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 #48273

The Relationship Between Student Sentiment and Academic Performance using Student Reflections from a Flipped, Mastery-Based Statics Course

Dr. Amie Baisley, University of Florida

Amie Baisley is currently the Thomas O. Hunter Rising Star Instructional Assistant Professor at the University of Florida teaching primarily 2nd year mechanics courses. Her teaching and research interests are alternative pedagogies, mastery-based learning and assessment, student persistence in their first two years, and faculty development.

Chiranjeevi Singh Marutla, University of Florida

The Relationship Between Student Sentiment and Academic Performance using Student Reflections from a Flipped, Mastery-Based Statics Course

This study explores the relationship between student sentiment on post-assessment reflections and academic performance in a flipped, mastery-based Statics course. Utilizing Natural Language Processing (NLP) techniques, we analyzed over 28,000 student comments collected from four semesters with each student contributing an average of 54 comments per semester. This data provides a rich and comprehensive source of qualitative data for each student. Sentiment analysis was applied to each written comment to categorize the tone as either positive, negative, or neutral. Following the assessments conducted biweekly, students submitted a reflection as part of the learning process, commenting on each part of their solution process. These comments provide insights into their cognitive and emotional engagement with the course material. To better understand each student's sentiment, we developed an approach to quantify their overall sentiment score for each assessment. This method enables us to track the evolution of each student's tone throughout the semester.

Our study will determine how this evolution in student sentiment correlates with their final grade in the course, identifying whether emotional tone in reflections is linked to academic performance. These trends could provide insights into how students perceive and engage with their work and how it aligns with the course metrics. This study also highlights new opportunities for targeted interventions in the course. Through leveraging NLP and reflective exercises, instructors gain access to more detailed and individualized insights into class progress. This can foster a better understanding of the connection between student attitudes and performance, enabling more personalized feedback and tailored interventions that can improve learning outcomes.

Introduction

Reflection is an important skill that contributes to continuous learning and understanding personal growth and can have major impacts when integrated into education. The use of reflection in engineering education closely aligns with ABET's criteria to develop lifelong learners [1]. Lifelong learning requires students to develop metacognitive skills including the ability to evaluate their knowledge, reflect on experiences, and process strategies to grow from it [2, 3]. It can create self-awareness, critical thinking, and the ability to better adapt which are essential skills for engineers. Recently, reflection opportunities have been integrated into engineering classrooms in a variety of ways including reflective journaling, reflective portfolios, and end of project or semester evaluations [4-6]. These methods have had benefits for the students even if the results are not explicitly clear in academic metrics.

There have been studies in engineering education to show that reflective practices can improve student problem-solving skills, increase their confidence, and enhance their understanding of

material [4, 5]. The types of reflections vary including structured self-assessments where students are asked to evaluate their performance [7]. However, there are many challenges with capturing reflections which are often done through journaling or other forms of writing, or through verbal communication which results in a large volume of student submissions in a classroom [6]. The process of reviewing these reflections individually is resource intensive; however, there is much value in it to be able to close the loop and support the students through feedback based on their reflection.

Advancements in artificial intelligence (AI), particularly natural language processing (NLP), have created pathways to identify meaning from written or communicated work [8]. NLP allows for the interpretation of qualitative date like student reflections at the scale of a classroom. NLP has been used in education research to help summarize student engagement online, evaluating open-ended test responses, and supporting course modifications [8]. An NLP application that has been used in education is sentiment analysis. The use of sentiment analysis can evaluate the emotional tone or polarity of a statement [9]. The polarity can be categorized as positive, neutral, or negative providing insights into emotions and attitudes for what students are reflecting on. Sentiment analysis has been used in one study to identify student emotions for assignments to help refine the assignment [10]. Another use of sentiment analysis for a design course involved tracking changes to student sentiments in reflections over time to gather insight into the evolution and development of the activities [11].

This study builds on the benefits and tools of reflection and sentiment analysis along with past work to systematically analyze the tone of student reflections in a flipped, mastery-based Statics course. The study will analyze how student tone changes over a semester and determine if that trend is related to the final grade the student earned in the course. The polarity of each student will be assigned as a score to determine trends in both emotional and cognitive attitude in the engineering course.

Course Context

Statics, a second-year engineering course, is taken by many engineering disciplines and serves as a foundational course for many subsequent theoretical engineering and design courses. It is one of the first technical courses the students take and integrates their math and science prerequisites with engineering applications. It is considered a challenging course in the engineering curriculum and due to its timing, often in the second year where expectations increase and attrition is common, makes it a pivotal part of their academic journey.

The content of Statics is well-established, but to foster student engagement and encourage deeper learning the course was redesigned. The redesign involved restructuring the course into a flipped classroom model utilizing a mastery grading system [12]. The flipped classroom shifts passive content delivery outside of class and allows in-class time to focus on active problem-solving and connecting with each student. The mastery grading system shifts the focus from earning grades to achieving mastery on the course mastery objectives. The course mastery objectives are the key

components to solve a Statics problem where each assessment requires the students to include work for 7-8 mastery objectives. To develop mastery in the subject, there are seven assessments given in a semester to track each student's progress with the course content, followed by detailed, individualized feedback of their progress. This student-centered approach helps students identify the areas that they are doing well on and areas that need improvement based on the key concepts of the course. The bi-weekly assessment allows for regular check-ins of each students progress along with the creation of manageable goals for the students to work towards every two weeks.

In addition to the course redesign, developing engagement opportunities and ways to understand student engagement has been a major focus in the course. To provide more opportunities for engagement, the course has integrated reflective practices through a Self-Assessment (SA) activity. The SA takes place every two weeks after each assessment where the students complete a reflection to evaluate their performance on the assessment compared to the instructor's solution and comment on their work for each of the 7-8 mastery objectives. The practice of doing the reflections encourage the students to develop self-awareness of their work in an environment that is safe and open ended for expression.

This study builds on prior research on the SA that included analyzing how students grade themselves compared to the instructor along with the initial sentiment categorization that was used for the reflections [7]. The sentiment analysis has been further refined to provide correlations of emotional tone with academic metrics such as final grades. This offers a generalized view of the relationship between student attitudes and course learning outcomes.

Data collection

The data for this study was collected over seven semesters of the Statics course taught by a single instructor with a total of 848 students. Participation in the SA was optional, but it was incentivized by providing participation points for completing it, resulting in most students completing all seven reflection opportunities each semester. In a single semester there was an average of 54 comments collected per student resulting in over 45,000 comments analyzed in this study. The prompts for the reflections were open-ended allowing students to express themselves emotionally about their performance or reflect cognitively on the specific concept or course material that was tested. This lead to a diverse range and variety of comment types creating a rich dataset for analysis.

Sentiment Analysis Framework

The use of sentiment analysis was key to analyzing the qualitative nature of the reflections and uses natural language processing (NLP) tools to assess the emotional polarity of the written work. The polarity was categorized on a five-point scale: negative, slight negative, neutral, slight positive, and positive. The reflections fell into two types of comments ranging from content-

specific reflections to emotional reflections. It was essential that the sentiment analysis could capture the polarity for both types of responses.

The development of the sentiment analysis framework involved testing and evaluating several pre-existing NLP models to find the one that was most accurate for the dataset [13]. After selecting the base model, it was fine-tuned using a rule book designed for this study. The rule book outlined the types of comments to include in each polarity category for both attitude and content-specific comments supported by student examples from the data. To validate the model, the research team manually labeled 1,500 comments and compared them to the program's outputs. The program was determined acceptable once it had an agreement rate over 75% and then was used on the entire dataset. The team moved forward with 75% agreement because the disagreements between the human label and computer label where often between slight positive and positive or slight negative and negative, so still within the positive or negative tone just not the same level was assigned by both.

Student Sentiment Scores

Using the refined model, each comment was assigned a polarity label. This resulted in 7-8 polarity labels for each of the mastery objectives per assessment per every student. To develop a holistic view of student sentiment, an algorithm was used to compute an overall sentiment score for each student per assessment. The overall sentiment score was computed by assigning numeric values to the polarity labels from -4 to +4 and then calculating the average sentiment score for each assessment. The polarity values were assigned ad -4:negative, -2:slight negative, 0:neutral, +2:slight positive, +4 positive. The average sentiment score for each assessment was related back to a polarity label to assign the student a single assessment sentiment based on their tone for that assessment. Then each student's overall average sentiment score was also computed by taking the average of their sentiment scores for each assessment. This results in one single sentiment score and label for each student capturing their average sentiment over the entire semester.

The overall sentiment score could be compared to final grade to determine the relationship along with tracking the single assessment sentiment score over the semester to determine trends in sentiment evolution. The single assessment and overall sentiment score were correlated with each student's final grade in Statics to determine if there is a relationship between emotional engagement and cognitive reflection with academic performance.

Results and Analysis

The analysis was completed for each semester in the study and then as a whole for all seven semesters included to show the relationship between overall sentiment score and final grade. The results are shown only for all the students combined and not broken down by each semester. There were subtle differences between each semester particularly comparing fall to spring semesters, but the general trends were constant so the cumulative results are shown in this paper. Figure 1 illustrates the average overall sentiment score for each final grade earned.

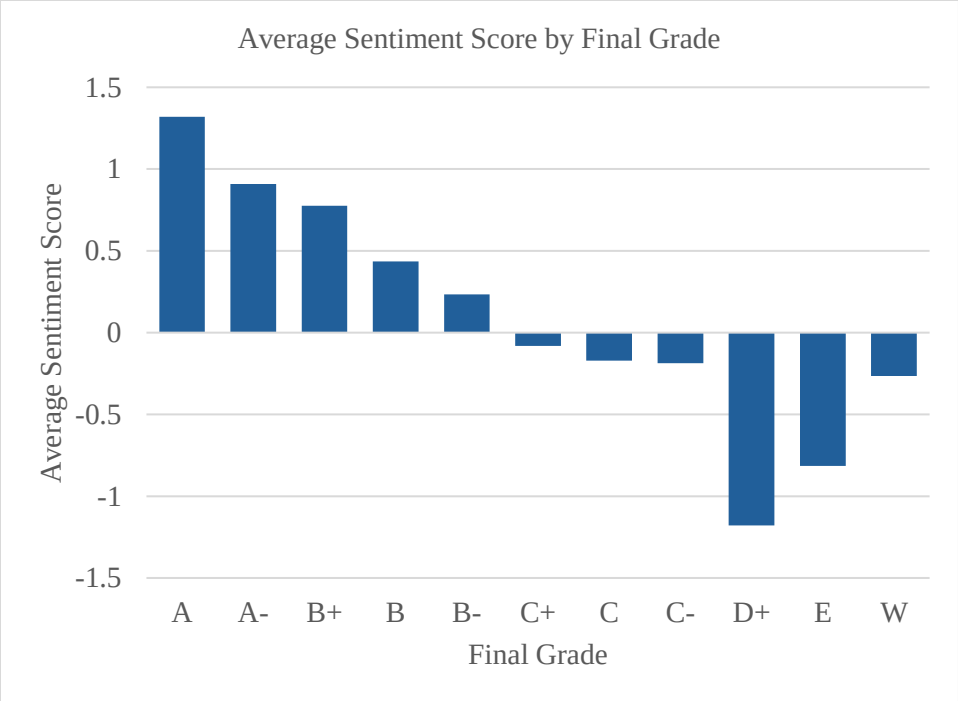

Figure 1: Average Sentiment Score by Final Grade Earned

There is a clear trend that students who earned high grades had higher sentiment scores meaning that students who earned A grades expressed the most positive sentiment in their reflections. These student's comments often included statements about their confidence with the material or getting the correct answer. There was a steady decrease in sentiment scores for students that earned A-, B+, B, and B- grades, which were all still on the positive side of neutral. Students that earned grades of C+ or lower had sentiment scores on the negative side of neutral with students that earned a grade of D+ having the most negative sentiment score. These students reflected more critically on their performance or expressed frustration about the problem or their understanding. These findings make sense intuitively that higher performing students would express more positive comments in their reflections. However, the results also emphasize the importance of addressing students that express negative sentiments early in the semester. These students often struggled academically, and early engagement and support based on their tone could improve their course outcome.

Further, the evolution of sentiment scores across the seven assessments provides deeper insights into how students engage with the material and their attitude towards the course. To explore these trends, students were divided into three groups where high-performing students are classified as those that earned an A- or A during the semester (A+ is not a grade given out at the university for this study), mid-performing students earned a B+, B, or B-, and low-performing students include all remaining grades. A student must earn a grade of C or better to move on from Statics, so C grades and below are considered low performers in this study when often C grades are considered average. Figure 2 presents the average sentiment score trend for each group over the seven assessments in a semester.

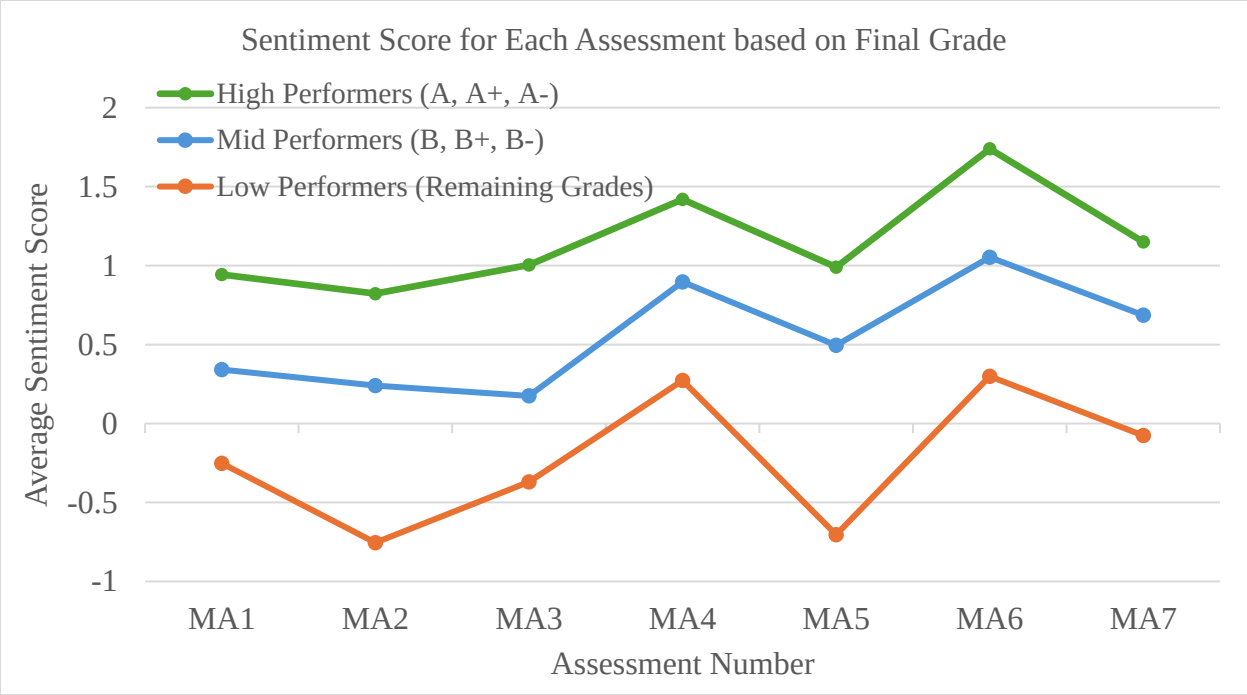

Figure 2: Sentiment Score Trend over Seven Assessments for High-, Mid-, Low-Performers

The high-performing students consistently had a positive sentiment across all assessments, this was also noted in the previous Figure 1. The lowest score for this group was for Assessment 2, while the highest score was for Assessment 6. The score fluctuations could be connected to the varying difficulty and timing of each assessment. Despite the small variations, the overall positive scores show the confidence of these students with their work and their constant engagement with the material. Approaching challenging topics with a positive sentiment is a skill that contributed to and will continue to contribute to their academic success.

Contrary to this, the low-performing students had sentiment scores in the slightly negative range for most assessments. There was some improvement towards the middle to end of the semester with the last few assessments except for a steep drop during Assessment 5. Assessment 2 and 5 had the most negative scores which was common for the three grade groups, but for the low performers the drop is much steeper for those assessments. These assessments are more challenging for students and highlight areas where more support could be valuable. Assessment 4, 6, and 7 had sentiment scores that reached into the slightly positive range. This trend reflects a combination of growth in understanding, increased familiarity with the material and course structure, and the result of consistent feedback and support throughout the semester. Unfortunately, these students seem to be in the positive area at the end resulting in lower final grades, but it is valuable to see how their attitude changes throughout a semester.

The positive sentiments for high-performing students reflects their attitude, engagement, and ability to positively work through challenging material. The mid-performing students were always positive as well but much closer to neutral. The low-performing students often had

negative sentiments which identify struggle and need for reinforcement to help with the material. There was still a positive ending for the low-performing students suggesting that mastery is still achievable for these students, but they need more time and support to reach it. This further emphasizes the importance for consistent feedback and support from the instructor with a possible need for targeted interventions during the points in the semester that had the most negative sentiments.

By identifying the students with consistently negative sentiments early in the semester along with the outcomes of each assessment, these students can receive additional support for the course material along with support to help build their confidence. These trends create the ability to better respond to each student's needs and create a more adaptive learning environment.

A final analysis was explored to evaluate how much student sentiment scores changed from the start of the semester (Assessment 1) to the end of the semester (Assessment 7). The change in sentiment score was calculated as the difference between the sentiment score assigned for Assessment 7 and that for Assessment 1. These individual score changes were then categorized based on the final grade earned. Figure 3 presents the distribution of sentiment score change for each final grade.

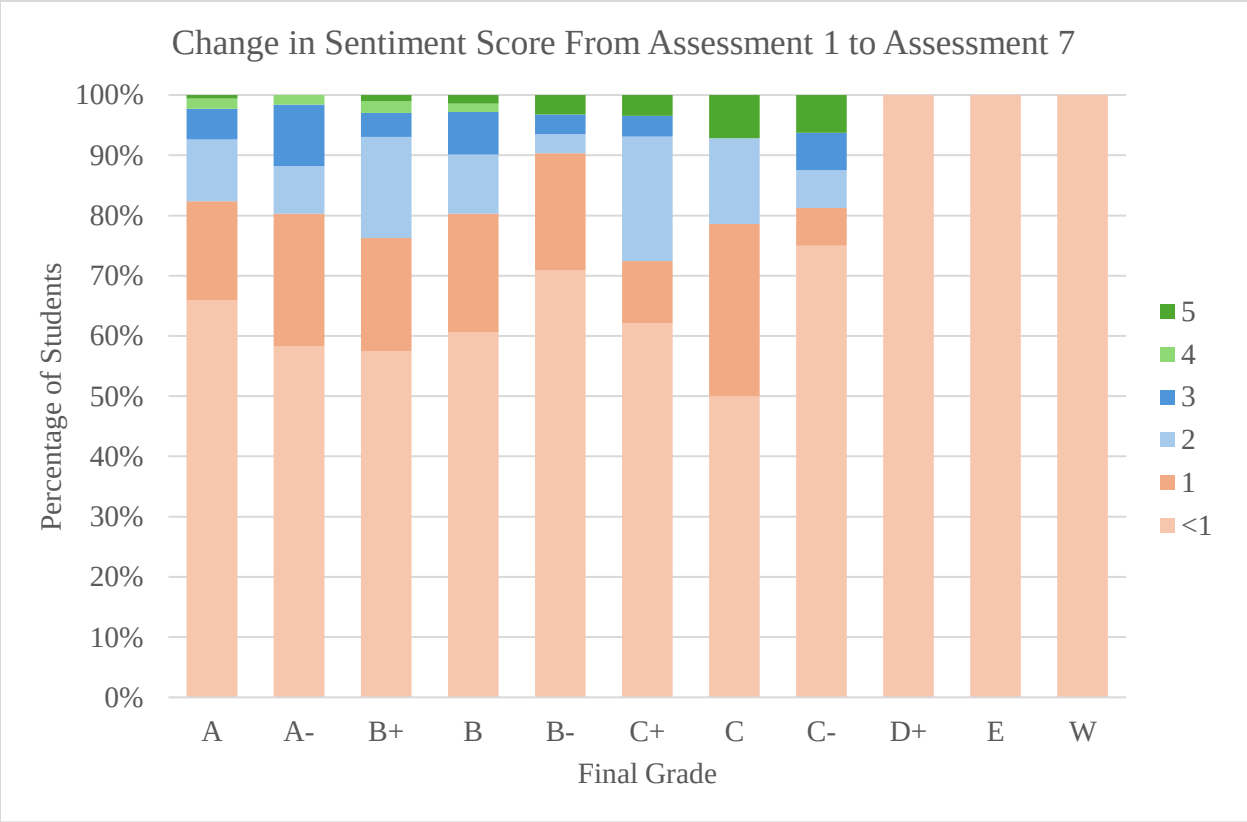

Figure 3: Change in Sentiment Score from Assessment 1 to Assessment 7 based on Final Grade

The analysis revealed that the majority of the students had minimal change to their sentiment score over the semester, particularly students who failed the course. This indicates that for most

students, their attitude or emotional tone for the reflections was consistent regardless of their academic performance. There was about 10% of students that earned an A- who had a sentiment score change of three or more points which corresponds to a shift across an entire sentiment category (e.g. from neutral to slight positive or from slight negative to neutral). This represents a group of students that notably had a significant shift in attitude over the semester.

These findings indicate that student sentiment scores remain consistent over a semester despite their academic performance. The high sentiment score changes that are seen in 10-20% of the students that earned higher grades identify opportunities to analyze how positive shifts in sentiment lead to improved engagement and academic performance. An additional analysis could determine if changes in sentiment score also correlate with academic improvements for the same assessments. An example would be if students that had a major change with their sentiment score also had a corresponding change with their assessment scores. These results can help identify and better support students that are in negative or stationary sentiment paths.

Discussion

This study shows a strong connection between student sentiment tone in their reflections and their academic success. From this connection, there are opportunities for improvements to teaching strategies and ways to engage the students. The use of sentiment analysis, by leveraging NLP tools, creates an accessible way to identify characteristics of reflections to better personalize student feedback and support.

The results have a clear relationship between sentiment tone and student final grades, where high-performing students had consistently positive tone. The positivity likely reflects their attitude to the material through identifying confidence in understanding and success with the problem. The students who received lower grades most often had negative sentiments, specifically for the more challenging assessments, that indicate disappointment or struggle with the course material or their understanding. The sentiment tone in this study represents both student cognitive engagement with the material and their attitude towards the course with the potential to impact their overall attitude and engagement with learning.

While the results give the overview of student tone, it still provides insight into how students interact with the course material, and patterns in this tone over time highlight moments of success and struggle. The assessments that saw a drop in sentiment are areas where students could use additional scaffolding, support, or most likely just more time to engage with the material. In real-time identifying student sentiment after each assessment will allow individual feedback for each student based on their tone within the context of the findings of this study. The students with positive sentiment can receive continual encouragement to continue their path while students with negative sentiment can be given targeted resources or interventions for the areas that they found challenging.

The interventions could vary but include one-on-one meetings with the instructor to discuss the student's challenges specific to each assessment but also in general for the course. It also creates opportunities to connect students with positive attitudes for the course with those that are more negative. There can be supplemental resources created to help focus on the challenging topics identified by topics or themes mentioned in the Self-Assessments. Finally, additional reflective practices that encourage progress and continual engagement and improvement can be encouraged to help reframe students with negative attitudes. This would involve adding mindset exercises into the course which will be pursued in a future project.

Currently, identifying student sentiment after each assessment provides opportunity for personalized feedback. In addition, to the detailed feedback the students receive on the technical components of their work they can be acknowledged for their reflective response. Students with positive sentiments can be encouraged to continue challenging themselves and commented on for their confidence. The students with a negative sentiment can be encouraged to communicate with the instructor to clarify the concepts and keep working through challenges. This will create a stronger connection with the student and a more supportive learning environment that incorporates affective qualities of each student's development.

Future Work

Further research will include identifying how sentiment tone can relate to student mindset. It is likely that a positive sentiment throughout a semester is indicative of a growth mindset which can lead to the ability to persevere through challenges and be successful [14]. While a negative sentiment could imply a fixed mindset that can result in frustration and disengagement leading to a less successful academic performance.

Conclusion

This study emphasizes the value of integrating reflective exercises into a technical engineering course with the use of NLP tools to support the interpretation of the reflections. Reflection is a key piece to encourage student metacognition to continue to engage and grow in their education. For a large classroom it is challenging though to use the reflections to their full potential due to information size, time constraints, and emphasis on covering the course content. The use of sentiment analysis has been a powerful tool to understand student attitude and confidence towards the course. The use of this NLP tool provides a scalable method for identifying these student characteristics. It will also inform more personalized feedback and adjustments to course resources. The students benefit from the regular reflection on their work, but to communicate how this reflection can shape their learning journey becomes critical for lifelong learning. The use of these qualitative insights with the quantitative academic metrics demonstrates how both can be used to support student outcomes. The integration of both results in creating a classroom environment that supports students both academically and emotionally using individualized data for each. The integration of technology to connect this information makes it possible in real time for large classrooms to enhance the overall learning experience for students.

References

[1] ABET, (2025), https://www.abet.org/accreditation/accreditation-criteria/criteria-for-accrediting-engineering-programs-2025-2026/

[2] S. A. Ambrose, "Undergraduate Engineering Curriculum: The Ultimate Design Challenge," in *The Bridge - Linking Engineering and Society*, vol. 43, no. 2, 2013, pp. 16–23.

[3] D. A. Schön, (1983). *The Reflective Practitioner: How Professionals Think in Action*. Basic Books.

[4] J. A. Turns, E. Cuddihy, and Z. Guan, "I thought this was going to be a waste of time: How portfolio construction can support student learning from project-based experiences." *Interdisciplinary Journal of Problem-Based Learning*, vol. 4, no. 2, 2010, pp. 63-93.

[5] R. Adams, J. Turns, and C. Atman, "Educating effective engineering designers: The role of reflective practice." *Design Studies*, vol. 24, no. 3, 2003, pp. 275-294.

[6] J. A. Turns, B. Sattler, K. Yasuhara, J. L. Borgford-Parnell, and C. J. Atman, "Integrating reflection into engineering education," *ASEE Annu. Conf. Expo. Conf. Proc.*, 2014.

[7] A. Baisley, and K. D. Hjelmstad, "The accuracy of student self-assessment in engineering mechanics", *Paper presented at 2022 ASEE Annual Conference & Exposition*, 2022, https://peer.asee.org/40837

[8] O. Zawacki-Richter, V. I. Marín, M. Bond, and F. Gouverneur, "Systematic review of research on artificial intelligence applications in higher education – Where are the educators?" *International Journal of Educational Technology in Higher Education*, vol. 16, no. 1, 2019.

[9] Y. Sun, T. Cui, S. Yuan, and X. Xiao, "Sentiment analysis in education research: A systematic review and future directions." *British Journal of Educational Technology*, vol. 52 no. 4, 2021, pp. 1432-1451.

[10] Y. Wu, Z. Ming, J. K. Allen, and F. Mistree, "Evaluation of Students' Learning Through Reflection on Doing Based on Sentiment Analysis," *J. Mech. Des.*, vol. 145, no. 3, pp. 1–12, 2023.

[11] J. Frich, N. Inie, K. Halskov, and P. Dalsgarrd, "A Sentiment Analysis of Design Reflections from Design Projects," *Conf. Proceedings of the 2018 CHI Conference on Human Factors in Computing System*, 2018, doi: 10.1145/3170427.3188444

[12] K.D. Hjelmstad, and A. Baisley, "The Mechanics Project: A Pedagogy of Engagement for Undergraduate Mechanics Courses," *Paper presented at 2020 ASEE Virtual Annual Conference,* 2020, doi: 10.18260/1-2—35346

[13] A. Baisley, and C. S. Marutla, "Use of Sentiment Analysis to Assess Student Reflections in Statics" *Paper presented at 2024 ASEE Annual Conference & Exposition*, Portland, Oregon, 2024, doi: 10.18260/1-2--48218

[14] C. Dweck, "Mindset: The New Psychology of Success" Ballantine Books, 2016.