Engineering Educators Bringing the World Together
2025 ASEE Annual Conference & Exposition
Palais des congrès de Montréal, Montréal, QC • June 22–25, 2025 ASEE
Paper ID #46248

Impact of Ungrading on Student Confidence in Capstone Engineering Design

Duncan Davis-Hall, Colorado School of Mines
Carter Moulton, Colorado School of Mines

Impact of Ungrading on Student Confidence in Capstone Engineering Design

Introduction

This study investigates the hypothesis that ungrading a capstone engineering design course will foster student confidence. Ungrading is an emergent approach to assessment that seeks to address the ineffectiveness of and potential harm caused by traditional grading practices [1]. Ungrading can take many forms, but in this case involves replacing points-based grades with a variety of alternative grading techniques. These techniques include *narrative and peer evaluations* that prioritize qualitative feedback on student work, *collaborative agreements* which allow the professor and students to develop rubrics and assessment criteria together, and *collaborative assessment* whereby students self-evaluate, justify, and discuss their final course grade using evidence from their work throughout the semester. In this work in progress submission, we ask: if grade-oriented environments promote a fear of failure and risk-taking even in high-achieving students [2, 3], how might ungrading impact student confidence?

An extensive body of work related to the efficacy of alternative grading has been published over the past few decades. Butler and Nisan published a seminal article in 1986 showing that prioritizing feedback over grades helped maintain intrinsic motivation in students. They found that traditional points-based grading led students to emphasize the importance of grades, disincentivized creativity, and decreased interest in learning as a process. Alternatively, students that received qualitative feedback without a grade attached performed significantly better on problem-solving tasks [4]. A review by Alfie Kohn states that students tend to think less deeply, avoid risks, and lose interest in learning when they are graded [5]. Another study found that grades negatively affect student confidence, and that uncertainty around grades can perpetuate assessment anxiety and disincentivize collaboration [6]. Grading in its traditional sense does not provide effective feedback and does not tend to change students' future efforts – particularly in courses that emphasize problem solving and creativity, like engineering design [7].

The capstone design course under evaluation emulated a realistic, authentic [8] work environment by incorporating self-evaluation and peer assessment along with instructor feedback. To further the goals of authentic assessment, assignments in this course were not graded with points. Instead, the instructor made comments and asked questions that engaged students' work rather than simply evaluating it. The instructor provided feedback in the form of narrative evaluations with a qualitative Excellent, Meets Expectations, Revisions Necessary, and Not Assessable (EMRN) grade scale [9]. Students were frequently asked to assess their own learning as the class progressed. The intention was to help students focus on operating in a more organic way that imitates work in a post-college environment.

Grading focused on qualitative feedback instead of quantitative point calculations. Below is language used in the course syllabus to help explain our reasoning:

> "This means I will not be grading your assignments with points, but rather responding and asking questions. There will be suggested avenues for you to pursue in the next iteration, but no points. You'll be doing these things for your teammates, yourselves, and your classmates too, as we engage with each other's projects as a learning community."

Overview of Intervention

We evaluated a bioscience-focused engineering capstone design class during the Fall 2024 semester. Confidence in various course learning outcomes was measured through a mixed methods design that consisted of surveys with open responses and five-point Likert scale questions. Participants' responses to surveys addressed the hypothesis by providing quantitative and qualitative data regarding confidence in course content during a semester of ungrading.

This ungraded course incorporated self-evaluation and peer evaluation for members in the same design team, which grounds the study methodology in the framework of social cognitive theory. Broadly speaking, social cognitive theory emphasizes interpersonal interactions as a determining factor for student behavior. Prior work has shown that peer feedback can be more impactful, and more relevant, than traditional instructor feedback while also developing students' metacognitive thinking [1, 7, 10].

Students used instructor feedback to self-evaluate their overall course performance in a collaborative assessment. The instructor did not assign A-F scale letter grades through narrative evaluations of each assessment submission, but students referenced the qualitative feedback in these narrative evaluations in a reflective and self-evaluative letter in which they proposed and justified a final letter grade using course learning outcomes – defined in the syllabus and at the start of each lesson – and evidence from their work. The instructor met with each student one-on-one to discuss these collaborative assessments and final grades.

Surveys were administered to students three times during the semester: week 5, week 10, and week 15. The questions in the surveys were identical during each iteration. The first section of each survey asked students to self-evaluate their performance on a scale of 1-5. Aspects of their performance included quality of work, team membership, and communications. Additionally, some context was provided as to what each end of the five-point Likert scale could represent:

- Quality of work: 1 – Has done minimal work or very poor-quality work. 5 – Has done exceptional work consistently and reliably. Exceeds expectations.
- Team membership: 1 – Works separately with little or no interaction with the team. 5 – Has been an integral and important team member; has made significant contributions to the overall solution; attends and participates in all team meetings.
- Communications: 1 – Produces no useful documentation of work done; communication skills poor. 5 – Documents work exceptionally well; has exceptional communication skills.

Each survey included a free response text box for students to specify something they were doing well and something that could use improvement with reference to the self-evaluation. Students were also asked to use the criteria above to evaluate their teammates' performance on the design project and given the same free response prompt for each teammate.

A series of five-point Likert scale questions (1 = strongly disagree to 5 = strongly agree) were used to assess student attitude towards the alternative grading methods in this course. Those questions are shown in the figures below and were designed to evaluate student confidence and the efficacy of instructor feedback, peer evaluation, and self-evaluation in course learning outcomes and progress on capstone design projects. A free response text box followed these questions and asked students whether there were any particular reasons for the confidence ratings they selected.

Finally, to align performance in class with institutional course grading policies, students were asked which letter grade (A, B, C, D, or F) best represented their course effort so far, and were asked to explain their reasoning for that grade. Actual final letter grades were proposed and justified in a separate letter submitted by each student at the end of the semester.

Findings

The three types of feedback in this course – self-evaluation, peer evaluation, and instructor-provided narrative assessments – were regarded as accurate and helpful by students. Trends indicated an overall increase in this perception as the semester progressed, although self-evaluated team membership saw a gradual decline in average Likert scale response. Notably, self-evaluated communication ability, perceived accuracy of feedback, and helpfulness of peer evaluation all increased over the semester (Figure 1, next page).

Students maintained high confidence in course learning outcomes throughout the semester. Importantly, students' confidence in their team's ability to complete their design project increased over time. Comments left by students reflect this confidence, with mostly positive sentiment in mid- and end term comments, but mostly neutral sentiment at the start of the semester (Figure 2, next page). Examples of student comments are shown below Figure 2.

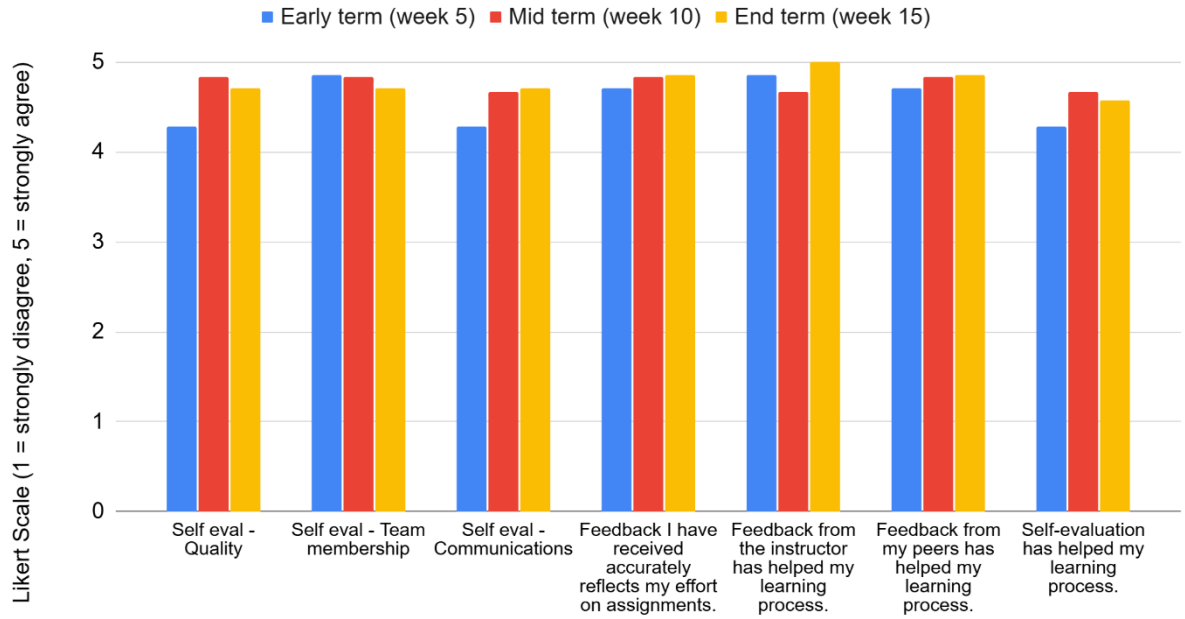

Figure 1. Self-evaluation, peer evaluation, and narrative assessments were effective forms of feedback. Students agreed or strongly agreed that feedback through alternative grading was helpful throughout the semester. Bars represent mean Likert scale score from n = 7 students.

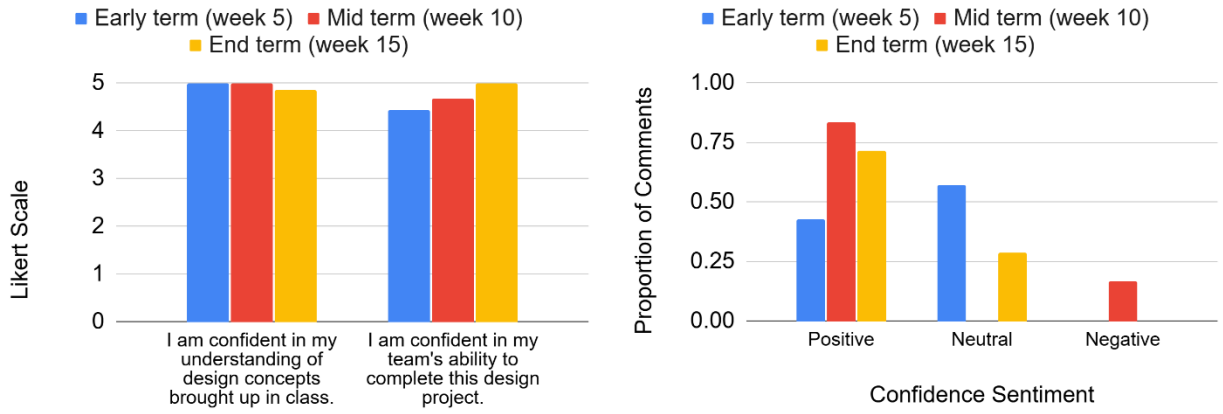

Figure 2. Students showed high confidence regarding course concepts and team success. (Left) students agreed or strongly agreed with statements about confidence related to design concepts (course learning outcomes) and their team's ability to complete a design project. Bars represent mean Likert scale from n = 7 students. (Right) free-response comments explaining Likert-scale confidence ratings were mostly positive or neutral throughout the semester. Bars represent proportion of comments from n = 7 students.

Below is an example of a neutral early term comment:

> "I think our team seems to have a strong conceptual understanding of the work so far and have shown up quite strong. I do have some concerns about the clarity of the project right

now in terms of the client, but I expect these to be resolved relatively soon as more communication occurs."

Included here is an example of a positive mid term comment showing confidence:

> "After the most recent client meeting I feel much more confident about the work ahead of us. Things feel much more clearly defined in terms of expected outcomes and in terms of scale this feels much more appropriate for the time frame.... In terms of design concepts, I feel like the lessons have done a good job of explaining them and giving opportunities to discuss or explore in different activities."

When asked to propose letter grades based on their progress at early, mid, and end term checkpoints, students evaluated themselves quite highly, almost entirely at the A level. Interestingly, comments meant to justify these grade proposals were not entirely positive until the end of the semester (Figure 3). Here is an example of an end term positive comment:

> "I think that at this point all of us have put a large amount of effort and time into this project and it turned out very well with tangible results, presentable information and overall good teamwork."

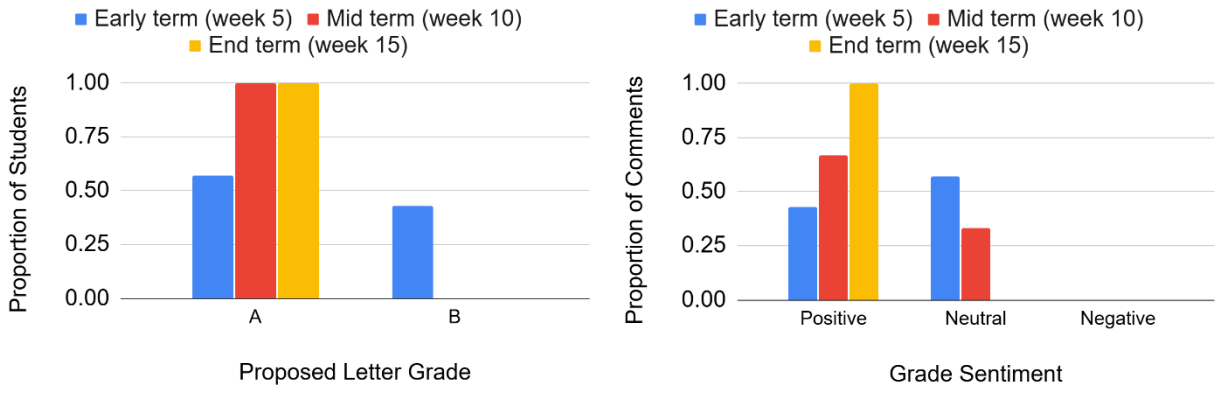

Figure 3. Students proposed high letter grades and reflected positively on their proposals. (Left) students proposed either A's or B's at the start of the semester, and only A's at the end of the semester. (Right) free-form responses justifying grade proposals were positive or neutral. Bars represent proportions of n = 7 students.

These findings demonstrate that alternative grading techniques – including self-evaluation, peer evaluation, narrative assessments, and collaborative assessments – were well-received by students. Students reported agreement with statements of value surrounding ungrading. Students also reported a high degree of confidence in design course learning outcomes and in their ability to complete a client-centered design project. These sentiments culminated in high proposed course grades, another indication of confidence.

Reflections/Next Steps

Next steps for this work-in-progress will involve enrolling more students in a larger class size to establish a more generalizable study population. Future survey questions will also be tailored to address the specific aspects of our ungrading scheme to delineate benefits and challenges presented by the system. A larger study population will necessitate continuous improvements in the experimental procedure, mostly centered around communication with students and the integration of client feedback. The EMRN grading structure used for narrative evaluations effectively focused student resubmissions on areas of assignments that needed refining, and consistently revisiting the EMRN criteria with students will be important to align expectations. Specifically, the difference between E (excellent/exemplary) and M (meets expectations) was a common area of confusion in student-professor conversations. Establishing clear and consistent expectations for narrative evaluations can be achieved by revisiting the collaborative agreements more frequently in future semesters. Students were encouraged, but not required, to make rubrics that established expectations for each assignment as part of collaborative agreements. However, students did not take this extra step – further iterations of this class would benefit from more structured collaborative agreements. Overall, student sentiment showed that this was a promising intervention in an engineering design course. Students were highly motivated – this comment demonstrates that perception:

> “I think I've contributed very strongly to our project and have been an active participant in both the lecture portions of the course, but also in our design work. I feel as if my contributions have represented a good understanding of design principles and different project management skills."

References

[1] J. Stommel, “How to ungrade,” in S. Blum, Ed., Ungrading: Why Rating Students Undermines Learning (and What to Do Instead), West Virginia University Press, 2020, pp. 25–41.

[2] C. Pulfrey, C. Buchs, and F. Butera, “Why grades engender performance-avoidance goals: The mediating role of autonomous motivation,” Journal of Educational Psychology, vol. 103, no. 3, pp. 683–700, 2011. [Online]. Available: https://doi.org/10.1037/a0023911

[3] M. Schultz-Bergin, “Grade anarchy in the philosophy classroom,” in S. Blum, Ed., Ungrading: Why Rating Students Undermines Learning (and What to Do Instead), West Virginia University Press, 2020, pp. 173–187.

[4] R. Butler and M. Nisan, “Effects of no feedback, task-related comments, and grades on intrinsic motivation and performance,” Journal of Educational Psychology, vol. 78, pp. 210–216, 1986.

[5] A. Kohn, “The trouble with rubrics,” The English Journal, vol. 95, no. 4, 2006.

[6] W. Howitz, K. J. McKnelly, and R. D. Link, “Developing and implementing a specifications grading system in an organic chemistry laboratory course,” Journal of Chemical Education, vol. 98, pp. 385–394, 2021. [Online]. Available: https://doi.org/10.1021/acs.jchemed.0c00450

[7] J. Schinske and K. Tanner, “Teaching more by grading less (or differently),” CBE–Life Sciences Education, vol. 13, no. 2, pp. 159–166, 2014. [Online]. Available: https://doi.org/10.1187/cbe.cbe-14-03-0054

[8] G. Wiggins. “The Case for Authentic Assessment.” ERIC Digest, 1990. Available: https://eric.ed.gov/?id=ED328611

[9] R. Stutzman and K. Race, “EMRF: Everyday rubric grading,” Mathematics Teacher, vol. 97, no. 1, pp. 34–39, 2004.

[10] B. Millis, “Using Metacognition to Promote Learning.” IDEA Paper no. 63, pp. 1-9, 2016.