ASEE AMERICAN SOCIETY FOR ENGINEERING EDUCATION

Paper ID #40993

Student and Instructor Feedback on an AI-Assisted Grading Tool

Dr. David Calamas, Georgia Southern University

David Calamas is an Associate Professor in the Department of Mechanical Engineering at Georgia Southern University. He received a BS in Mechanical Engineering from Clemson University and an MS and PhD in Mechanical Engineering from the University of Alabama.

Student and Instructor Feedback on an AI-Assisted Grading Tool

Abstract

A browser-based AI-assisted grading tool was deployed in an undergraduate mechanical engineering course, offering the instructor online grading capabilities, time savings, a transparent and unbiased evaluation process, personalized feedback, and detailed analytics for student performance visualization. A Likert-scale-based survey instrument gauged responses from a cohort of undergraduate mechanical engineering students, probing the efficiency of Gradescope, grading transparency, bias reduction, and overall satisfaction. Preliminary analysis indicates positive sentiment, with students expressing satisfaction with the speed and transparency of grading. The platform's ability to provide detailed feedback on problem-solving steps emerged as a significant benefit, contributing to a more standardized and unbiased grading experience. This study offers nuanced insights into student perspectives on technology-enhanced grading tools, contributing to discussions on digital platforms in academia.

Introduction

Recent years have witnessed a transformative shift in educational assessment with technology-driven solutions. Acknowledging the need for efficient, transparent, and unbiased grading methods, this paper explores Gradescope, a browser-based AI-assisted grading tool, in a Heat Transfer course within a Bachelor of Science in Mechanical Engineering program. Gradescope offers advantages in grading efficiency, transparency, and bias mitigation, aligning with the evolving landscape of educational assessment.

From the author's perspective, Gradescope presents three main advantages: (1) an efficient grading workflow, (2) transparency in assessment, and (3) bias mitigation. The tool streamlines grading, replacing traditional manual methods with an automated and time-efficient approach. It allows rapid AI-assisted assessment of concept questions and offers partial credit allocation for problem-solving steps. Gradescope enhances transparency by providing detailed feedback on each question, fostering a culture of continuous improvement. Its rubric-based approach ensures consistent evaluation, reducing potential biases associated with manual assessment.

Research Questions

This paper addresses three main research questions:

1. How does Gradescope impact grading efficiency in an undergraduate Heat Transfer course?
2. To what extent does Gradescope contribute to transparency in assessment, and how do students perceive the feedback provided?
3. How does Gradescope mitigate grading biases compared to traditional manual methods?

Understanding the implications of integrating Gradescope in an educational framework is crucial for informing future pedagogical practices. This study contributes valuable insights that can be applied to undergraduate engineering programs and the broader academic community seeking

effective tools for assessment enhancement. By examining Gradescope's impact on student learning, feedback mechanisms, and grading efficiency, this paper aims to contribute to the ongoing discourse on leveraging technology to advance the quality of engineering education.

Course Information and Assessment

Gradescope was implemented in Heat Transfer, a required course in most Bachelor of Science in Mechanical Engineering degree programs. The course is a traditional lecture-based course with three credit and three contact hours. The prerequisites for the course are Fluid Mechanics and Heat Transfer. In the author's course, exams consist of ten concept questions and three problem-solving questions. The concept questions consist of multiple-choice, true/false, fill-in-the-blank-style questions worth one point each. The problem-solving questions are worth thirty points, and partial credit is awarded for procedural steps. Historically, the author graded the exams "by hand." While the concept questions could be assessed with scannable answer sheets (e.g., Scantron), the problem-solving questions were graded by hand. Unfortunately, the grading of the problem-solving questions was time-consuming. More importantly, because the exams weren't anonymous, the elimination of implicit (or explicit) bias in the assessment could not be demonstrated to students. Finally, providing detailed feedback efficiently was difficult. So, the author piloted Gradescope to assess the exams and distributed a survey to garner student perceptions of the browser-based assessment tool. A sample of the feedback report generated by Gradescope for a student exam can be found in the Appendix.

Gradescope Workflow

Gradescope facilitates the evaluation of paper-based exams, quizzes, bubble sheets, and homework by instructors. Gradescope also allows instructors to create online assignments, including programming tasks, for student responses on the platform. It is worth noting that this paper exclusively focuses on paper-based assignments. Within this context, instructors have the flexibility to choose from diverse question types, such as open-ended responses, proofs, diagrams, multiple-choice, true/false, and fill-in-the-blank. Auto-grading applies to specific question types like multiple-choice and fill-in-the-blank.

For paper-based exams, the grading process involves several key steps. Instructors begin by uploading a template, essentially a blank copy of the exam. Subsequently, they delineate the question regions on the template where students are expected to demonstrate their work. Following this, instructors establish a rubric for each question, consisting of descriptive criteria and assigned point values. The next step involves uploading scanned copies of the student exams, which are then auto-assigned to students based on their name or university ID number. Instructors proceed to grade the student work using a dynamic rubric. Once the grading is finalized, instructors can publish grades and communicate them to students via email. Additionally, the platform supports online submission of regrade requests by students. Instructors also have the capability to export grades and access question and rubric-level statistics, enabling a more nuanced assessment of student learning outcomes.

Survey Design

A Likert scale was used to measure the students' attitudes, opinions, and perceptions regarding using Gradescope as an exam assessment tool. Using a Likert scale provided a structured way to collect quantitative data, and the standardized format ensured consistency in data collection. The survey included questions on efficiency, transparency, bias reduction, and overall satisfaction. The survey also had several open-ended questions to receive qualitative feedback. A copy of the survey can be found in the Appendix.

Student Background and Demographics

Of the forty-four enrolled students, thirty-nine (88.64%) participated in the survey. The students' self-reported race/ethnicity and gender identity (as defined in Stanford's IDEAL Diversity, Equity, and Inclusion Survey [1]) can be found in Table 1. Most students (76.92%) had no prior experience with online grading tools, but 92.31% were comfortable with technology-assisted grading (see Table 2 in the Appendix).

Table 1. Self-Reported Student Demographic Data

Gender Identity			Race/Ethnicity		
	n	%		n	%
Man	31	79.49%	Asian or Asian American	3	7.69%
Woman	7	17.95%	Black or African American	11	28.21%
No Response	1	2.56%	Hispanic or Latino/a	1	2.56%
Total	39		White or European	23	58.97%
			No Response	1	2.56%
			Total	39	

Quantitative Feedback

Quantitative results from the survey can be found in Table 2 in the Appendix. The majority of students (92.31%) expressed a positive sentiment regarding the efficiency of Gradescope in the grading process for the course. A rating of 5 (strongly agree) was given by a substantial number (62.23%), with only a few providing lower scores. This suggests that most students believed that the grading process was made more efficient through Gradescope. The results also revealed that most students agreed that assessment via Gradescope was more transparent than manual "paper-based" assessment. Specifically, 87.18% of the students expressed agreement with the statement, providing ratings of 4 or 5. A small percentage (12.82%) chose a neutral rating. This indicates a positive perception among the students regarding the transparency of the grading process with the use of Gradescope. Overall, the sentiment leans towards a positive perception of Gradescope contributing to transparency in the grading process compared to traditional manual methods. Finally, the majority of students (84.62%) agreed or strongly agreed that the use of Gradescope

had reduced bias in grading. Some students (10.26%) expressed a neutral stance on this aspect, while 2.56% expressed disagreement with the statement.

Qualitative Feedback

Student feedback on the open-ended questions was centered around four coherent themes: consistency and standardization, anonymity and bias reduction, clarity and transparency, and efficiency and turnaround time. The responses reflected a consensus among students on the perceived fairness of grading using Gradescope compared to traditional manual methods. Students appreciated the uniformity in grading standards, emphasizing that everyone was subject to the same criteria, eliminating potential biases. The anonymity of the grading process was highlighted as a key factor contributing to fairness, ensuring equal treatment for all without the influence of personal factors.

The efficiency of Gradescope was acknowledged, with students expressing satisfaction with the standardized and automated grading process. The system's ability to provide clear and transparent criteria for point allocation was seen as advantageous, allowing students to understand where they gained or lost points in a detailed manner. Breaking down the grading into individual parts was particularly valued, providing a structured and transparent view of the assessment.

Some students noted that traditional manual grading might offer a deeper understanding of individual students and their problem-solving mindset. However, the majority of responses emphasized the benefits of Gradescope's standardized approach, minimizing subjectivity and inconsistencies associated with manual grading. The overall sentiment suggested that Gradescope enhanced fairness by providing all students a clear, standardized, and unbiased grading experience.

Conclusions

This study reveals positive student perceptions of Gradescope in a mechanical engineering course, emphasizing its efficiency, transparency, and perceived fairness. Survey results indicated strong agreement (92.31%) that Gradescope made grading more efficient, with 87.18% agreeing on increased transparency and 84.62% perceiving bias reduction. Overall satisfaction with Gradescope reached 97.44%, reflecting a positive student reception. Student feedback garnered from the open-ended questions clustered around four themes: consistency and standardization, anonymity and bias reduction, clarity and transparency, and efficiency and turnaround time. Students appreciated Gradescope's consistent grading criteria, the anonymity fostering fairness, transparent feedback, and the system's overall efficiency. The findings underscore the importance of student feedback in shaping and optimizing grading technologies. While Gradescope proves valuable, ongoing refinement is essential to address concerns, ensuring the continued effectiveness of automated grading tools in academic settings. This study contributes to the broader conversation on enhancing grading processes through technology in engineering education. Future work could delve deeper into addressing specific student concerns, refining the system's recognition of diverse problem-solving approaches, and ensuring that the benefits of automation do not compromise the nuanced understanding that manual grading may provide.

References

[1] “IDEAL Diversity Equity and Inclusion Survey,” idealdeisurvey.stanford.edu. https://idealdeisurvey.stanford.edu/ (accessed Nov. 24, 2023)

Appendix

Survey

Demographics

1. (Optional) Self-reported race/ethnicity (as specified in Stanford's IDEAL Survey). My race/ethnicity is best described as:
    - American Indian or Alaska Native
    - Asian or Asian American
    - Black or African American
    - Hispanic or Latino/a
    - Middle Eastern or North African
    - Native Hawai`ian or Pacific Islander
    - White or European
    - Prefer not to say
2. (Optional) Self-reported gender identity (as specified in Stanford's IDEAL Survey). My gender or gender identity is best described as:
    - Gender nonconforming
    - Genderqueer
    - Man
    - Nonbinary
    - Questioning
    - Woman
    - Trans
    - Prefer not to say

Prior Experience with Technology-Assisted Grading Tools

> Please rate the following statement on a 5-point Likert scale (1: very uncomfortable, 2: uncomfortable, 3: neutral, 4: comfortable, 5: very comfortable).

3. How comfortable are you with technology-assisted grading tools like Gradescope?
4. Have you taken any other courses with online grading tools before (e.g., Gradescope)?
    a. Yes
    b. No

Perceptions of Gradescope

> Please rate the following statements on a 5-point Likert scale (1: strongly disagree, 2: disagree, 3: neutral, 4: agree, 5: strongly agree).

5. The use of Gradescope has made the grading process in this course more efficient.
6. I believe the grading using Gradescope is more transparent than manual grading.
7. The use of Gradescope has reduced bias in grading.

Open-Ended Questions

8. Please share any specific feedback or comments about your experience with Gradescope in this course. What aspects do you find most beneficial, or do you have any concerns?
9. How do you perceive the fairness of grading on Gradescope compared to traditional manual grading methods? Explain your reasoning.
10. Are there any specific challenges or difficulties you've encountered when using Gradescope for assignments/exams in this course? If so, please describe.
11. If you have any suggestions for improving the use of Gradescope in this course, please share them here.

Overall Satisfaction

Please rate the following statement on a 5-point Likert scale (1: very dissatisfied, 2: dissatisfied, 3: neutral, 4: satisfied, 5: very satisfied).

12. How satisfied are you with the use of Gradescope in this course?
13. Would you recommend the use of Gradescope for grading to other course instructors?
    a. Yes
    b. No

Table 2. Quantitative Data from Likert-Scale-Based Questions

	n	1		2		3		4		5				
	n	n	%	n	%	n	%	n	%	n	%	Positive	Neutral	Negative
How comfortable are you with technology-assisted grading tools like Gradescope?	39	0	0.00%	0	0.00%	3	7.69%	13	33.33%	23	58.97%	92.31%	7.69%	0.00%
The use of Gradescope has made the grading process in this course more efficient.	39	0	0.00%	1	2.56%	2	5.13%	9	23.08%	27	69.23%	92.31%	5.13%	2.56%
I believe the grading using Gradescope is more transparent than manual grading.	39	0	0.00%	0	0.00%	5	12.82%	11	28.21%	23	58.97%	87.18%	12.82%	0.00%
The use of Gradescope has reduced bias in grading.	39	0	0.00%	2	5.13%	4	10.26%	6	15.38%	27	69.23%	84.62%	10.26%	5.13%
How satisfied are you with the use of Gradescope in this course?	39	0	0.00%	0	0.00%	1	2.56%	16	41.03%	22	56.41%	97.44%	2.56%	0.00%

Table 3. Data from Yes/No Response Questions

		Yes		No	
	n	n	%	No	%
Have you taken any other courses with online grading tools before (e.g., Gradescope)?	39	9	23.08%	30	76.92%
Would you recommend the use of Gradescope for grading to other course instructors?	39	38	97.44%	1	2.56%

MENG 3233 Exam 2

[redacted]

TOTAL POINTS

94 / 100

QUESTION 1

1 C1 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: A

QUESTION 2

2 C2 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: A

QUESTION 3

3 C3 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: B

QUESTION 4

4 C4 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: B

QUESTION 5

5 C5 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: D

QUESTION 6

6 C6 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: C

QUESTION 7

7 C7 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: C

QUESTION 8

8 C8 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: A

QUESTION 9

9 C9 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: A

QUESTION 10

10 C10 **1 / 1**

✓ **+ 1 pts** *Correct*

+ 0 pts Correct Answer: C

QUESTION 11

11 PS1 **27 / 30**

+ 30 pts Correct

+ 3 pts Characteristic length

✓ **+ 3 pts** *Biot number*

✓ **+ 3 pts** *Criteria for lumped system analysis*

✓ **+ 3 pts** *Inverse of time constant, _b_*

✓ **+ 3 pts** *Lumped system analysis equation*

✓ **+ 3 pts** *Solve for time*

✓ **+ 3 pts** *Units and unit conversions are present and correct (for example, 1 N = 1 kg·m/s²).*

✓ **+ 3 pts** *Equations are presented in their variable form (for example, F = ma)*

✓ **+ 3 pts** *Equations are shown with numbers inserted (for example, F = ma = (10 kg)(9.81 m/s²)).*

✓ **+ 3 pts** *Calculations are complete for all attempted steps (for example, F = ma = (10 kg)(9.81 m/s²) = 98.1 kg·m/s² = 98.1 N).*

+ 0 pts Problem not attempted or entirely incorrect.

QUESTION 12

12 PS2 **30 / 30**

✓ **+ 30 pts** *Correct*

+ 1.5 pts Surface area

+ 3 pts Equation for drag force due to friction

+ 3 pts Solve for coefficient of friction

+ 3 pts Momentum heat transfer analogy

+ 1.5 pts Definition of Stanton number

+ 3 pts Solve for heat transfer coefficient

+ 3 pts Heat transfer rate

+ 3 pts Units and unit conversions are present and correct (for example, 1 N = 1 kg·m/s²).

+ 3 pts Equations are presented in their variable form (for example, F = ma)

+ 3 pts Equations are shown with numbers inserted (for example, F = ma = (10 kg)(9.81 m/s²)).

+ 3 pts Calculations are complete for all attempted steps (for example, F = ma = (10 kg)(9.81 m/s²) = 98.1 kg·m/s² = 98.1 N).

+ 0 pts Problem not attempted or entirely incorrect.

QUESTION 13

13 PS3 **27 / 30**

+ 30 pts Correct

✓ **+ 4.5 pts** *Maximum velocity*

✓ **+ 3 pts** *Reynolds number*

✓ **+ 3 pts** *Nusselt number*

✓ **+ 3 pts** *Correction factor from table*

✓ **+ 1.5 pts** *Corrected Nusselt number*

+ 3 pts Heat transfer coefficient

✓ **+ 3 pts** *Units and unit conversions are present and correct (for example, 1 N = 1 kg·m/s²).*

✓ **+ 3 pts** *Equations are presented in their variable form (for example, F = ma)*

✓ **+ 3 pts** *Equations are shown with numbers inserted (for example, F = ma = (10 kg)(9.81 m/s²)).*

✓ **+ 3 pts** *Calculations are complete for all attempted steps (for example, F = ma = (10 kg)(9.81 m/s²) = 98.1 kg·m/s² = 98.1 N).*

+ 0 pts Problem not attempted or entirely incorrect.