



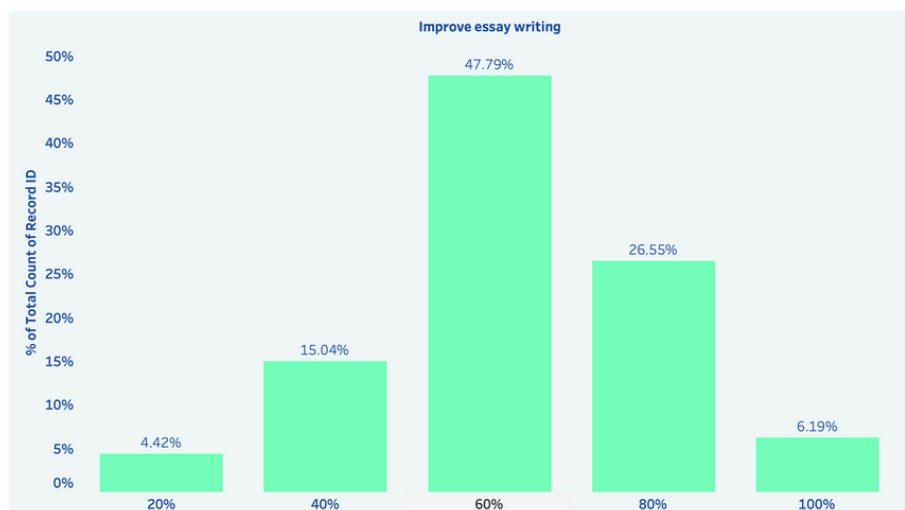
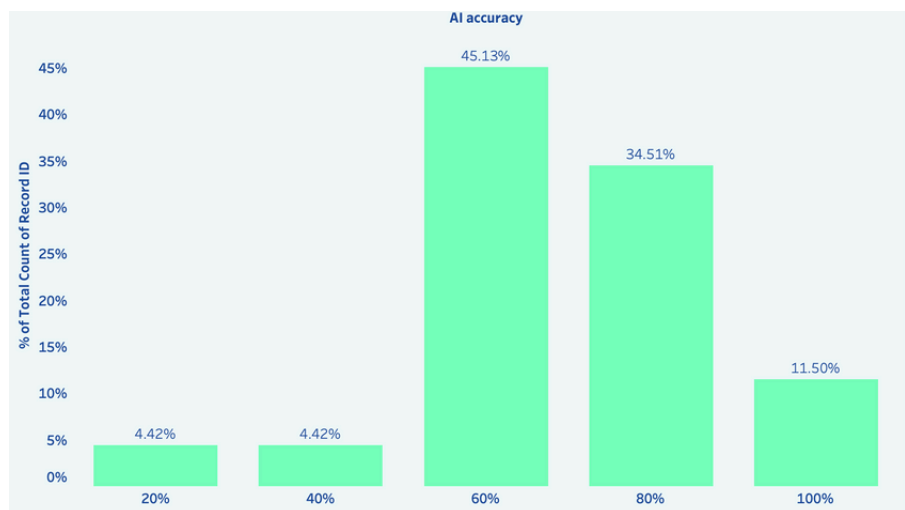
## Transformative Learning: AI's Role in Nurturing Thoughtful Citizens through Disciplinary Skills

In January of 2020, Thinking Nation officially launched as a 501c3 with a mission to cultivate thinking citizens. Our goal is to shift the paradigm of social studies education. To do so, we realized we needed to scaffold the process for students to begin to write essays that reflect their ability to analyze and think critically. We also knew that we needed to support teachers through professional development, and to alleviate the time spent grading essays, while minimizing bias. We created standardized rubrics for the assessment of essays and hired part-time staff (teachers) to grade our Curated Research Papers (CRPs). In Summer 2023, Thinking Nation began work to incorporate AI into our grading process. In December of 2023, we launched our AI Grading. The results of this switch to AI demonstrate the benefits to the students from more timely and accurate feedback. AI grading norms unbiased results and provides teachers with immediate/accurate data, which they can utilize in honing their lesson plans. The following data is from a survey conducted to which over 200 students responded.

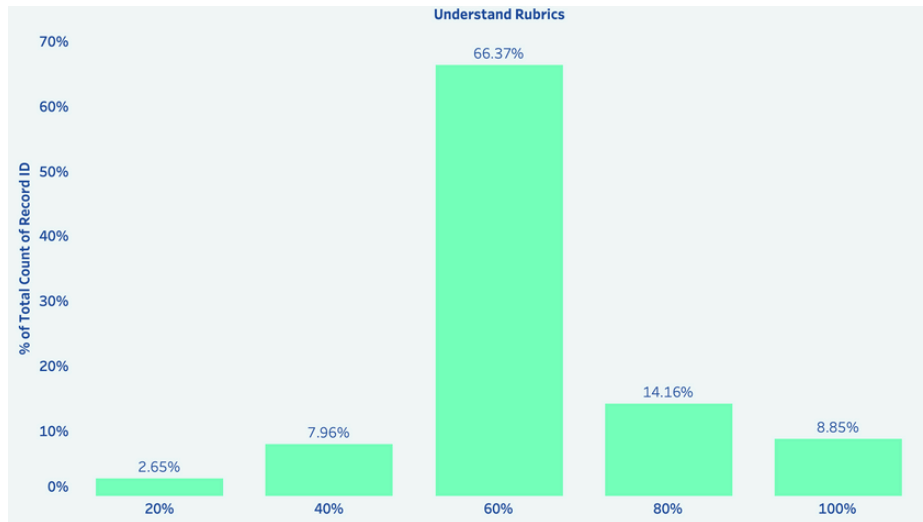
### Student Reflections

Most students perceived Thinking Nation's AI grading system to be consistently more accurate than human graders, as it typically shows at least a **4%** improvement across all accuracy levels, with approximately **11%** of students noting a significant increase in accuracy, particularly at the **100% level**.

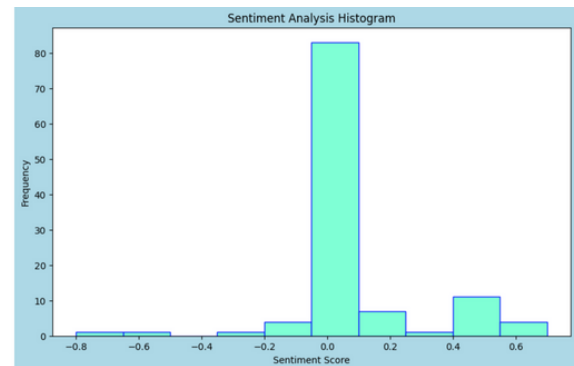
Student respondents indicated that they believe AI-based grading has significantly helped to improve their essay writing skills, with a weighted average improvement of approximately **57.39%** across all reported cases."



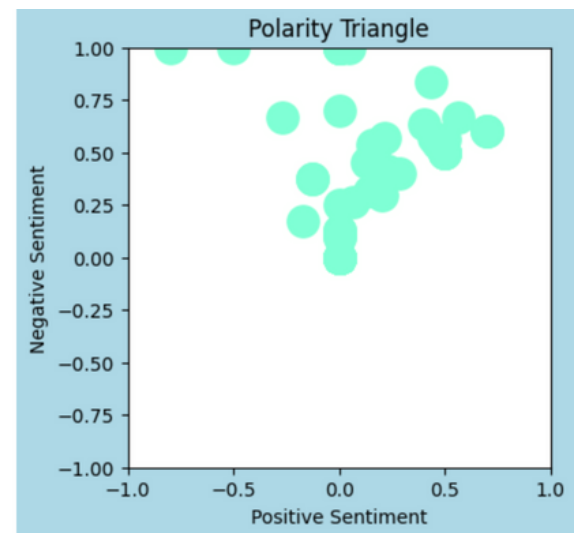
Analysis of student feedback indicates that our AI-based grading system has significantly enhanced students' understanding of rubrics, with a weighted average improvement of approximately **63.72%** across all reported cases.



In addition to survey question asking respondents to answer using a Likert Scale, we asked an open-ended question: **“How do you think AI should be used in school now, and in the future”** Examining the polarity triangle and histogram reveals a prevailing sentiment among students, largely indicating a mixture of neutral and positive attitudes towards the integration of AI in educational environments. This observation underscores a general acceptance or curiosity among students regarding the potential benefits and implications of AI technology within the realm of schooling.



Among our surveyed students, a notable **80%** acknowledged that AI grading significantly contributed to a substantial **70%** enhancement in their essay writing skills, indicating their belief that this advancement can effectively aid in their journey toward becoming better writers.

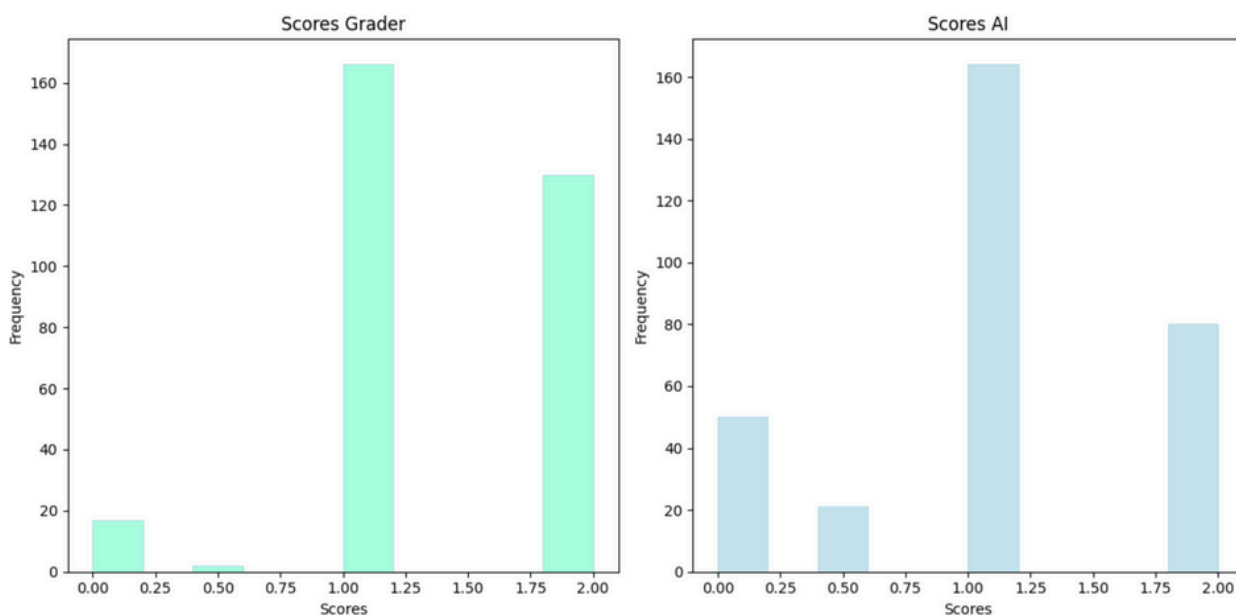


# Data and Teacher Feedback



While we did survey teachers for their thoughts on AI grading, the population of teachers who have had students submit CRPs totals 22, not a significant enough pool for the data we plan to gather on an ongoing basis. We did however notice some initial trends in their responses, and have looked at data from previously human-graded CRPs to confirm if the scoring is more accurate and thorough.

- **Consistency:** The distribution in the "Scores AI" histogram appears more concentrated around the peak score, suggesting that the AI system may provide more consistent grading outcomes this is beneficial as it ensures fairness and reliability.
- **Balanced Distribution:** The "Scores AI" histogram shows frequencies distributed more evenly across the score range, implying that the AI system evaluates assignments comprehensively and thoroughly across the entire spectrum of scores.
- **Potential for Error Reduction:** Absence of scores in certain bins in the "Scores AI" histogram suggests that the AI system may be less susceptible to grading errors or biases, potentially leading to more accurate assessment outcomes and reduced grading inconsistencies.



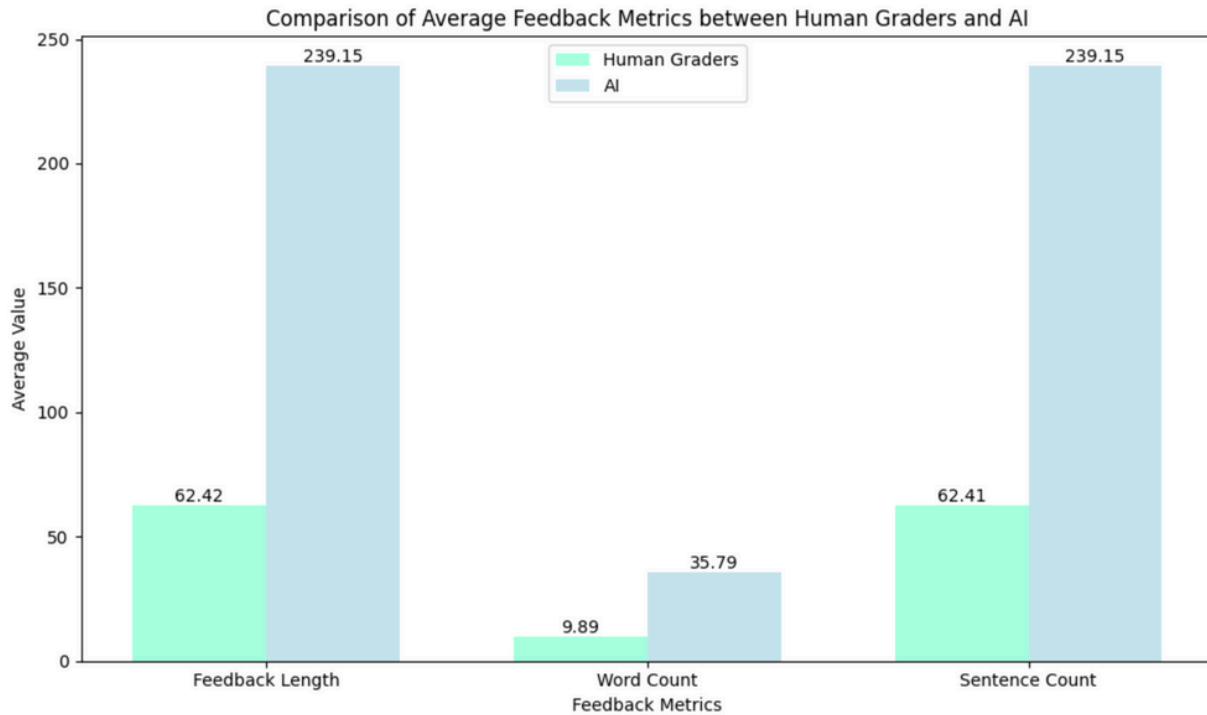
75% of the interviewed teachers expressed that the feedback provided by AI is approximately 80% accurate, and they also believe it to be 77% more accurate than feedback from humans.

88% of teachers surveyed said that they have been able to utilize and act upon the feedback given by AI 90% of the time.

# Comparison of Feedback Metrics



The detailed feedback offered by AI, with longer explanations and deeper analyses, holds promise for enhancing students' understanding of assignments. This comprehensive feedback provides clearer insights into strengths and areas for improvement, fostering deeper comprehension and proficiency in the subject matter.



## Thesis Feedback

### Grader

"Agriculture shaped early civilization by allowing people to get a surplus of food, have permanent houses, and create in their free time. However, on the negative side, it can also lead to diseases"

### AI

"The essay provides a clear thesis in the introduction that aligns with the prompt by claiming that agriculture allowed early civilizations to have a surplus of food, live in permanent houses, create during their free time, and that it had negative effects such as the spread of diseases, which can be defended with textual evidence throughout the essay."

## Outside Evidence Feedback

### Grader

None used

### AI

"The essay does not incorporate any outside evidence to strengthen its argument. The student only uses information provided in the textual evidence section and their prior knowledge of the topic."