–6). *Standards-based grading: Preliminary studies to quantify changes in affective and cognitive student behaviors*. 2012 Frontiers in Education Conference Proceedings (pp. 1–5), Seattle, WA, USA.
- Carlisle, S. (2020). Simple specifications grading. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, *30*, 926–951. [CrossRef]
- Cerkez, E. B. (2024). Application of specifications grading to an analytical chemistry lab. *Journal of Chemical Education*, *101*, 4268–4275. [CrossRef]
- Clark, D., & Talbert, R. (2023). *Grading for growth: A guide to alternative grading practices that promote authentic learning and student engagement in higher education*. Taylor & Francis.
- Closser, K. D., Hawker, M. J., & Muchalski, H. (2024). Quantized grading: An ab initio approach to using specifications grading in physical chemistry. *Journal of Chemical Education*, *101*, 474–482. [CrossRef]
- Copp, D. A. (2024, June 23–26). *Specifications grading in an undergraduate engineering dynamics course*. 2024 ASEE Annual Conference & Exposition, Portland, OR, USA.

- Cosoroaba, E. (2020, June 22–26). *Helping students write it right: Instilling good report-writing habits in a linear circuit lab course*. 2020 ASEE Virtual Annual Conference Content Access Proceedings, ASEE Conferences, Online.
- Dabney, B. W., & VanDerWoude, C. J. (2023). Implementing a specifications grading system in a nursing course. *Nurse Educator*, *48*, 187–191. [CrossRef] [PubMed]
- Dennen, V., & Bagdy, L. (2020, April 14–16). *Shifting to specifications grading: Two design cases*. TCC 2020 Conference Papers, TCC Worldwide Online Conference, Online.
- Donato, J. J., & Marsh, T. C. (2023). Specifications grading is an effective approach to teaching biochemistry. *Journal of Microbiology & Biology Education*, *24*, e00236-22. [CrossRef]
- Dupree, L. H., Augustine, J. M., & Miller, S. W. (2024). How did we get here? Evolution of specifications grading in a required skills-based course series. *Currents in Pharmacy Teaching and Learning*, *16*, 102124. [CrossRef]
- Durm, M. W. (1993). An A Is Not an A Is Not an A: A History of Grading. *The Educational Forum*, *57*, 294–297. [CrossRef]
- Dweck, C. S. (2013). *Self-theories: Their role in motivation, personality, and development* (1st ed.). Psychology Press.
- Earl, D. (2021). Two years of specifications grading in philosophy. *Teaching Philosophy* *45*, 23–64. [CrossRef]
- Elkins, D. M. (2016). Grading to learn: An analysis of the importance and application of specifications grading in a communication course. *Kentucky Journal of Communication*, *35*, 2016.
- Elliot, A. J., & McGregor, H. A. (2001). A 2 × 2 Achievement goal framework. *Journal of Personality and Social Psychology*, *80*, 501–519. [CrossRef] [PubMed]
- Elsinger, J., & Lewis, D. (2020). Applying a standards-based grading framework across lower level mathematics courses. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, *30*, 885–907. [CrossRef]
- Espasa, A., Guasch, T., Mayordomo, R. M., & Martinez-Melo, M. (2022). Prior experience with online feedback: Its influence on students' engagement. *Distance Education*, *43*, 444–465. [CrossRef]
- Evensen, H. (2022, June 26–29). *Specifications grading in general physics and engineering physics courses*. 2022 ASEE Annual Conference & Exposition, ASEE Conferences, Minneapolis, MN, USA.
- Eyler, J. R. (2024). *Failing our future*. Johns Hopkins University Press.
- Fernandez, T., Martin, K., Mangum, R., & Bell-Huff, C. (2020, June 22–26). *Whose grade is it anyway?: Transitioning engineering courses to an evidence-based specifications grading system*. 2020 ASEE Virtual Annual Conference Content Access Proceedings, ASEE Conferences, Online.
- Fierke, K. L. (2024). Utilizing specification grading in sport management classes. *Sport Management Education Journal*, *18*, 151–156. [CrossRef]
- Fredricks, J. A., Wang, M. -T., Schall Linn, J., Hofkens, T. L., Sung, H., Parr, A., & Allerton, J. (2016). Using qualitative methods to develop a survey measure of math and science engagement. *Learning and Instruction*, *43*, 5–15. [CrossRef]
- Gargac, J. (2022a, June 6–8). *Incorporating industry-based metrics into a specifications grading system for capstone assessment*. Capstone Design Conference 2022, Capstone Design Community, Dallas, TX, USA.
- Gargac, J. (2022b, June 26–29). *Manufacturing Engagement: Improving Student Learning through Modifying Content Delivery and Assessment*. 2022 ASEE Annual Conference & Exposition, ASEE Conferences, Minneapolis, MN, USA.
- Gay, J. L., & Poproski, R. (2023). Specifications grading for graduate students: Motivated, less stressed, and achieving in public health. *Pedagogy in Health Promotion*, *9*, 237–243. [CrossRef]
- Gestwicki, P. (2021). Godot engine and checklist-based specifications: Revising a game programming class for asynchronous online teaching. *Journal of Computing Sciences in Colleges*, *37*, 30–40.
- Gratwick, R., Kinneer, G., & Wood, A. K. (2020). An online course promoting wider access to university mathematics. In R. Marks (Ed.), *Proceedings of the British society for research into learning mathematics* (1st ed., Vol. 40). Available online: https://www.pure.ed.ac.uk/ws/portalfiles/portal/158931832/BSRLM_CP_40_1_04.pdf (accessed on 25 October 2023).
- Hackerson, E. L., Slominski, T., Johnson, N., Buncher, J. B., Ismael, S., Singelmann, L., Leontyev, A., Knopps, A. G., McDarby, A., Nguyen, J. J., Condry, D. L. J., Nyachwaya, J. M., Wissman, K. T., Falkner, W., Grieger, K., Montplaisir, L., & Hodgson, A. (2024). Alternative grading practices in undergraduate STEM education: A scoping review. *Disciplinary and Interdisciplinary Science Education Research*, *6*, 15. [CrossRef]
- Haddaway, N. R., Page, M. J., Pritchard, C. C., & McGuinness, L. A. (2022). PRISMA2020: An R package and shiny app for producing prisma 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Systematic Reviews*, *18*, e1230. [CrossRef] [PubMed]
- Hammoudi Halat, D., Soltani, A., Dalli, R., Alsarraj, L., & Malki, A. (2023). Understanding and fostering mental health and well-being among university faculty: A narrative review. *Journal of Clinical Medicine*, *12*, 4425. [CrossRef]
- Harrington, B., Galal, A., Nalluri, R., Nasiha, F., & Vadarevu, A. (2024, March 20–23). *Specifications and contract grading in computer science education*. 55th ACM Technical Symposium on Computer Science Education V. 1, Portland, OR, USA.
- Harsy, A., Carlson, C., & Klamerus, L. (2021). An analysis of the impact of mastery-based testing in mathematics courses. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, *31*, 1071–1088. [CrossRef]

- Hassencahl, F. (1979). Contract Grading in the Classroom. *Improving College and University Teaching*, 27, 30–33. [CrossRef]
- Helmke, B. P. (2019, June 15–19). *Specifications grading in an upper-level BME elective course*. ASEE Annual Conference & Exposition, ASEE Conferences, Tampa, FL, USA.
- Henriksen, M., Kotas, J., & Wentworth, M. (2020). Specifications-based grading reduces anxiety for students of ordinary differential equations. *CODEE Journal*, 13, 1. [CrossRef]
- Hofmeister, E. H., Fogelberg, K., Conner, B. J., & Gibbons, P. (2022a). Specifications grading for veterinary medicine. *Journal of Veterinary Medical Education* 50, 15–18. [CrossRef] [PubMed]
- Hofmeister, E. H., Fogelberg, K., Conner, B. J., & Gibbons, P. (2022b). Specifications grading in a cardiovascular systems course: Student and course coordinator perspectives on the impacts on student achievement. *Journal of Veterinary Medical Education*, 50, 172–182. [CrossRef] [PubMed]
- Hollinsed, W. C. (2018). Applying innovations in teaching to general chemistry. In *Increasing Retention of Under-Represented Students in STEM through Affective and Cognitive Interventions* (Vol. 1301, pp. 145–152). American Chemical Society.
- Houseknecht, J. B., & Bates, L. K. (2020). Transition to Remote Instruction Using Hybrid Just-in-Time Teaching, Collaborative Learning, and Specifications Grading for Organic Chemistry 2. *Journal of Chemical Education*, 97, 3230–3234. [CrossRef]
- Howitz, W. J., Frey, T., Saluga, S. J., Nguyen, M., Denaro, K., & Edwards, K. D. (2023). A specifications-graded, sports drink-themed general chemistry laboratory course using an argument-driven inquiry approach. *Journal of Chemical Education*, 100, 672–680. [CrossRef]
- Howitz, W. J., McKnelly, K. J., & Link, R. D. (2021). Developing and implementing a specifications grading system in an organic chemistry laboratory course. *Journal of Chemical Education*, 98, 385–394. [CrossRef]
- Hunter, R. A., Pompano, R. R., & Tuchler, M. F. (2022). Alternative Assessment of Active Learning. In *Active Learning in the Analytical Chemistry Curriculum* (Vol. 1409, pp. 269–295). American Chemical Society.
- Johanesen, K. E., Claiborne, L. L., Falk, E. S., Hubbard, K. P., Kohfeld, K. E., Nadin, E. S., & Schmidt, A. H. (2024). Common-Sense teaching for the 2020s: Ungrading in response to COVID-19 and beyond. *Journal of Geoscience Education*, 72, 422–437. [CrossRef]
- Johnson, S. (2023). An alternative grading approach in undergraduate anatomy & physiology courses. *HAPS Educator*, 27, 99–108. [CrossRef]
- Jones, P. A. (2020). Implementing specifications grading in MPA courses: A potential strategy for better work-life balance. *Journal of Public Affairs Education*, 26, 531–547. [CrossRef]
- Joseph, M. L., Miller, S. W., Diec, S., & Augustine, J. M. (2023). Successes and challenges in implementing specifications grading in skills-based laboratory courses: Experiences at two colleges of pharmacy. *Currents in Pharmacy Teaching and Learning*, 15, 186–193. [CrossRef]
- Joshi, A. (2023, October 18–21). *Assessing the impact of specifications grading on a data visualization course*. 2023 IEEE Frontiers in Education Conference (FIE) (Vol. 25, pp. 1–6), College Station, TX, USA.
- Katzman, S. D., Hurst-Kennedy, J., Barrera, A., Talley, J., Javazon, E., Diaz, M., & Anzovino, M. E. (2021). The effect of specifications grading on students' learning and attitudes in an undergraduate-level cell biology course. *Journal of Microbiology & Biology Education*, 22, e00200–e00221. [CrossRef]
- Kelz, J. I., Uribe, J. L., Rasekh, M. F., Takahashi, G. R., Gibson, W. S., Link, R. D., McKnelly, K. J., & Martin, R. W. (2023). Implementation of specifications grading in an upper-division chemical biology lecture course. *The Biophysicist*, 4, 11–29. [CrossRef]
- Kiefer, S. F., & Earle, A. J. (2023, June 25–28). *Work in progress: Specifications grading in a system modeling course*. 2023 ASEE Annual Conference & Exposition, ASEE Conferences, Baltimore, MD, USA.
- Kinnear, G., Wood, A. K., & Gratwick, R. (2022). Designing and evaluating an online course to support transition to university mathematics. *International Journal of Mathematical Education in Science and Technology*, 53, 11–34. [CrossRef]
- Knight, M., & Cooper, R. (2019). Taking on a new grading system: The interconnected effects of standards-based grading on teaching, learning, assessment, and student behavior. *NASSP Bulletin*, 103, 65–92. [CrossRef]
- Kulik, C. -L. C., Kulik, J. A., & Bangert-Drowns, R. L. (1990). Effectiveness of mastery learning programs: A meta-analysis. *Review of Educational Research*, 60, 265. [CrossRef]
- Largent, D. L. (2024). Specifications grading meets computer science: Implementations in skill-and project-based courses. In A. D. Tomlin, & C. M. Nowik (Eds.), *Effective alternative assessment practices in higher education*. Research, Theory, and Practice Within Academic Affairs, Information Age Publishing.
- LeHew, M. (2019). Specifications grading: Experimenting with a new form of student assessment. *Journal for Academic Excellence*, 6, 12–15.
- Leslie, P., & Lundblom, E. (2020). Specifications grading: What it is, and lessons learned. *Seminars in Speech and Language*, 41, 298–309. [CrossRef] [PubMed]
- Lewis, D. (2022). Impacts of standards-based grading on students' mindset and test anxiety. *JoSoTL*, 22. [CrossRef]
- Lillard, A. S., & Taggart, J. (2022). Reimagining assessment in a large lecture: An alternative approach inspired by Thomas Jefferson and Maria Montessori. *College Teaching*, 72, 168–180. [CrossRef]

- Lovell, M. D. (2018, June 24–27). *Board 44: Defining and assessing competencies in an undergraduate reinforced concrete design course*. 2018 ASEE Annual Conference & Exposition, ASEE Conferences, Salt Lake City, UT, USA.
- Ludwigsen, D. (2017). Effective and efficient grading: Tips for assessing student learning. *Proceedings of Meetings on Acoustics*, 30, 025001. [[CrossRef](#)]
- Martin, K. M., Fernandez, T. M., & Mangum, R. (2021, July 26–29). *Work in progress: Examining engineering seniors students' perception of justice and fairness of grading practices*. 2021 ASEE Virtual Annual Conference Content Access, ASEE Conferences, Virtual.
- Martin, L. J. (2019). Introducing components of specifications grading to a general chemistry I course. In *Enhancing Retention in Introductory Chemistry Courses: Teaching Practices and Assessments* (Vol. 1330, pp. 105–119). American Chemical Society.
- Mattfeld, R. S. (2023). Improving student motivation through an alternative grading system. *Journal of Computing Sciences in Colleges*, 39, 86–95.
- McKnelly, K. J., Howitz, W. J., Thane, T. A., & Link, R. D. (2023). Specifications grading at scale: Improved letter grades and grading-related interactions in a course with over 1000 Students. *Journal of Chemical Education*, 100, 3179–3193. [[CrossRef](#)]
- McKnelly, K. J., Morris, M. A., & Mang, S. A. (2021). Redesigning a “writing for chemists” course using specifications grading. *Journal of Chemical Education*, 98, 1201–1207. [[CrossRef](#)]
- Mendez, J. (2018a, June 24–27). *Standards-based specifications grading in a hybrid course*. 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, USA.
- Mendez, J. (2018b, June 24–27). *Standards-based specifications grading in thermodynamics*. 2018 ASEE Annual Conference & Exposition, Salt Lake City, UT, USA.
- Mendez, J. (2019, November 9–15). *Development of a hybrid heat and mass transfer course*. ASME 2018 International Mechanical Engineering Congress and Exposition, Pittsburgh, PA, USA.
- Mendez, J. (2023, April 1–2). *Specifications grading in undergraduate fluid mechanics*. 2023 ASEE IL-IN Section Conference, Edwardsville, IL, USA.
- Mendez, J. (2024, April 13). *Work in progress: Specifications grading in mechanical engineering design courses*. ASEE IL-IN Section Conference, Chicago, IL, USA.
- Mio, M. J. (2024). Alternative grading strategies in organic chemistry: A journey. *Frontiers in Education*, 9, 1400058. [[CrossRef](#)]
- Mirsky, G. M. (2018). Effectiveness of specifications grading in teaching technical writing to computer science students. *Journal of Computing Sciences in Colleges*, 34, 104–110.
- Mirth, J. A. (2017, June 25–28). *Applying lean thinking to the structure and delivery of a kinematics course*. 2017 ASEE Annual Conference & Exposition, Columbus, OH, USA.
- Moster, C. A., & Zingales, S. K. (2024). Use of specifications-based grading in an online, asynchronous graduate organic chemistry course. *Frontiers in Education*, 9, 1379216. [[CrossRef](#)]
- Nilson, L. B., & Stanny, C. J. (2014). *Specifications grading: Restoring rigor, motivating students, and saving faculty time* (Reprint ed.). Stylus Publishing.
- Noell, S. L., Rios Buza, M., Roth, E. B., Young, J. L., & Drummond, M. J. (2023). A bridge to specifications grading in second semester general chemistry. *Journal of Chemical Education*, 100, 2159–2165. [[CrossRef](#)]
- Pascal, J., Vogel, T. J., & Wagstrom, K. (2020, June 22–26). *Grading by competency and specifications: Giving better feedback and saving time*. 2020 ASEE Virtual Annual Conference Content Access, ASEE Conferences, Online.
- Pope, L., Parker, H. B., & Ultsch, S. (2020). Assessment of specifications grading in an undergraduate dietetics course. *Journal of Nutrition Education and Behavior*, 52, 439–446. [[CrossRef](#)] [[PubMed](#)]
- Prasad, P. V. (2020). Using revision and specifications grading to develop students' mathematical habits of mind. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 30, 908–925. [[CrossRef](#)]
- Quintana, R., & Quintana, C. (2020). When classroom interactions have to go online: The move to specifications grading in a project-based design course. *Information and Learning Sciences*, 121, 525–532. [[CrossRef](#)]
- Reck, R. M. (2022, October 8–11). *Adopting alternative grading in an upper-level laboratory course in bioengineering*. 2022 IEEE Frontiers in Education Conference (FIE) (pp. 1–5), Uppsala, Sweden.
- Ring, J. (2017). ConfChem conference on select 2016 BCCE presentations: Specifications grading in the flipped organic classroom. *Journal of Chemical Education*, 94, 2005–2006. [[CrossRef](#)]
- Rojas, C., & Quan, G. M. (2023, June 25–28). *Mastery grading in a software engineering course*. 2023 ASEE Annual Conference & Exposition, ASEE Conferences, Baltimore, MD, USA.
- Rupakheti, C. R., Hays, M., Mohan, S., Chenoweth, S., & Stouder, A. (2018). On a pursuit for perfecting an undergraduate requirements engineering course. *Journal of Systems and Software*, 144, 366–381. [[CrossRef](#)]
- Saluga, S. J., Burns, A. M., Li, Y., Nguyen, M. M., & Edwards, K. D. (2023). A specifications-graded, spice-themed, general chemistry laboratory course using an argument-driven inquiry approach. *Journal of Chemical Education*, 100, 3903–3915. [[CrossRef](#)]
- Sanft, K. R., Drawert, B., & Whitley, A. (2021). Modified specifications grading in computer science: Preliminary assessment and experience across five undergraduate courses. *Journal of Computing Sciences in Colleges*, 36, 34–46.

- Santucci, A., & Golas, J. C. (2023). Grading for learning at the university of Rhode Island. *NEFDC Exchange*, 37, 7–12.
- Schinske, J., & Tanner, K. (2014). Teaching more by grading less (or differently). *CBE—Life Sciences Education*, 13, 159–166. [CrossRef] [PubMed]
- Seymour, E. (1997). *Talking about leaving: Why undergraduates leave the sciences* (1st ed.). Westview Press.
- Shields, K., Denlinger, K., & Webb, M. (2019). Not missing the point (s): Applying specifications grading to credit-bearing information literacy classes. In M. N. Mallon, L. Hays, C. Bradley, R. Huisman, & J. Belanger (Eds.), *The grounded instruction librarian: Participating in the scholarship of teaching and learning*. Association of College and Research Libraries.
- Smeding, A., Darnon, C., Souchal, C., Toczek-Capelle, M. -C., & Butera, F. (2013). Reducing the socio-economic status achievement gap at university by promoting mastery-oriented assessment. *PLoS ONE*, 8, e71678. [CrossRef] [PubMed]
- Sorensen-Unruh, C. (2024). The Ungrading learning theory we have is not the ungrading learning theory we need. *CBE Life Sciences Education*, 23, es6. [CrossRef]
- Streifer, A. C., & Palmer, M. S. (2021). Is specifications grading right for me?: A readiness assessment to help instructors decide. *College Teaching*, 71, 244–251. [CrossRef]
- Stutzman, R. Y., & Race, K. H. (2004). EMRF: Everyday rubric grading. *The Mathematics Teacher*, 97, 34–39. [CrossRef]
- Suresh, M. A. (2023, June 25–28). *Challenges and experiences in implementing a specifications grading system in an upper-division undergraduate computer networks course*. 2023 ASEE Annual Conference & Exposition, ASEE Conferences, Baltimore, MD, USA.
- Talbert, R. (n.d.). *The EMRN rubric*. Available online: <https://rtalbert.org/emrn/> (accessed on 25 October 2023).
- Tamés, D. (2021, July 27–30). *Implementation and evaluation of specifications grading in a video production course* [Paper presentation]. 75th University Film and Video Association (UFVA) Conference: Empowering Spaces, Online.
- Taylor, H. (1980). *Contract grading*. EPIC Clearinghouse on Tests, Measurement, and Evaluation.
- Toledo, S., & Dubas, J. M. (2017). A Learner-centered grading method focused on reaching proficiency with course learning outcomes. *Journal of Chemical Education*, 94, 1043–1050. [CrossRef]
- Trachsler, T. A., Morris, E., & Mahoney, T. Q. (2023). Specifications grading in the sport management classroom: Breakdown of the System and Reflections Upon Implementation With Relation to Outcomes Assessment. *Sport Management Education Journal*, 17, 121–128. [CrossRef]
- Tsoi, M. Y., Anzovino, M. E., Lin Erickson, A. H., Forringer, E. R., Henary, E., Lively, A., Morton, M. S., Perell-Gerson, K., Perrine, S., Villanueva, O., Whitney, M., & Woodbridge, C. M. (2019). Variations in implementation of specifications grading in STEM courses. *Georgia Journal of Science*, 77, 10.
- Tuson, E., & Hickey, T. (2022). Mastery learning and specs grading in discrete math. In *Proceedings of the 27th ACM conference on innovation and technology in computer science education* (Vol. 1, pp. 19–25). Association for Computing Machinery.
- Tuson, E., & Hickey, T. (2023, March 15–18). *Mastery learning with specs grading for programming courses*. 54th ACM technical symposium on computer science education V. 1 (pp. 1049–1054), Toronto, ON, Canada.
- Villalobos, C., Garza, G., Setayesh, S., Fernandez, L., Balogh, A., Serbin, K. S., & Huber, T. J. (2024). Our never-ending pathway to innovate calculus 1: Course coordination and active learning to specifications grading and growth mindset. *Problems, Resources, and Issues in Mathematics Undergraduate Studies, Advance online publication*, 1–17. [CrossRef]
- Vitale, S. E., & Concepción, D. W. (2021). Improving student learning with aspects of specifications grading. *Teaching Philosophy*, 44, 29–57. [CrossRef]
- Walden, P. R. (2022). Student motivation, anxiety and pass/fail grading: A SoTL project. *Teaching and Learning in Communication Sciences & Disorders*, 6, 13. [CrossRef]
- Wasniewski, E., Munro, T., & Tandon, T. (2021). Exploring Online Specifications Grading: An Undergraduate Course Case Study. In T. J. Bastiaens (Ed.), *Proceedings of the EdMedia + innovate learning* (pp. 179–184). Association for the Advancement of Computing in Education (AACE).
- Williams, K. (2018). Specifications-based grading in an introduction to proofs course. *Problems, Resources, and Issues in Mathematics Undergraduate Studies*, 28, 128–142. [CrossRef]
- Winstone, N., Bourne, J., Medland, E., Niculescu, I., & Rees, R. (2021). “Check the grade, log out”: Students’ engagement with feedback in learning management systems. *Assessment & Evaluation in Higher Education*, 46, 631–643. [CrossRef]
- Yang, M. A., & Korsnack, K. (2024). Pairing a bioinformatics-focused course-based undergraduate research experience with specifications grading in an introductory biology classroom. *Biology Methods and Protocols*, 9, bpae013. [CrossRef] [PubMed]
- Yik, B. J., Machost, H., Streifer, A. C., Palmer, M. S., Morkowchuk, L., & Stains, M. (2024). Students’ perceptions of specifications grading: Development and evaluation of the perceptions of grading schemes (pgs) instrument. *Journal of Chemical Education*, 101, 3723–3738. [CrossRef]

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.