Beyond the Grade: Cultivating Deep Learning and Student Agency in a Thermodynamics Classroom

Andrea Giorgioni, The Pennsylvania State University - York

Andrea Giorgioni is an Assistant Teaching Professor of Engineering at Penn State York. He holds a Laurea in Ingegneria Chimica (M.S. in Chemical Engineering) from the Universita' degli Studi di Bologna, Italy. With over two decades of teaching experience at institutions including the New Jersey Institute of Technology, his current work focuses on innovative pedagogical strategies. He is actively involved in curriculum development and explores alternative assessment methods, such as ungrading, to enhance academic integrity and critical thinking in engineering education.

Redesigning Thermodynamics: A Hybrid Ungrading and Mastery Approach to Promote Integrity and Engagement

Abstract

In response to growing academic integrity concerns within engineering education, a new pedagogical approach has been developed to counter the challenges posed by technology and online resources. In a field where the rise of artificial intelligence (AI) tools hinders the development of critical thinking skills, this course redesign moves away from traditional grading. Instead, it implements a blend of ungrading and mastery grading to foster genuine student effort and a deeper understanding of the subject matter. This innovative strategy was piloted over two semesters in a Thermodynamics course for mechanical engineers with a small cohort (N=18), allowing for a focused learning environment.

This teaching method aims to enhance students' intrinsic motivation by shifting the focus from grades to the learning process itself. Ungrading minimizes the use of points, emphasizing detailed feedback to support improvement, while mastery grading structures the course around specific learning objectives. In this model, students demonstrate their proficiency by taking on the role of an instructor via "Teaching Presentations." This approach fosters deep learning, as the act of teaching a concept to others, known as the "protégé effect", requires a thorough understanding. To demonstrate mastery, students present step-by-step solutions for pre-assigned problems, gaining agency in how they communicate their reasoning to their peers.

Quantitative and qualitative feedback collected from the pilot demonstrates the efficacy of this hybrid model. Survey results indicate that 83% of students were satisfied with the knowledge gained, and 78% perceived the assessment system as fair. By verbalizing and structuring information for others, students reported reinforced knowledge and reduced assessment anxiety. This paper outlines the specifics of the course redesign, presents empirical evidence of its impact on engagement and integrity, and discusses strategies for scaling this approach to larger enrollment courses.

1. Introduction

Academic integrity has long been a concern in engineering education, but the rapid proliferation of online resource repositories (e.g., Chegg, CourseHero) and GenAI tools (e.g., ChatGPT, Claude) has fundamentally altered the landscape. In fields like Thermodynamics, which require heavy mathematical rigor, the pressure to achieve high grades often incentivizes students to bypass the "struggle" of learning in favor of finding the "right answer" online.

While GenAI offers pedagogical opportunities, it poses a threat to the validity of traditional take-home assessments. Dron argues that because AI can now emulate 'soft technique', the creative and idiosyncratic processes previously unique to humans, educators must adapt their process of assessment [1]. When a student can generate a correct solution in seconds, grading based on homework accuracy becomes a measure of resourcefulness rather than understanding. This paper presents a course redesign that moves away from traditional grading. Instead, it implements a blend of ungrading and mastery grading [2] to foster genuine student effort. The goal was to enhance intrinsic motivation, reduce assessment anxiety, and disincentivize academic dishonesty by focusing on the learning process.

2. Pedagogical Framework

The course redesign integrates two distinct pedagogical strategies:

A. Ungrading

Ungrading challenges the assumption that grades are the best metric for learning. Research suggests that grades can dampen intrinsic motivation and encourage students to focus on performance metrics rather than intellectual growth [3]. Crogman et al. argue that abandoning institutionalized assessment protocols in favor of feedback-oriented strategies can significantly improve student outcomes [4]. In this course, ungrading was employed to create a "safe to fail" environment, a method supported by findings on student perceptions of ungraded courses [5]. Homework and in-class exercises were assigned but not graded for accuracy; credit was awarded for completion and effort. This shifted the focus from "earning points" to utilizing feedback, an approach successfully piloted in various STEM contexts [6].

B. Mastery Grading and the "Student as Instructor"

To mitigate the utility of AI-generated solutions, this course introduced a "learning-through-teaching" model. Mastery was not demonstrated through written exams, but through "Teaching Presentations." This method is grounded in the "protégé effect," where the act of teaching material to others creates a social and cognitive commitment that reinforces the teacher's own understanding and effort [7]. A student relying on memorized AI solutions would struggle to explain the conceptual "why" during live questioning, making this a robust filter for authentic understanding.

3. Course Methodology

The pilot was conducted for two semesters in a sophomore-level Thermodynamics course for Mechanical Engineers (N=18: nine students per semester).

A. Removal of Traditional Exams and Graded Homework

Traditional high-stakes exams were eliminated. Homework was assigned but not graded for accuracy. Instead, students received full credit for attempting and submitting the work. Feedback was provided through collaborative in-class sessions where the instructor and students worked through problems together, addressing misconceptions in real-time.

B. In-Class FE-Style Exercises

To ensure students remained prepared for standardized engineering assessments, ungraded in-class exercises modeled on the Fundamentals of Engineering (FE) exam were introduced. These encouraged group work and provided immediate feedback on the students' grasp of core concepts without the anxiety of a recorded grade.

C. Teaching Presentations (The Mastery Component and Academic Integrity)

The primary summative assessments were two "Teaching Presentations" scheduled throughout the semester. In these sessions, students acted as the instructor, guiding the class through a step-by-step solution to a pre-assigned complex problem. This method serves as an oral defense of knowledge, requiring the student to explain the "why" behind every step and respond to real-time inquiries from peers and the instructor.

A critical requirement of this process is a mandatory formative checkpoint: students must meet with the instructor prior to their presentation. This meeting serves two essential mastery functions. First, it acts as a proficiency threshold; if the student's work does not meet the required standard, they are provided with targeted feedback and required to iterate on their solution until mastery is demonstrated. Second, this iterative process serves as a robust safeguard for academic integrity. By requiring a preliminary one-on-one review and a subsequent live oral defense, the assessment makes the student's learning journey visible. This structure ensures that the work is the student's own and not the result of "cognitive offloading" to Generative AI. Because students must justify their reasoning and adapt to unscripted questions in person, they cannot rely on AI-generated outputs. Mastery, in this context, is verified not just by the accuracy of the final answer, but by the student's ability to authentically own and communicate the logic behind it.

4. Results and Student Feedback

To evaluate the impact of this redesign, anonymous feedback was collected via Microsoft Forms at the end of each semester. 18 students responded to the survey in total.

A. Student Satisfaction and Perception of Fairness

The response to this redesigned model was overwhelmingly positive.

Satisfaction: When asked, "How satisfied are you with the knowledge you gained?", 83% of students (15 out of 18) reported being "Satisfied" or "Very Satisfied."

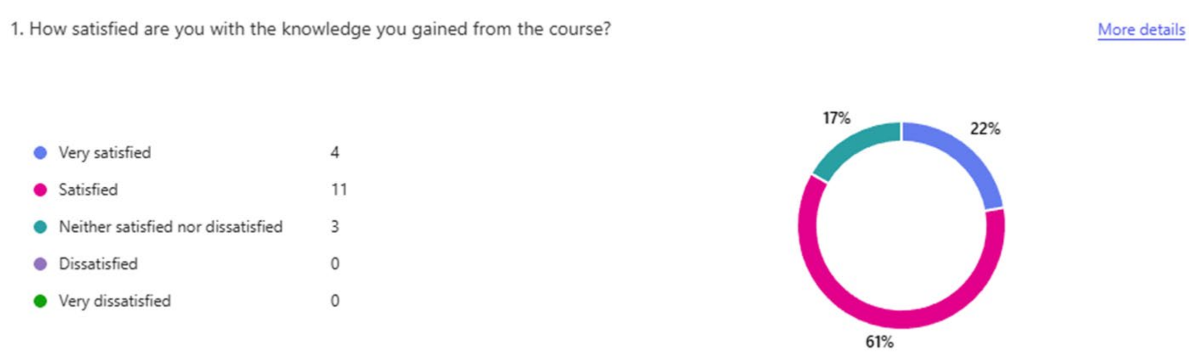

Figure 1: Question 1 of the Feedback Survey

Fairness: When asked if they were assessed fairly, 78% (14 out of 18) responded "Yes," with only one student responding "No."

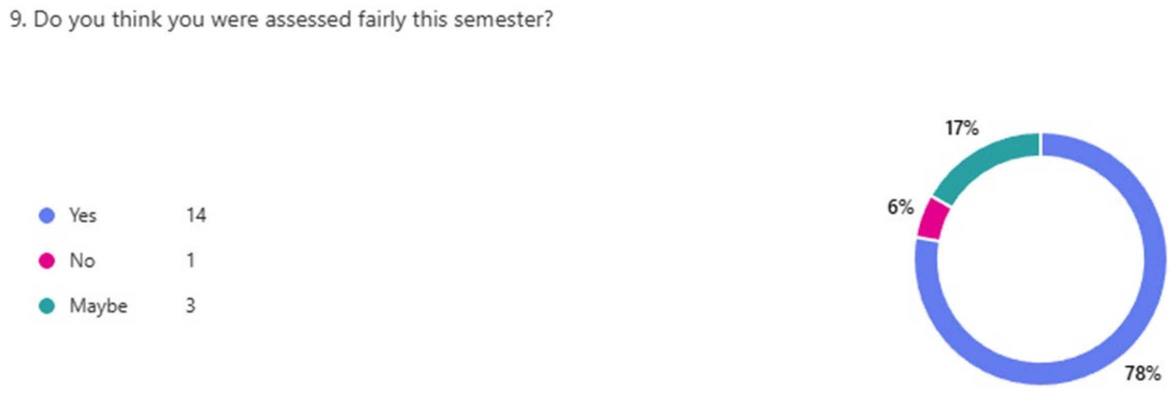

Figure 2: Question 9 of the Feedback Survey

Qualitative feedback highlighted the reduction of stress as a primary benefit:

"Tests stress me out, so having this class be more collaborative than anything really helped me figure out how to solve problems."

"I feel like the grading/assessment in this class was very conducive to a good learning environment. It helped to limit stress while focusing on the material itself."

B. Shifting Motivation

A core goal was shifting focus from grades (extrinsic) to learning (intrinsic). The survey asked what typically motivates the student's best work. The cohort was split:

39% (7 students) selected "Getting a good grade."

61% (11 students) selected options related to personal interest, academic interest, or meeting course expectations.

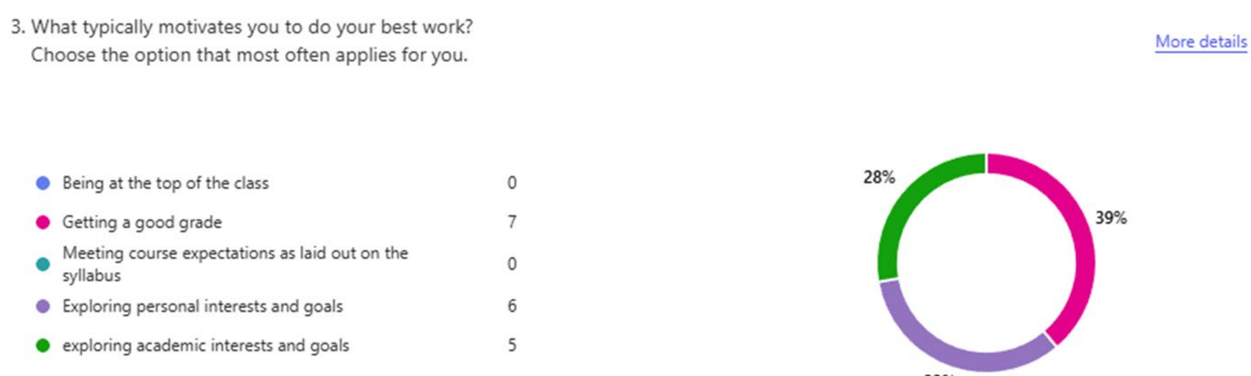

Figure 3: Question 3 of the Feedback Survey

However, regarding this specific course structure, 72% (13 out of 18) indicated that the grading system encouraged them to do their best work.

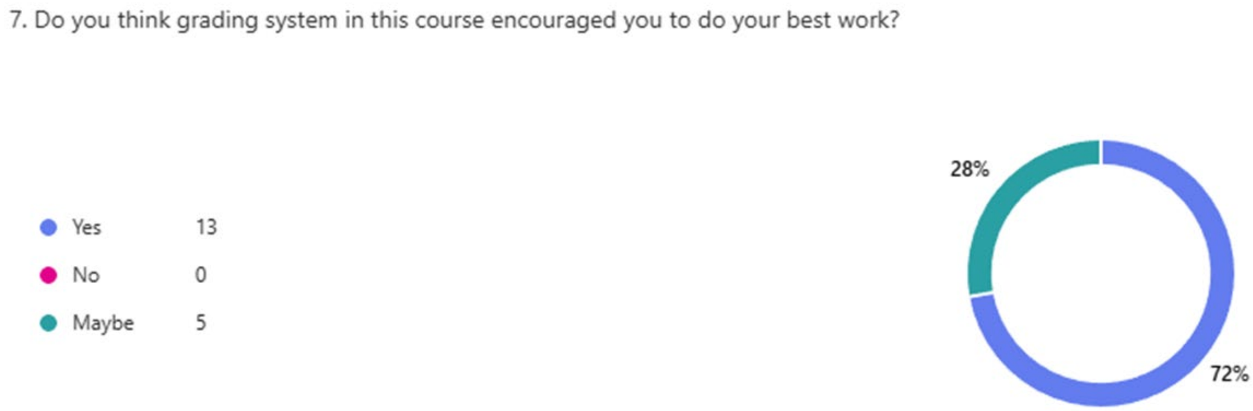

Figure 4: Question 7 of the Feedback Survey

Students noted that removing the fear of failure actually increased their engagement:

"It encouraged me to do my best work because I felt like I had to attempt each question fully to eventually get a genuine understanding of the concept... whether it was that night or in class when it was explained."

"I was not focused on getting the questions right, but rather learning the concepts and asking why?"

C. Effectiveness of the Teaching Presentations

Students ranked the course components by effectiveness. The Oral Presentations were ranked #1 (most effective), followed by Homework (#2) and Classwork (#3). Students recognized that the presentations forced a deeper level of study than traditional exams:

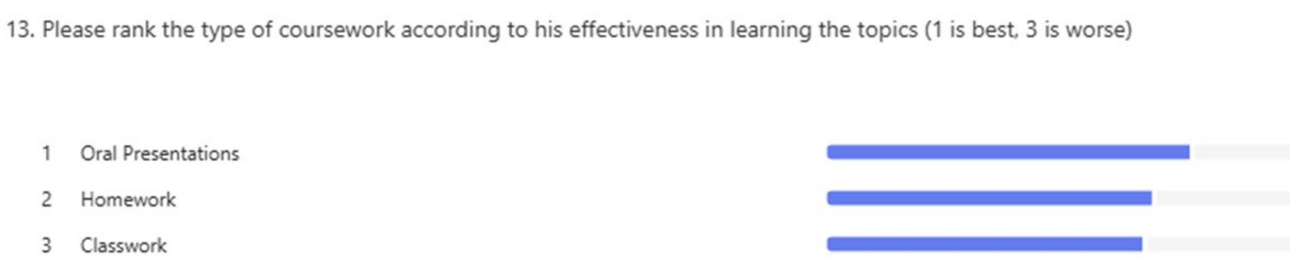

Figure 5: Question 13 of the Feedback Survey

"The oral exams are a good method of examination... I find them more challenging because when I study for them or prepare my presentation, I find gaps in my understanding that I had no idea existed until I have to recall it for those purposes."

D. Areas for Improvement

The survey also revealed valid critiques.

Accountability Gaps: Some students felt that without the pressure of graded quizzes, they struggled to gauge their own retention.

"I feel like the course would benefit from small, low stakes quizzes... There were times in the course when I thought that I had understood the material... only for you to later reference back to that material... where I then realized I had not understood it."

In-Class Review Time: Some students noted that reviewing ungraded homework in class was time-consuming.

"I think that by assigning so much homework, it took away from class time... I would spend less time going over homework and classwork in class, as that took away from time we could have used to introduce new topics."

5. Discussion

A. Integrity through Oral Defense

The survey results validate the hypothesis that oral presentations serve as a viable alternative to written exams in the AI era. Students acknowledged that the presentations verified their knowledge in ways that simple completion grading could not. The student who noted that preparation revealed "gaps in understanding" highlights the metacognitive benefit of this mastery approach.

B. Balancing "Ungrading" with Accountability

While the reduction in stress was universal, the data suggests that a purely ungraded model may leave some students unsure of their academic standing until the summative presentation. As suggested by the student feedback, a hybrid approach incorporating low-stakes, automated quizzes (e.g., Canvas quizzes) could provide the necessary "checkpoints" for students to gauge their retention without reintroducing the anxiety of high-stakes exams.

C. Scalability and Resource Management

A limitation of this study is the class size (9 students per semester). The "student as instructor" model is time-intensive requiring significant instructor time for the formative checkpoint and class time for the live presentation. For larger cohorts, where faculty workload and scheduling constraints become prohibitive, this approach would require adaptation. Potential scaling strategies could include:

- Peer Review: Using a validated rubric, peers could assess the clarity and accuracy of the teaching explanation, reducing the instructor's grading load.
- Recitation Sections: conducting presentations in smaller break-out groups led by Graduate Teaching Assistants or the student themselves if using a peer review system.
- Video Submissions: Students could record their teaching presentations, allowing for asynchronous grading.

The critique regarding class time is significant, yet it highlights a tension between efficiency and deep learning. The in-class homework review was designed not merely for error correction, but as a synchronous "modeling" phase where the instructor and students collaboratively navigate the engineering problem-solving process. While some students, particularly those who successfully solved the problems independently, may view this time as superfluous, removing it entirely risks losing a critical feedback loop. To address the time concern without sacrificing this pedagogical benefit, future iterations will adopt a "triage" model. By using pre-class polls to identify which specific problems caused the most difficulty, synchronous class time can be reserved for deep dives into those stumbling blocks, while straightforward solutions could be relegated to asynchronous video resources.

6. Conclusion

The goal of this redesign was not to ban AI, but to render it insufficient for passing. As AI tools evolve, engineering education must move toward assessments that require human communicative agency and critical defense of ideas, skills that generative AI currently cannot authentically replicate in a live environment.

This pilot demonstrates that a hybrid ungrading and mastery model can successfully reduce student anxiety and foster deep learning in a Thermodynamics course. 83% of students reported satisfaction with their knowledge gain, and the majority felt the system encouraged their best work. By shifting the summative assessment to an oral defense (Teaching Presentations), the course effectively neutralized the threat of AI-generated cheating while empowering students to take ownership of their learning. Future iterations will refine the model by introducing low-stakes checkpoints to improve student self-assessment and optimizing in-class time management. A possible future direction for this work is to establish a comparison framework, such as using concept inventories, to assess direct learning outcomes. Another interesting line of investigation involves examining how extensive instructor contact time, as opposed to traditional graded homework, influences the degree of student independence and preparation required for successful teaching presentations.

References

[1] J. Dron, "The Human Nature of Generative AIs and the Technological Nature of Humanity: Implications for Education," Digital, vol. 3, no. 4, p. 319, 2023.

[2] R. Campbell, D. Clark, and J. O'Shaughnessy, "Introduction to the special issue on implementing mastery grading in the undergraduate mathematics classroom," PRIMUS, vol. 30, no. 8–10, pp. 837–848, 2020.

[3] E. Dosmar and J. Williams, "PUT DOWN THE GRADE BOOK," ASEE Prism, vol. 32, no. 1, p. 53, 2022.

[4] H. T. Crogman et al., "Ungrading: The case for abandoning institutionalized assessment protocols and improving pedagogical strategies," Education Sciences, vol. 13, no. 11, p. 1091, 2023.

[5] D. Guberman, "Student Perceptions of an Online Ungraded Course," Teaching and Learning Inquiry, vol. 9, no. 1, pp. 86-98, 2021.

[6] C. Von Renesse and S. A. Wegner, "Two examples of ungrading in higher education in the United States and Germany," PRIMUS, vol. 33, no. 9, p. 1035, 2023.

[7] C. C. Chase, D. B. Chin, M. A. Oppezzo, and D. L. Schwartz, "Teachable agents and the protégé effect: Increasing the effort towards learning," J. Sci. Educ. Technol., vol. 18, no. 4, pp. 334–352, Aug. 2009.