TCC 2020 Conference Papers

Shifting to Specifications Grading: Two Design Cases

Vanessa P. Dennen
Florida State University, USA
vdennen@fsu.edu

Lauren M. Bagdy
Florida State University, USA
lbagdy@fsu.edu

Abstract: Most formal learning experiences result in some sort of assessment and acknowledgement of learning outcomes, which may range from a certificate of completion to a grade on a rating scale. Whatever system is used, instructors and students benefit from clear expectations and well-aligned learning objectives and assessment measures. Specifications grading is one approach that promotes a focus on achieved learning outcomes by clearly articulating expectations and the relationship between competencies and grades. This paper presents two university-level course design cases, one undergraduate and one graduate, in which a shift was made to a specifications grading system. The redesign required careful consideration of key competencies, competency indicators, mastery thresholds, and revision opportunities. At each course level there were different challenges to address, reflecting different levels of course difficulty and anticipated student maturity. Issues such as providing student feedback in a specifications system, handling situations where students challenge the system, and conversion to standard university grading scales also are discussed. Benefits included increased student self-regulation and ownership of the learning and assessment processes.

Introduction

At the end of a course, a grade provides a single indicator of student performance in a course. However, the meaning behind this indicator can vary. In some instances, a final grade may be the average of student scores on objective tests. In these cases, the grade is a direct indicator of performance on the tests. It represents a student's overall attainment of course learning outcomes to the extent that the tests were valid measures of these outcomes. In other courses, final grades are a multifaceted measure that consider factors beyond just student achievement. For example, grades may reflect student work that is completed on time, and active participation during class activities. Although the notion that grades may represent more than just mastery of learning objectives may sound problematic initially, the combined elements of a grading system, inclusive of work habits and disposition, may be a good predictor of future success (Brookhart et al., 2016).

Grades and the grading process are often a source of anxiety for both student and instructors, and can be the greatest source of conflict in a course context (Placier, 1995). From the student perspective, the grading process may lack transparency and feel

subjective. Students submit their assignments and hope for a good grade, but they may lack confidence in that grade until they see the grade posted by the instructor. In turn, instructors struggle with issues such as determining how many points to deduct for different problems and justifying those point deductions in terms that students understand. Well-constructed rubrics can help mitigate some of these challenges by articulating expectations and related point values or ranges (Jonsson & Svingby, 2007). Still, students occasionally submit assignments that defy a rubric, or a rubric may not fully capture the potential of a creative student assignment (Popham, 1997), leaving instructors with little guidance for assessing that assignment. Even when clear justification is provided or when rubrics are used, students may try to negotiate grades on individual assignments, seeking small point increases in an attempt to reach a higher letter grade on an individual assignment or to influence a cumulative grade total.

This paper proposes specifications grading as a solution to these issues. Specifications grading is a form of grading that firmly anchors the assessment process in student learning outcomes (Nilson, 2015). The concept is simple: Students are provided with clearly written, measurable specifications that their assignments must meet, and then their work is graded according to those specifications. By tying grades directly to achieved outcomes, this approach adds rigor to the grading process and can be designed to communicate the attainment of necessary competencies in professional fields (Bonner, 2016). Specifically, this paper discusses the design process used to integrate specifications grading into two courses, focusing on how existing course policies, grading structures, and assessments were modified and then implemented in a learning management system (LMS).

Specifications Grading

Specifications grading can focus on whether a student has met outcomes, with a binary pass or fail grade provided by the instructor. A threshold or benchmark can be established to determine what percentage or combination of the specifications the students must successfully meet in order to pass. The specifications part of specifications grading may feel like a natural approach for anyone who already engages in a backwards design (Wiggins, Wiggins, & McTighe, 2005) or traditional instructional systems design (e.g., Dick, Carey, & Carey, 2005) approach. Similarly, the pass/fail part of this system will be familiar to people who have engaged in mastery learning experiences. For example, both written and practical driving tests are assessed on the basis of sufficient mastery. On the written component of a driving test, there is typically a minimum pass score representing a certain percentage of correct answers on questions of cognitive knowledge. Learners earning the minimum passing score or a perfect score all have the same outcomes. On the practical component, a different approach is taken. Not all demonstrated skills are considered equal, and failure to demonstrate competency on a single major skill might result in a failed test even if all other skills are executed correctly.

From the instructor perspective, specifications grading streamlines the grading process. Students either meet the specifications or do not. Assignments that do not meet the specifications can simply be returned to the student as an assignment not passed, with no

agonizing over points or partial grades. There is empirical support for shifting to pass-fail or similar grading systems. Students graded using a pass-fail system reported lower stress than students who were graded on an A-F system (Rohe et al., 2006). In medical education, some schools have shifted to pass-fail systems and have reported positive effects on student well-being along with no problems related to academic outcomes (Spring, Robillard, Gehlbach, & Moore Simas, 2011).

A specifications grading system may initially appear to be a high stakes system for students. However, students who do not at first succeed can have second chances through the use of a token system. Students can be provided with tokens at the beginning of the course. These tokens can be redeemed for the opportunity to revise an assignment that did not meet specifications, submit late work, skip a minor assignment, or drop a low grade. Students also may be given the opportunity to earn tokens. In one study, students responded positively to a token economy in which active course participation yielded tokens to be redeemed for other purposes within the course (Boniecki & Moore, 2003). Instructors can design a token system in whatever manner best suits the needs of the students within a specific course, ensuring that there is enough opportunity for students to have second chances without heading down the path of endless revision and feedback.

Method

These design cases were drawn from field notes kept during the design process, course documents, observations, and interactions with students. The redesign process was carefully documented for the two courses, with close attention paid to decision points about the overall structuring of assessments for the courses. Analysis focused on reconstructing the phases of the overall project, from the initial impetus for redesign through the initial implementation and outcomes. Next steps in this work are to systematically collect data on student reactions and outcomes in the undergraduate class.

Assessment Redesign Context

Two existing classes were redesigned to adopt a specifications grading system. One was an online, graduate-level course focused on applied learning technologies, and the other was an undergraduate educational technology course for preservice teachers. Both courses were completely designed with detailed course assessments and grading criteria prior to the redesign. The graduate level course was redesigned first, and then, using lessons learned in that process, the undergraduate course was redesigned.

In both cases, the impetus for assessment redesign was a sense that student work and grades at times could unevenly reflect achievement of learning outcomes, along with a desire to lessen instructor frustration and fatigue related to denoting and justifying every point deduction. For example, in the graduate course some assignments required two weeks of planning and work to do well. Some students clearly tried to backfill required parts of the assessment and complete it in its entirety a few days before the due date. Consequently, the assignment would fit what was asked for in format, but the breadth of data the student would work with and the depth of insight offered in their data analysis

would be less than what was expected. Providing detailed comments about these assignment shortcomings and determining a suitable grade was time-consuming and demoralizing. In the undergraduate class, students tended to compare point deductions with classmates and challenge grades even when the feedback already made clear why points had been deducted.

The authors were both involved in the assessment redesign process. The first author was the instructor of record for the graduate course and the supervising faculty member for the undergraduate course. The second author was the teaching assistant for the graduate course and the lead instructor for the undergraduate course.

Redesign Process

The assessment redesign process included both conceptual tasks and technological tasks. The first conceptual task was to decide how the overall grade system would work. Specifications grading can be done at multiple levels; individual assignments, assignment groups, or whole courses can apply the approach (Nilson, 2015). In both of these courses, we felt that it would be too big of a cultural change at the university if specifications grading were used to determine final grades, and neither class was approved as a pass-fail course. Thus, we needed to have a system in which students could earn a range of grades based on specifications met. In the graduate course, we decided that each major assignment would have a point value, and two variants (concept, full) worth different points. In the undergraduate course, we opted for point values with two levels (pass, high pass). The point values for each assignment were calibrated so that a student completing the lower level for each assignment would earn a "C" in the undergraduate class, or a "B" in the graduate class.

Next, we had to design the specifications for each assignment. Fortunately, the existing assignments had detailed grading criteria that became the specifications. Had these criteria not already existed, this task might have been more involved. All grading criteria were reviewed for clarity, and we also considered how effectively they aligned with the course learning objectives. Here are sample specifications for a graduate-level literature search and annotated bibliography assignment:

- Search approach is systematic and planned
- Documentation of search captures and reports information about number and type of hits from various sources
- Accurate characterization of research in two topical areas (identification of authors, themes, research designs, journals, etc.)
- Five relevant articles from each topical area are succinctly summarized following the provided example
- Accurate APA formatting for all in-text citations and references
- Accurate identification and classification of empirical research articles

Once the specifications were listed, the second conceptual task was to determine which specifications must be satisfied in order to earn a particular grade. In the undergraduate

class, this meant determining which specifications represented a pass, and which represented a high pass. Most of the undergraduate assignments focused on demonstrating technical and design skills, and those were split into two lists. Students who completed all of items on the basic skills list receive a pass and students completing all advanced skills could earn a high pass.

In the case of the graduate class, where the original assignments required detailed participation across time in order to successfully fulfill the specifications, the solution was to develop an alternate assignment that could be completed in less time while still meting the baseline specifications. In other words, students who choose to complete the concept version can develop and demonstrate the same core knowledge and skills, albeit in a less rich manner. For example, the following are the assignment overviews for two versions of an assignment focused on communities of practice:

> FULL VERSION
> Participate in 2 or more social media-based communities of a related genre/topic/focus for 2 weeks. In one you should be a consumer/lurker and in one you should be an active contributor. Write a 5 page double-spaced paper comparing and contrasting various elements of the communities, focusing on membership, trajectories, norms, moderation, knowledge sharing, privacy, and ethics. Include a log of your community interactions. Support your findings with at least 4 references from journal articles.
>
> CONCEPT VERSION
> Read 4 journal articles about online communities. In a 4-5 page double-spaced paper, synthesize the main points about membership, trajectories, norms, moderation, knowledge sharing, privacy, and ethics.

In the full version, students engage with authentic communities of practice for 2 weeks and connect their observations to what they read in the literature. In contrast, in the concept version students simply read and synthesize the literature. Both versions have students fully engaged with the learning objectives, just with different levels of depth.

Finally, a token system was designed in each class to facilitate giving students second chances. In the graduate class, each student started with three tokens that could be used to revise assignments or submit assignments three days late. The undergraduates received seven tokens and could also use them for a free pass on a blog post or in-class assignment. The instructors were told that they could offer opportunities to earn extra tokens if they felt their students needed them.

Once the specifications system was designed, it was time to sort out the technological components and figure out how it would function within Canvas, the course LMS. The assignments were easily set up, with rubrics designating the three levels of assignment completion (e.g., no pass, pass, high pass). Additionally, students were asked to indicate via submission comments which assignment version they were submitting. This approach

allows for a streamlined gradebook with only on submission option per assignment, while acknowledging the different versions and their respective point values.

Tokens were set up as ungraded quizzes. By using the quiz function it became easy to track how many tokens each student had used. When students requested to use a token, it would show up as an item to be graded. The instructor would then mark it complete and put a comment in the assignment submission (e.g., “token applied for late submission” or “token applied for assignment revision”) so the token would be considered when grading.

Results

Students have reacted positively to the specifications system. Initially there was a learning curve as they adjusted to the expectations, but tokens were helpful for this purpose. Students who did not attend to specifications on their first assignment receive a “no pass” grade, submit a token, and revise to a higher grade. Then they learned the expectations to read carefully and be thorough in their work. This was a bigger issue for undergraduate students than it was for graduate students.

Students reported to their instructors that they enjoyed the extra control, knowing what grade they would earn if they followed the specifications and that they could have a second chance if they did not succeed on their first try. In both classes, few students used all of their tokens. Those who used all did so deliberately. Interestingly, grade distributions did not change after moving to a specifications system, but students who earned lower grades did so by submitting higher quality work. In other words, they excelled at addressing the specifications level that they chose to focus on. Also, students appeared to be satisfied with their grades, which clearly represented work completed.

Conclusion

This foray into specifications grading provided the opportunity to carefully consider both learning objectives and student expectations in both classes. In the system that we designed, although pass grades were awarded for the attainment of competencies, there was also an element of rewarding good student behavior, demonstrating the same complexities of many contemporary grading systems (Brookhart et al., 2016). Specifications grading systems can empower students by taking the mystery out of grading and allowing them to choose what work they will do for their grades. The design process could be time-consuming for instructors who do not already have extensive and clearly articulated grading criteria or rubrics. However, it ultimately can yield a more pleasant grading experience for instructors and satisfying learning experiences for students.

References

Boniecki, K. A., & Moore, S. (2003). Breaking the silence: Using a token economy to reinforce classroom participation. *Teaching of Psychology, 30*(3), 224-227.

Bonner, M. W. (2016). Grading rigor in counselor education: A specifications grading framework. *Educational Research Quarterly, 39*(4), 21-42.

Brookhart, S. M., Guskey, T. R., Bowers, A. J., McMillan, J. H., Smith, J. K., Smith, L. F., . . . Welsh, M. E. (2016). A century of grading research. *Review of Educational Research, 86*(4), 803-848. doi:10.3102/0034654316672069

Dick, W., Carey, L., & Carey, J. O. (2005). *The systematic design of instruction* (6th ed.). New York, NY: Allyn & Bacon.

Jonsson, A., & Svingby, G. (2007). The use of scoring rubrics: Reliability, validity and educational consequences. *Educational Research Review, 2*(2), 130-144. doi:10.1016/j.edurev.2007.05.002

Nilson, L. B. (2015). *Specifications grading: Restoring rigor, motivating students, and saving faculty time*: Stylus Publishing.

Placier, M. (1995). "But I have to have an A": Probing the cultural meanings and ethical dilemmas of grades in teacher education. *Teacher Education Quarterly, 22*(3), 45-62.

Popham, W. J. (1997). What's wrong — and what's right — with rubrics. *Educational Leadership, 55*(2), 72-75.

Rohe, D. E., Barrier, P. A., Clark, M. M., Cook, D. A., Vickers, K. S., & Decker, P. A. (2006). *The benefits of pass-fail grading on stress, mood, and group cohesion in medical students.* Paper presented at the Mayo Clinic Proceedings.

Spring, L., Robillard, D., Gehlbach, L., & Moore Simas, T. A. J. M. e. (2011). Impact of pass/fail grading on medical students' well-being and academic outcomes. *45*(9), 867-877.

Wiggins, G., Wiggins, G. P., & McTighe, J. (2005). *Understanding by design*. Alexandria, VA: ASCD.