



Human vs. AI Feedback on EFL Students' Essays: A Comparative Study

Yasemin Esma Şahin^{a 1*}, Fatma Gürman Kahraman^b

^a Bursa Uludag University, MA Programme Türkiye

^b Bursa Uludag University, School of Foreign Languages, Türkiye

Suggested citation: Sahin, Y. E. & Kahraman, F. G. (2025). Human vs. AI Feedback on EFL Students' Essays: A Comparative Study. *Language Education and Technology*, 5(1), 71-87.

Article Info

Date submitted: 25/06/2025

Date accepted: 25/07/2025

Date published: 28/07/2025

Abstract

This study compares the effectiveness of human and AI-generated feedback on improving EFL student essays. Twenty students from a state university in Türkiye participated, producing 19 argumentative and 20 opinion essays. Initial scores ranged from 2.00 to 2.15. Essays were scored with a four-band analytic rubric adapted from Reid (1993) and expanded with grammar-error categories from Schenck & Daly (2012). Human feedback significantly improved essay quality by addressing grammar, structure, and content with personalized comments, encouraging deeper engagement. In contrast, AI feedback focused mainly on grammatical corrections and clarity, lacking comprehensive and personal touch. The findings suggest a hybrid approach, utilizing AI for basic corrections and human instructors for detailed feedback, which could optimize student learning outcomes.

Research Article

Keywords: AI feedback, educational assessment, essay improvement, feedback effectiveness, human feedback

Introduction

Recently, AI in science has taken another significant leap forward with the release of the large language model software ChatGPT, a publicly accessible tool developed by OpenAI, a research center founded in 2015 (Hill-Yardin, Hutchinson, Laycock, & Spencer, 2023). This tool can engage in conversation with users in a seemingly natural and intuitive way, bridging gaps between human-like interaction and machine responses (Rudolph et al., 2023). This revolutionary advancement has enabled billions of individuals to access an expansive, open-source repository of information, fostering intellectual engagement and satisfaction on a global scale.

As the capabilities of ChatGPT emerged, the world was grappling with the aftermath of COVID-19, seeking new ways to adapt and recover. The introduction of ChatGPT marked a significant technological milestone,

^{1*}Yasemin Esma Sahin. English Language Teaching, Bursa Uludag University, MA Programme, Türkiye.
e-mail adress: y.caner19@gmail.com

capturing the attention of diverse sectors, including education. However, this innovation has sparked a dichotomy within the educational community.

The swift integration of ChatGPT into various domains ignited a debate on the effectiveness of AI feedback versus human feedback, especially given its rapid response capability. Feedback is a crucial component of the educational process at every level because it informs students about how to improve their performance by manipulating their current knowledge and proactively changing their consequent behavior (Boud & Molloy, 2013; Hounsell, 2007). The rapid proliferation of AI in the educational field has raised significant questions about the future of language learning and teaching, affecting millions of students, parents, teachers, principals, and educational field workers. A central concern has been whether AI can provide feedback that genuinely fosters the learning process.

One of the primary challenges noted is the significant time and effort required to provide effective feedback, especially to numerous students across various classes, which can be overwhelming for educators and may even discourage some from offering the essential feedback needed to support student development (Applebee & Langer, 2011; Kiuahara, Graham, & Hawken, 2009; Graham, 2019). This emerging debate has created an urgent need to understand the advantages and disadvantages of AI-generated feedback and how it can best be utilized to complement or replace traditional human feedback.

This study aims to bridge this gap by providing a detailed comparison of AI-generated and human-provided feedback on student essays in an English as a Foreign Language (EFL) setting by assessing the technical accuracy of AI in identifying grammar and punctuation errors. By doing so, the study will offer insights into the potential of AI to supplement or even enhance traditional teaching methods, thereby informing future educational strategies and technology integration in language learning.

Research Questions

1. How does the effectiveness of AI-generated feedback compare to human-provided feedback in improving grammar, punctuation, and structure in EFL students' essays?
2. Does the effectiveness of feedback (AI vs. human) vary across different types of essays, such as opinion or argumentative?

Literature Review

AI in Education and Feedback Mechanisms

Artificial Intelligence (AI) has seen transformative advancements across various fields, including education. The continuous evolution of educational technologies has led to increased interest and debate over AI's potential to replace human language teachers. AI-powered tools, such as intelligent tutoring systems, language learning apps, and conversational agents, offer personalized learning experiences and adaptive feedback to learners (Nushi et al., 2020). These applications utilize natural language processing (NLP) and machine learning algorithms to foster engaging and interactive environments for language learners.

One of the key areas where AI has made significant contributions is in the generation of feedback. AI systems can efficiently handle large volumes of data and provide instant feedback, which is particularly

beneficial in educational settings with high student-to-teacher ratios. AI can generate detailed feedback on grammar, punctuation, and structure, helping students make immediate improvements to their writing. Bulut et al. (2022) highlighted that AI-driven feedback can offer personalized learning experiences tailored to individual learning styles and needs, which can be more difficult for human teachers to achieve consistently.

AI-powered language learning platforms like Duolingo, Babbel, and Rosetta Stone have become popular due to their capacity to provide tailored learning pathways and instant feedback to users. These applications utilize NLP and machine learning algorithms to create interactive and engaging environments for language learners (Lu et al., 2021). AI-driven virtual tutors and chatbots facilitate real-time conversational practice, mimicking human interactions to aid learners in improving their language skills.

The efficiency of AI in processing large datasets and providing consistent feedback is a significant advantage. AI can quickly identify patterns and common errors, allowing for more targeted and effective feedback (Wongvorachan & Bulut, 2022). Moreover, AI can help manage administrative tasks, freeing human teachers to focus on more complex educational activities.

Providing high-quality feedback requires substantial time and effort from educators. Teachers often face the challenge of delivering detailed, individualized feedback to a large number of students, which can be overwhelming and discouraging (Applebee & Langer, 2011; Graham, 2019). This workload can impact their ability to offer the essential support needed to foster student development. Ferguson (2011) found that students value detailed and timely feedback, but teachers often struggle to meet these expectations due to their heavy workloads. The study emphasized that high-quality feedback is essential for student improvement, yet delivering such feedback is time-consuming and labor-intensive for educators.

The integration of AI systems can alleviate this burden by automating the feedback process, allowing teachers to allocate their time and energy more effectively (Debus & Lawley, 2016). AI can handle routine and repetitive tasks, providing immediate feedback on assignments and enabling teachers to focus on more personalized and in-depth interactions with students (Bijlsma et al., 2019). This not only enhances the efficiency of the feedback process but also ensures that students receive the necessary guidance to improve their performance.

Several studies have compared the quality of feedback provided by AI versus human instructors. One study found that while AI can effectively identify and correct grammatical errors, it often struggles with nuanced language issues that require a deeper understanding of context and intent (Perkins et al., 2022). Another study highlighted that students generally preferred human feedback for its perceived empathy and personal touch, which they felt were lacking in AI-generated feedback (Wongvorachan & Bulut, 2022). However, AI feedback was still considered highly useful and nearly as effective as human feedback in many cases (Bulut et al., 2022).

The integration of AI into educational settings raises several ethical and logistical challenges. Concerns such as data privacy, potential biases in AI algorithms, and the effects of diminished human interaction on students' social development are crucial considerations (Yang, 2024). Ensuring that AI systems are transparent and accountable is essential to maintaining trust and efficacy in educational environments (Kostka & Toncelli, 2023). Additionally, the deployment of AI requires significant infrastructure and ongoing maintenance, which can be resource-intensive (Shi & Zhang, 2023).

The literature suggests that while AI has the potential to significantly enhance language education by providing personalized and adaptive learning experiences, the unique human elements of emotional intelligence, cultural understanding, and personalized mentorship remain irreplaceable. The future of language education should focus on integrating AI to support and augment human teaching, ensuring a balanced and holistic approach to student development. AI feedback, being nearly as effective as human feedback, can serve as a valuable tool to complement the efforts of educators, particularly in managing time and providing consistent, high-quality feedback.

Methodology

Research Design

This study followed a pre-test/post-test within-subjects design to compare the effectiveness of human and AI feedback on EFL learners' writing. Each student produced two drafts of the same essay type (argumentative or opinion): an initial draft (baseline) and a revised draft after feedback. The independent variable was the feedback source (instructor vs. ChatGPT-4.0); the dependent variable was the change in Grammar and Spelling Analysis Rubric (GSAR) scores.

Participants/Sampling

Participants were 20 preparatory-school students at a state university in Türkiye. All participants were B2 level according to the CEFR. They submitted 19 argumentative and 20 opinion essays during the semester. A convenience sample was used because the intact class volunteered to participate.

Instruments/Materials

Grammar and Spelling Analysis Rubric (GSAR)

The GSAR was adapted from Reid (1993). It retains Reid's five macro-criteria—Content, Organization, Diction, Language Use, and Mechanics—each scored on a four-band scale (Poor = 1 to Excellent = 4). Language Use was subdivided into eight grammar-error categories (sentence structure, verb tense and form, pronoun usage, prepositions, article use, modifiers, run-on sentences, sentence fragments) derived from Schenck and Daly (2012). Mechanics was expanded into five spelling-related categories (spelling accuracy, consistency, homophones, compound words, syllable division) developed by the researchers for the present study. The full rubric appears in Appendix A.

AI feedback tool

ChatGPT-4.0 (OpenAI, 2024) generated automated comments; system parameters and version details are listed in Appendix B.

Procedure

The study began with the initial assessment of the writing samples using the GSAR to establish baseline data. Following this, feedback was provided to the students, with each essay receiving detailed feedback from both AI and the students' human instructor.

Feedback Process

Each essay received feedback from both a human instructor and an AI system (ChatGPT 4.0). The human instructor provided detailed and personalized feedback, addressing various aspects of writing quality,

including grammar, structure, and grammar and mechanics. AI feedback was generated using a specific prompt designed to elicit comprehensive feedback on the essays.

AI Feedback Prompt

The prompt used to obtain AI feedback was as follows:

You're an English teacher, and this is one of your student's essays. You'll be assessing the essay in terms of content, organization, grammar, and vocabulary. The students' level is B1. Feedback should be between 100-150 words. The tone should be formal yet motivating. Address to the student. Identify the mistakes. Give back the student's essay but make the necessary changes like a human teacher would. If you encounter any mistake, italicize the incorrect word/phrase and type the appropriate correction code right after the incorrect word in parentheses. For example; refards (regards)

Data Collection and Analysis

The data analysis process in this study involved a solely quantitative approach to evaluate the effectiveness of human and AI feedback on student writing. The GSAR was employed to assess the writing samples before and after feedback was provided.

Initial Assessment

The initial assessment of the writing samples was conducted using the GSAR to establish a baseline score. This provided a benchmark for comparing the improvements after receiving feedback.

Post-Feedback Assessment

After the feedback was incorporated into the writing samples, a post-feedback assessment was conducted using the same rubric. This step involved re-evaluating each writing sample using the same criteria to determine the extent of improvement in grammar and spelling.

Results and Discussion

Human Feedback

Research supports the effectiveness of human feedback in educational settings. Shute (2008) emphasizes that effective feedback is specific, timely, and actionable, which aligns with the comprehensive feedback provided by instructors in this study. Human comments targeting verb-tense and sentence-fragment errors corresponded to a +1.9-point gain (see Appendix C). For instance, in Arg_Essay1_Student2, the human instructor's feedback improved the essay's structure and clarity by addressing verb tense errors, sentence fragments, and spelling issues (Shute, 2008).

AI Correction

AI feedback effectively identified and corrected basic grammatical errors, such as spelling mistakes, verb tense issues, and sentence fragments.

Several studies have explored the benefits and limitations of AI feedback in education. Perikos and Hatzilygeroudis (2017) examined how AI systems provide immediate and consistent feedback within intelligent tutoring environments, noting their effectiveness in correcting grammatical forms. In our analysis, AI feedback effectively corrected grammatical errors but did not match the depth of human feedback in structural improvements and personalized guidance.

Comparative Analysis

Argumentative Essays

The comparative analysis revealed that while both human and AI feedback improved the essays, the extent and nature of the improvements varied. Human feedback led to more significant enhancements in sentence structure, clarity, and overall coherence due to its detailed and personalized nature.

In contrast, AI feedback primarily focused on grammatical corrections and sentence clarity. The post-feedback scores for essays that received AI feedback were generally lower compared to those that received human feedback.

Table 1.

19 Argumentative Essays and Their Respective Scores

Essay ID	Student ID	Baseline Score	Human Feedback Score	AI Feedback Score
Arg_Essay1_Student2	S2	2,00	3,90	2,90
Arg_Essay2_Student3	S3	2,00	4,00	3,00
Arg_Essay3_Student4	S4	2,00	3,80	3,00
Arg_Essay4_Student5	S5	2,15	3,90	3,00
Arg_Essay5_Student7	S7	2,15	3,80	2,90
Arg_Essay6_Student8	S8	2,08	4,00	2,90
Arg_Essay7_Student9	S9	2,08	3,80	2,90
Arg_Essay8_Student10	S10	2,08	4,00	3,00
Arg_Essay9_Student11	S11	2,08	3,90	3,00
Arg_Essay10_Student12	S12	2,08	3,90	2,80
Arg_Essay11_Student14	S14	2,08	4,00	3,00
Arg_Essay12_Student15	S15	2,08	3,90	2,80
Arg_Essay13_Student16	S16	2,15	3,90	3,00
Arg_Essay14_Student18	S18	2,15	3,80	2,90
Arg_Essay15_Student19	S19	2,15	4,00	2,90

The average baseline score for the argumentative essays was 2.09, indicating a need for significant improvement in the initial drafts. Following the provision of feedback, the average scores increased $\approx 3,91$ for human feedback and $\approx 2,93$ for AI feedback.

Opinion Essays

The results for the opinion essays were similar to those for the argumentative essays. Human feedback again proved to be more effective in improving the overall quality of the essays. The detailed and motivational nature of human feedback encouraged students to engage more deeply with the revision process, leading to higher post-feedback scores.

Table 2.

20 Opinion Essays and Their Respective Scores

Essay ID	Student ID	Baseline Score	Human Feedback Score	AI Feedback Score
Op_Essay1_Student1	S1	2,15	3,90	2,90
Op_Essay2_Student2	S2	2,10	4,00	3,00
Op_Essay3_Student3	S3	2,00	4,00	3,00
Op_Essay4_Student4	S4	2,15	3,80	2,80
Op_Essay5_Student5	S5	2,15	3,90	3,00
Op_Essay6_Student6	S6	2,00	4,00	2,80
Op_Essay7_Student7	S7	2,15	4,00	2,90
Op_Essay8_Student8	S8	2,00	4,00	3,00
Op_Essay9_Student9	S9	2,15	3,90	2,90
Op_Essay10_Student10	S10	2,10	3,90	3,10
Op_Essay11_Student11	S11	2,10	3,90	3,00
Op_Essay12_Student12	S12	2,10	3,90	3,00
Op_Essay13_Student13	S13	2,15	4,00	3,00
Op_Essay14_Student14	S14	2,00	4,00	3,10
Op_Essay15_Student15	S15	2,15	3,80	2,90
Op_Essay16_Student16	S16	2,00	4,00	3,00
Op_Essay17_Student17	S17	2,15	4,00	2,90
Op_Essay18_Student18	S18	2,15	4,00	3,10
Op_Essay19_Student19	S19	2,10	3,90	3,00
Op_Essay20_Student20	S20	2,15	4,00	3,00

AI feedback for opinion essays also focused primarily on grammatical corrections and making sentences more concise. While useful for initial corrections, it did not provide the comprehensive support needed for substantial improvements in the essays. As a result, the post-feedback scores for essays that received AI feedback were generally lower compared to those that received human feedback.

The average baseline score for the 20 opinion essays was 2.10, indicating a need for significant improvement in the initial drafts. Following the provision of feedback, the average scores increased to $\approx 3,95$ for human feedback and $\approx 2,97$ for AI feedback.

The findings of this study underscore the superiority of human feedback over AI feedback in improving the quality of student essays. Human feedback is inherently more comprehensive and personalized, addressing a broader range of writing issues, including grammar, structure, and content. Human feedback's detailed and motivational nature plays a crucial role in encouraging students to engage deeply with the revision process and make substantial improvements to their essays. This is supported by literature emphasizing the importance of detailed and context-specific feedback in fostering student learning and improvement (Bulut et al., 2022).

AI feedback, while beneficial for initial grammatical corrections and sentence clarity, needs to provide more depth and breadth of feedback necessary for significant writing improvement. The primary limitation of AI feedback is its lack of personalization and context-specific suggestions, which are essential for helping students understand and rectify their writing weaknesses (Alnajashi, 2023). Moreover, AI systems like ChatGPT, although capable of processing large volumes of text quickly and accurately, still need to match the nuanced understanding and adaptive feedback capabilities of human instructors.

The study's findings align with previous research that highlights the potential of AI to support but not replace human instructors in the feedback process. AI can be a valuable tool for initial error correction and providing consistent, objective feedback on specific aspects of writing (e.g., grammar and spelling) (Wongvorachan & Bulut, 2022). However, the role of human instructors remains crucial for offering comprehensive, personalized feedback that addresses higher-order writing skills such as argumentation, coherence, and overall essay structure (Dai et al., 2023).

The findings underscore that the effectiveness of feedback (AI vs. human) varies across different types of essays, such as opinion or argumentative. Both types of feedback improved essay quality, but human feedback was markedly more effective. For argumentative essays, human feedback led to significant enhancements in sentence structure, clarity, and overall coherence, raising the average scores from 2.09 to ≈ 3.91 compared to ≈ 2.93 for AI feedback.

Similarly, for opinion essays, human feedback improved the scores from 2.10 to ≈ 3.95 , while AI feedback increased the scores to ≈ 2.97 . These results suggest that while AI feedback is beneficial for initial corrections, human feedback provides the comprehensive support necessary for significant writing improvement.

Conclusion

This study demonstrates that human feedback is more effective than AI feedback in improving the quality of student essays. Human feedback provides comprehensive support that addresses multiple aspects of writing, including grammar, structure, and content. In contrast, AI feedback, while useful for initial corrections, lacks the depth and personalization required for significant improvements in student writing. By combining both types of feedback, educators can optimize the feedback process, ensuring that students receive timely and comprehensive support while reducing the workload on teachers. Future research should explore the integration of AI feedback systems in educational practice and their long-term impacts on student learning.

Limitations and Suggestions for Further Research

First, the study relied solely on quantitative rubric scores; the researchers were unable to collect students' qualitative reflections on the two feedback sources, so learner perceptions and preferences remain unexamined. Second, the sample was limited to 20 upper-intermediate students from a single Turkish university, which restricts generalisability to other proficiency levels or educational contexts. Third, the intervention covered a single semester and two essay genres; future work should incorporate long-term follow-up and additional writing tasks to determine the durability and breadth of the observed gains.

References

- Applebee, A. N., & Langer, J. A. (2011). "EJ" extra: A snapshot of writing instruction in middle schools and high schools. *English Journal*, 100(6), 14–27.
- Bijlsma, H. J. E., Visscher, A., Dobbelaer, M., & Veldkamp, B. (2019). Does smartphone-assisted student feedback affect teachers' teaching quality? *Technology, Pedagogy and Education*, 28(2), 217–236. <https://doi.org/10.1080/1475939X.2019.1572534>
- Boud, D., & Molloy, E. (2013). Rethinking models of feedback for learning: the challenge of design. *Assessment & Evaluation in Higher Education*, 38(6), 698–712. <https://doi.org/10.1080/02602938.2012.691462>
- Bulut, O., Cutumisu, M., Aquilina, A. M., & Singh, D. (2022). Feedback Generation Through Artificial Intelligence. *Open/Technology in Education, Society, and Scholarship Association Conference Proceedings*, 2(1), 1-9. <https://doi.org/10.18357/otessac.2022.2.1.125>
- Bulut, O., & Wongvorachan, T. (2022). Feedback generation through artificial intelligence. *Open TESOL*. <https://doi.org/10.18357/otessac.2022.2.1.125>
- Dai, W., Lin, J., Jin, F., Li, T., Tsai, Y., Gasevic, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. *OSF Preprints*. <https://doi.org/10.35542/osf.io/hcgzj>
- Debus, J., & Lawley, M. (2016). Benefits and drawbacks of computer-based assessment and feedback systems: Student and educator perspectives. *British Journal of Educational Technology*, 47(2), 294–301. <https://doi.org/10.1111/bjet.12232>
- Ferguson, P. (2011). Student perceptions of quality feedback in teacher education. *Assessment & Evaluation in Higher Education*, 36(1), 51–62. <https://doi.org/10.1080/02602930903197883>
- Graham, S. (2019). Changing How Writing Is Taught. *Review of Research in Education*, 43(1), 277–303. <https://doi.org/10.3102/0091732X18821125>
- Hill-Yardin, E. L., Hutchinson, M. R., Laycock, R., & Spencer, S. J. (2023). A Chat(GPT) about the future of scientific publishing. *Brain, Behavior, and Immunity*, 110, 152–154. <https://doi.org/10.1016/j.bbi.2023.02.022>
- Hollands, F. M., & Tirthali, D. (2020). *MOOCs in higher education: Institutional goals and paths forward*. Palgrave Macmillan.
- Kostka, I., & Toncelli, R. (2023). Exploring Applications of ChatGPT to English Language Teaching: Opportunities, Challenges, and Recommendations. *TESL-EJ*, 27(3).
- Lu, L., Meng, X., Mao, Z., & Karniadakis, G. E. (2021). DeepXDE: A Deep Learning Library for Solving Differential Equations. *SIAM Review*, 63(1), 208–228. <https://doi.org/10.1137/19M1274067>
- Nushi, M., & Orouji, F. (2020). Investigating EFL Teachers' Views on Listening Difficulties Among Their Learners: The Case of Iranian Context. *Sage Open*, 10(2). <https://doi.org/10.1177/2158244020917393>

- Perikos, I., & Hatzilygeroudis, I. (2017). Examining the efficiency of feedback types in a virtual reality educational environment for learning search algorithms. In C. Frasson & G. Kostopoulos (Eds.), **Brain Function Assessment in Learning** (LNCS vol. 10512, pp. 169–175). Springer.
- Perkins, J. et al. (2022). Comparing the quality of human and ChatGPT feedback. *Journal of Applied Learning and Teaching*, 6(1).
- Reid, J. M. (1993). *Teaching ESL writing*. London, UK: Prentice Hall Englewood Cliffs.
- Rudolph, J., Tan, S., & Tan, S. (2023). ChatGPT: bullshit spewer or the end of traditional assessments in higher education? *Journal of Applied Learning and Teaching*, 6(1).
- Schenck, A., & Daly, E. (2012). Building a better mousetrap: Replacing subjective writing rubrics with more empirically-sound alternatives for EFL learners. *Creative Education*, 3, 1320-1325.
- Shi, J., & Zhang, X. (2023). Integration of AI with Higher Education Innovation: Reforming Future Educational Directions. *International Journal of Science and Research (IJSR)*. <https://doi.org/10.21275/sr231023183401>
- Wongvorachan, T., & Bulut, O. (2022). Feedback Generation Through Artificial Intelligence. *Open/Technology in Education, Society, and Scholarship Association Conference Proceedings*, 2(1), 1-9. <https://doi.org/10.18357/otessac.2022.2.1.125>
- Yang, A. (2024). Challenges and Opportunities for Foreign Language Teachers in the Era of Artificial Intelligence. *International Journal of Education and Humanities*.

Appendices

Appendix A: Grammar and Spelling Analysis Rubric

Grammar

Criteria	Poor (1)	Fair (2)	Good (3)	Excellent (4)
Sentence Structure	Frequent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Verb Tense and Form	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Pronoun Usage	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Prepositions	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Article Use	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Modifiers	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Run-on Sentence	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Sentence Fragment	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.

Spelling

Criteria	Poor (1)	Fair (2)	Good (3)	Excellent (4)
Spelling Accuracy	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Consistency	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Homophones	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Compound Words	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.
Syllable Division	Persistent errors, very unclear.	Frequent errors, occasionally unclear.	Some errors, but generally clear.	No errors.

Appendix B: Argumentative Essays

Essay ID	Human Feedback Details	AI Feedback Details
Arg_Essay1_Student2	Corrected verb tense, addressed sentence fragments, spelling, and suggested clearer thesis.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay2_Student3	Corrected verb tense, addressed sentence fragments, spelling, suggested clearer thesis, and noted irrelevant sentences.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay3_Student4	Corrected verb tense, addressed sentence fragments, spelling, suggested clearer thesis, and noted irrelevant sentences.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay4_Student5	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.

Arg_Essay5_Student7	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay6_Student8	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay7_Student9	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay8_Student10	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay9_Student11	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay10_Student12	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay11_Student14	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay12_Student15	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay13_Student16	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.

Arg_Essay14_Student18	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.
Arg_Essay15_Student19	Corrected verb tense, addressed sentence fragments, spelling, praised vocabulary variety, and suggested structural improvements.	Corrected spelling, verb tense, sentence fragments, provided corrections to make sentences clearer and more concise.

Opinion Essays

Essay ID	Human Feedback Details	AI Feedback Details
Op_Essay1_Student1	Corrected verb tense, spelling, suggested additional content, noted irrelevant sentences, praised vocabulary variety, and pointed out areas for improvement.	Corrected spelling, verb tense, sentence fragments, suggested clearer thesis, provided corrections to make sentences clearer and more concise.
Op_Essay2_Student2	Corrected verb tense, spelling, suggested clearer thesis statement, emphasized the need for topic sentences and more detailed examples, and provided overall encouragement.	Corrected spelling, verb tense, suggested clearer thesis, provided corrections to make sentences clearer and more concise.
Op_Essay3_Student3	Provided grammar and spelling corrections, emphasized the need for more detailed examples, and praised the clarity of the main argument.	Corrected grammar and spelling errors, suggested more concise sentence structures, and provided general comments on improving clarity and coherence.
Op_Essay4_Student4	Suggested improvements to essay structure, corrected grammatical errors, and highlighted areas needing more evidence and support.	Focused on correcting grammatical errors, providing clearer sentence structures, and suggesting a more coherent thesis statement.
Op_Essay5_Student5	Noted strong vocabulary use, corrected grammar and spelling, and suggested additional details to support the main argument.	Corrected grammar and spelling errors, recommended more precise language, and suggested improvements to the clarity of the thesis statement.

Op_Essay6_Student6	Emphasized the need for clearer topic sentences, provided grammar and spelling corrections, and suggested more detailed examples.	Focused on correcting grammatical errors, suggested more concise and clearer sentence structures, and provided general feedback on improving essay coherence.
Op_Essay7_Student7	Corrected verb tense, suggested additional supporting evidence, praised the organization of ideas, and highlighted areas for improvement.	Corrected grammatical errors, suggested clearer sentence structures, and provided general feedback on improving the thesis statement and overall coherence.
Op_Essay8_Student8	Provided detailed grammar and spelling corrections, suggested improvements to essay structure, and highlighted areas needing more detailed support and evidence.	Focused on correcting grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the thesis statement.
Op_Essay9_Student9	Emphasized the need for clearer topic sentences, provided detailed grammar and spelling corrections, and suggested more detailed examples and supporting evidence.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay10_Student10	Corrected verb tense and spelling errors, praised the organization of ideas, and suggested additional supporting evidence and more detailed examples.	Corrected grammatical errors, suggested clearer sentence structures, and provided general feedback on improving the coherence and clarity of the thesis statement.
Op_Essay11_Student11	Noted strong vocabulary use, corrected grammatical errors, emphasized the need for more detailed supporting evidence, and provided overall encouragement.	Focused on correcting grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay12_Student12	Corrected grammar and spelling, suggested improvements to essay structure, and highlighted areas needing more detailed support and evidence.	Corrected grammatical errors, suggested more concise and clearer sentence structures, and provided general feedback on improving the clarity of the thesis statement and overall coherence.

Op_Essay13_Student13	Provided detailed grammar and spelling corrections, suggested clearer topic sentences, and emphasized the need for more detailed examples and supporting evidence.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay14_Student14	Emphasized the need for clearer topic sentences, corrected grammatical errors, and highlighted areas needing more detailed supporting evidence and examples.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay15_Student15	Noted strong vocabulary use, corrected grammatical errors, suggested additional supporting evidence, and provided overall encouragement.	Corrected grammatical errors, suggested clearer sentence structures, and provided general feedback on improving the clarity and coherence of the thesis statement.
Op_Essay16_Student16	Provided detailed grammar and spelling corrections, suggested clearer topic sentences, and emphasized the need for more detailed examples and supporting evidence.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay17_Student17	Corrected verb tense and spelling errors, suggested improvements to essay structure, and highlighted areas needing more detailed support and evidence.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the thesis statement.
Op_Essay18_Student18	Emphasized the need for clearer topic sentences, corrected grammatical errors, and highlighted areas needing more detailed supporting evidence and examples.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the essay.
Op_Essay19_Student19	Provided detailed grammar and spelling corrections, suggested improvements to essay structure, and emphasized the need for more detailed supporting evidence and examples.	Corrected grammatical errors, suggested clearer and more concise sentence structures, and provided general feedback on improving the coherence and clarity of the thesis statement.

Op_Essay20_Student20	Corrected verb tense and spelling errors, praised the organization of ideas, and suggested additional supporting evidence and more detailed examples.	Corrected grammatical errors, suggested clearer sentence structures, and provided general feedback on improving the coherence and clarity of the thesis statement.
-----------------------------	---	--

Appendix C:

Example of an Argumentative Essay with Human Feedback

(Title?)

Everybody knows, one of the important things in sports are passion. But what about money? Also money is an important aspect in sports because athletes need equipment for doing their sports and they need money for equipment. In these(At this) point, sponsorship is a great opportunity to make money while doing sports. Some people think, sponsors and companies disturbe(disturb) sports sprits. But I don't agree with this. Sponsorship and ads are an important part of expensive sports like Formula 1 and fencing.

Firstly, some sports are too expensive, and they(?) need help to keep doing that(them). Lets(Let's) take Formula 1. Formula 1 are(is) composed of cars, and these cars need maintaumamce (maintenance). These cars have too expensive parts like engines and tires. For the tires, in every race, the cars require a couple of new tires set. In(At) this point, tires company can handle it. There is(are) lots(a lot) of companies for every spare parts, and tires are just an(one) example.

Secondly, the other expensive sport is fencing. The armor and swords is(are) a big trouble for athletes. But the-company(companies) says, "if you win these(this) game and you carry our brand on your back, we are going to pay you very well". That element keeps athletes motivated, and they always work hard for winning(to win). This motivation is important for succes(success).

Last one is (Lastly), sponsors and athletes encourage teens who are interent(interested) in doing sports. Young people see that adrenaline, excitement and success they want to do they(?). Also, companies help these children and teens to do(realize) they(their) dreams. They made(organize) some events and tournaments for young people. It's a great opportunity for teens.

In conclusion, some people think sponsorship has negative effect on sports. But companies has(have) lots of positive side(effects/impacts) on sports. They help teams for handle financial problems. Also, they motivate athletes to be successful and hard-working. The(Their) other one(influence) is, they help teams who want to be(have) a-successful athletes. I think sponsors are not as useless as we tought(think).

Example of an Argumentative Essay with AI Feedback

(Title?)

Everybody knows, one of the important things in sport *are passion* (is passion). But what about money? Also, money is an important aspect in sport because athletes need equipment for doing their sports and they need money for equipment. *At these point* (this point), sponsorship is a great opportunity to make money while doing sport. Some people think sponsors and companies *disturbe* (disturb) sports *sprits* (spirits). But

I don't agree with this. Sponsorship and ads are an important part of expensive sports like Formula 1 and fencing.

Firstly, some sports are too expensive and they need help to keep doing that. Let's take Formula 1. Formula 1 *are composed of* (consists of) cars, and these cars need *maintaumamce* (maintenance). These cars have expensive parts like engines and tires. For the tires, every race the cars require a new set of tires. At this point, a tire company can handle it. There *is lots* (are lots of) companies for every spare part; tires (are) just an example.

Secondly, another expensive sport is fencing. The armor and swords *is* (are) a big trouble for athletes. But the company says, "if you win *these game* (this game) and you carry our brand on your back, we are going to pay you very well." This element keeps athletes motivated and they always work hard for winning. This motivation is important for *succes* (success).

Last one is, (Lastly,*) sponsors and athletes encourage teens who are *interent* (interested) in doing sports. Young people see that adrenaline, excitement, and success; they want to do *they* (emulate them). Also, companies help these children and teens *to do they dreams* (achieve their dreams). They organize events and tournaments for young people. It's a great opportunity for teens.

In conclusion, some people think sponsorship has a negative effect on sports. But companies have lots of positive sides in sports. They help teams handle financial problems. Also, they motivate athletes to achieve success and work hard. *Other one is* (Another one is), they assist teams who want to become successful athletes. I think sponsors are not as useless as we *tought* (thought).