American Journal of Humanities and Social Sciences Research (AJHSSR)
e-ISSN : 2378-703X
Volume-09, Issue-12, pp-43-50
www.ajhssr.com
Research Paper Open Access

Blending Human and Artificial Intelligence: An Empirical Study of Student Perceptions of AI-Generated Feedback in EFL writing Instruction

Merrilee Brinegar
(Graduate School of TESOL, Hankuk University of Foreign Studies, South Korea)

ABSTRACT: This study examines Korean university students' perceptions of AI-generated feedback compared to traditional instructor feedback in an EFL writing course. Thirty-one undergraduate students received blended feedback from both their instructor and GPT-4 on essay assignments. Survey data assessed students' ability to distinguish between feedback sources, perceived helpfulness, comprehension levels, and preferences. Results indicate that 83.9% of students found AI feedback helpful or very helpful, while 41.9% could not distinguish between AI-generated and teacher feedback. Notably, 67.7% preferred receiving both AI and teacher feedback rather than a single source. Students particularly valued grammar corrections, organizational suggestions, and concrete examples. These findings suggest that AI-generated feedback can effectively complement traditional instructor feedback when thoughtfully integrated into writing pedagogy.

Keywords : *AI feedback, EFL writing, student perceptions, automated writing evaluation, generative AI, English language teaching*

I. INTRODUCTION

The influx of artificial intelligence into education has caused many changes, especially in terms of teaching writing in a second language. With English as a lingua franca, there is an ever-growing need for improved English as a Foreign Language (EFL) instruction. This puts a lot of pressure on teachers, who have to juggle giving detailed feedback to large classes. While the traditional methods of providing writing feedback may still prevail, many teachers lack sufficient time or resources to provide ample detailed feedback to each student.

With the rise of large language models like GPT-4, a whole new world of possibilities has been opened for dealing with such writing challenges. Unlike older automated writing evaluation systems, today's AI tools can analyze and comment on areas including organization, argumentation, and content development. However, as these technologies are becoming more pervasive, research on how effective they really are, especially in terms of how students engage with AI feedback versus traditional instructor feedback, has not quite been fully explored. Thus, this study investigates how Korean university students experience a mix of feedback from both human instructors and AI-generated comments. Korea has some of the highest rates of English language learners in the world, and these students face unique hurdles when it comes to academic writing, as English differs from Korean in terms of using articles correctly, keeping verb tenses consistent, and different rhetorical styles.

This study is driven by several research questions: Can students tell the difference between feedback from AI and their instructors? How do they feel about the usefulness and clarity of AI feedback? What types of feedback do they find most helpful? What is student's view on blending AI with human input? By investigating these questions, this research contributes to understanding both the theory behind using AI in education and offers practical guidance for teachers considering incorporating these tools into their writing classes.

II. LITERATURE REVIEW

1. Evolution of Automated Writing Evaluation and Generative AI

Technology's role in giving feedback on student writing has changed dramatically over the years. In its infancy, automated writing evaluation systems mainly focused on basics such as spelling, grammar, and simple sentence structure. They relied on rule-based algorithms and pattern matching to get the job done, but as natural language processing grew more sophisticated, tools like Grammarly and Pigai started offering deeper linguistic analysis.

The research around traditional automated writing evaluation (AWE) systems has shown mixed results. For instance, Strobl et al. (2019) examined 44 digital tools for providing automated feedback and discovered that most of them mainly focused on micro-level features like grammar and spelling errors. In contrast, Fu et al. (2022) reviewed 48 studies and found that while automated feedback can give writing a boost, it generally does not provide the insights you would get from human feedback. Interestingly, a meta-analysis by Fleckenstein et al. (2023) showed that automated feedback does have a medium positive effect on writing performance (d = 0.62), but there is a large amount of variation depending on different contexts and learner traits.

Now, with generative AI, there are massive upgrades in what AI-powered writing tools can do. Unlike older AWE systems that stuck to set rules, large language models like GPT-4 can generate text that is not only contextually relevant but also sounds human-like and can offer nuanced feedback on various writing issues. Escalante et al. (2023) compared GPT-4's feedback with what human tutors provided and found that students thought the AI-generated comments were clearer and more detailed than their instructors' notes. However, students who favored face-to-face interactions mentioned emotional benefits that just cannot be matched by AI.

In recent studies focused on English as a Foreign Language (EFL), researchers have explored how generative AI affects learning outcomes. Song and Song (2023) found that Chinese EFL learners using ChatGPT saw notable improvements in their writing skills and felt more motivated because they appreciated getting real-time, personalized feedback. Similarly, research from Dai et al. (2023) highlighted that students enjoyed the specific and clear nature of comments generated by ChatGPT, reinforcing earlier findings that AI feedback can often be richer than what instructors provide.

2. Student Perspectives and Blended Approaches

Understanding how students view AI-generated feedback is important for making it work effectively. A review by Shi and Aryadoust (2024) looked at 11 different studies and found that most students had a mostly positive outlook on automated feedback. Only two studies pointed to a preference for traditional human feedback, while five had mixed results; students liked grammar-related feedback but were not satisfied with the content-related comments.

When it comes to generative AI tools, the research has shown mostly positive student perceptions. Utami et al. (2023) discovered that AI tools helped Indonesian EFL learners boost their academic writing and kept them engaged. The quick turnaround on feedback was especially appreciated; however, Dai et al. (2023) found that even though AI feedback was clear and detailed, some students still wanted interpersonal interaction from a tutor. Thus, the best approach might be using AI as a helpful addition rather than trying to replace human input altogether.

There is also growing interest in blending AI with human feedback. Asadi et al. (2025) investigated how using ChatGPT alongside teacher feedback helped EFL learners improve their writing skills. They noted that the way teachers perceive these tools plays a huge role in how well they actually get used. In another study, Wiboolyasarin et al. (2024) explored AI-enhanced collaborative writing with Thai undergrads and found it useful, highlighting how effective AI can be when woven into the revision process.

Nonetheless, there are still some gaps in the research. Not enough studies have directly compared how students experience AI versus human feedback in real classroom settings. Additionally, there is limited research specifically looking at Korean university students' experiences. It is also unknown which types of feedback students find most helpful. Lastly, whether EFL students can actually distinguish between AI-generated and human feedback has not yet been extensively explored, which is a critical factor related to trust and transparency in student-teacher relationships in an increasingly tech-driven learning environment.

III. METHODOLOGY

1. Research Context and Participants

This study took place at a major university in Seoul, South Korea, in the fall of 2025. Thirty-two undergraduate students were enrolled in an academic reading and writing course for three hours a week over the span of 15 weeks. The participants were mostly between 19 and 23 years old, mainly first-year students, and all were native Korean speakers with English skills ranging from intermediate to advanced. A survey showed that 90.3% of them used AI platforms either weekly or even daily, which underscores how comfortable they are with these technologies.

2. Course Design and Feedback Integration

The course included both reading and writing and had students practice both skills each week covering various topics. For this study, one of their academic essays written in class received a mix of feedback that combined traditional teacher comments with insights generated by AI using Claude. The feedback given to students did not specify which comments came from the teacher and which came from the AI, appearing as a mix of in-text changes, highlighting of errors, margin notes, and comments. The in-text changes and highlighting portions focused on issues covered repeatedly in class such as the use of informal language, conjunctions starting a sentence, and subject-verb agreement. The comments were related to the scoring criteria, which included topic sentences, claims and arguments, supporting evidence, transitions and flow, vocabulary and word variety, paragraph structure, and analysis and conclusion. The aim was for the AI to be constructive and encouraging, providing clear and actionable feedback suitable for intermediate to advanced EFL learners,

including strengths, areas for improvement, and specific suggestions for revision. Students were informed that the feedback was from both their teacher and an AI assistant but without any labels identifying who said what. This setup allowed for exploration into whether students could tell the difference between the sources without bias.

2. Data Collection and Analysis

Data was gathered through an online survey after the students were given feedback on their academic essay. This timing gave them enough experience to share meaningful opinions. The survey mixed closed-ended and open-ended questions. While 32 students participated in the survey, one declined to be part of the research. Thus, 31 student responses were analyzed. The survey used Likert scales and multiple-choice questions to find out how students distinguished between different feedback sources, how helpful they found AI feedback, how they rated the amount of feedback, and how easy it was for them to understand. They also got to choose their preferred type of feedback, pinpoint which types were the most helpful, and share how often they used AI tools. Additionally, open-ended questions were included where students could suggest ways to improve feedback and share their thoughts on using AI as a feedback tool.

IV. RESULTS

1. Distinguishing Between AI and Teacher Feedback

When students were asked if they could determine which comments came from AI and which ones were from their teacher, 41.9% (n=13) said they could not tell the difference at all, and 25.8% (n=8) said they were not really sure. In contrast, 29% (n=9) said they could sometimes figure it out, but only one person claimed they could always tell them apart. It is interesting to note that 67.7% felt like the feedback from AI was equal to what they would get from a human teacher in terms of quality, despite their not knowing which feedback came from the teacher or AI.

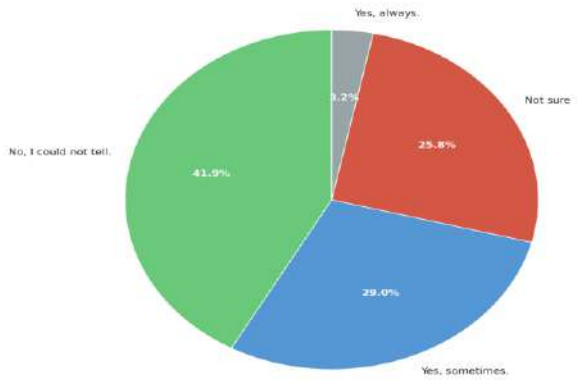

Figure 1: Could students distinguish AI from teacher feedback?

2. Perceived Helpfulness of AI Feedback

Students had a generally positive view of AI feedback. When asked how helpful they found it, about 19.4% (n=6) said it was "very helpful," while 64.5% (n= 20) rated it as "helpful." Only 16.1% (n=5) thought it was "a little helpful." Interestingly, no participants responded by saying it was "not helpful" at all (see Fig. 2). Thus, 83.9% of students felt that AI feedback was substantially beneficial.

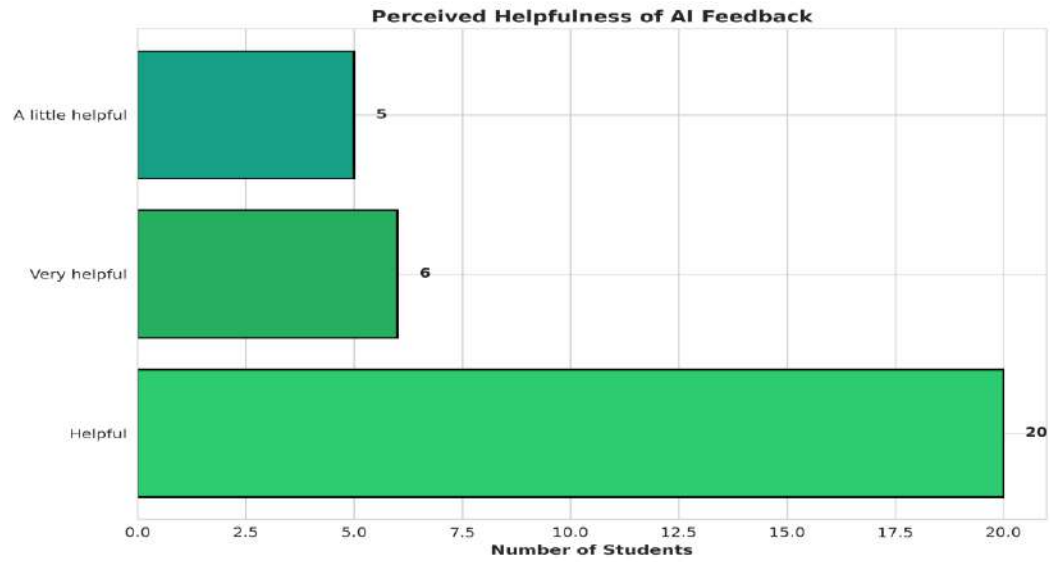

Figure 2: Perceived helpfulness of AI feedback

3. Feedback Comprehensibility and Amount

All 31 students felt the total amount of feedback was appropriate, suggesting this blended approach provided enough guidance without making the students feel overwhelmed. In terms of how easy the feedback was to understand, 64.5% (n=20) thought both types were equally clear. In contrast, 22.6% (n=7) answered that the teacher's feedback was easier to grasp, while 9.7% (n=3) said they preferred the AI feedback. Interestingly, just one student thought both were equally difficult to understand. Notably, almost two-thirds felt the AI-generated feedback was just as clear as the feedback they judged to be from the teacher.

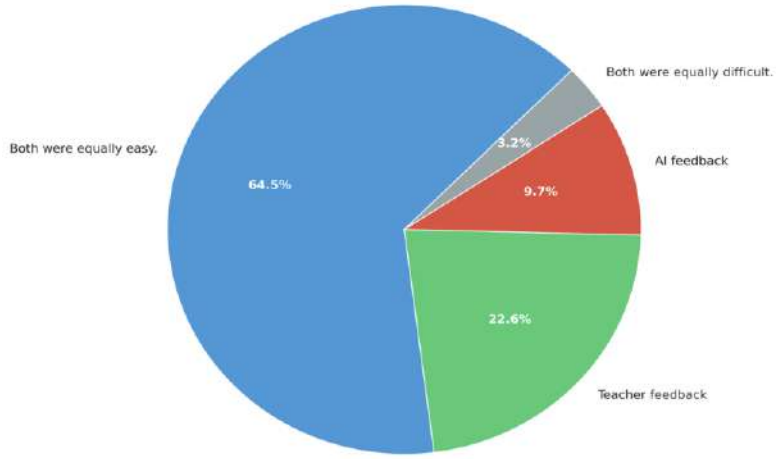

Figure 3: Which feedback was easier to understand?

4. Student Preferences for Feedback Type

The majority, 67.7% (n=21) said they would prefer to get feedback from both AI and teachers. Moreover, 22.6% (n=7) did not express a strong preference either way. In contrast, only 6.5% (n=2) wanted only teacher feedback, and one student opted for only AI feedback (see Fig. 4). This strong preference for both types of feedback indicates that students see the value of having different viewpoints on their writing.

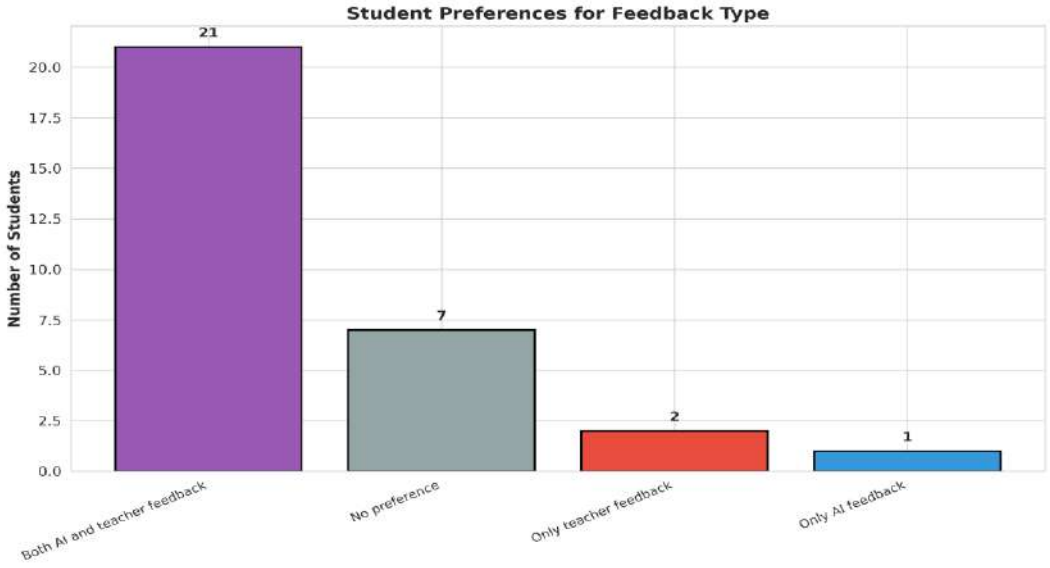

Figure 4: Student preferences for feedback type

5. Most Helpful Type of Feedback

In regards to the most helpful type of feedback, students chose various kinds from a list provided on the survey. The most popular choice was grammar corrections, at 54.8% (n=170, followed by comments on organization at 48.4% (n=15). Examples of better sentences was chosen as most helpful by 41.9% (n=13), feedback on ideas and content was chosen by 29% (n=9), and vocabulary suggestions received 25.8% (n=8). It is clear that all five types of feedback were useful for some students, showing the importance of addressing various aspects of writing with thorough feedback.

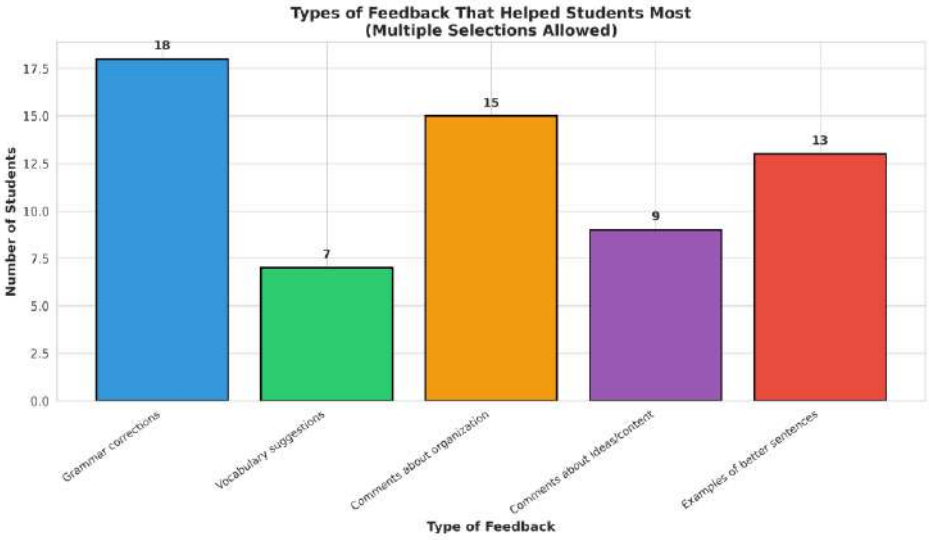

Figure 5: Types of feedback that helped most

6. Engagement with Feedback and AI Usage

The students showed high engagement with the feedback they received: 61.3% (n=19) reported having read all the feedback, and 25.8% (n=8) looked at most of it. However, 9.7% (n=3) read just some of the feedback, and one student read very little of it (see Fig. 6). Thus, 87.1% of students engaged substantially with the feedback.

Examining often the learners used AI tools, it is notable that 54.8% (n=17) are using them weekly, while 35.5% (n=11) utilize AI tools daily. A smaller group, 6.5% (n=2), uses them monthly, and just one student rarely uses AI tools. Thus, when considering that nearly 90.3% use AI frequently, it is not surprising how comfortable these students are with AI feedback in general.

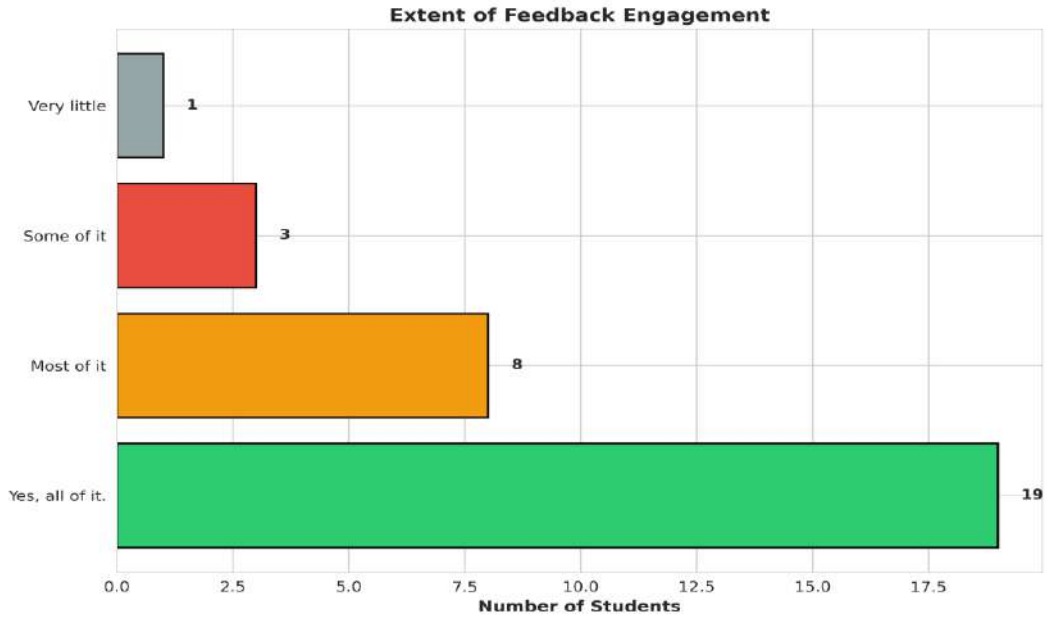

Figure 6: Extent of feedback engagement

7. Qualitative Themes

After examining the open-ended responses, a few key themes emerged. First, students definitely want more examples and models. Several mentioned that having concrete examples of improved sentences or restructured paragraphs would make the feedback much more helpful. Additionally, there is a desire for feedback to be more detailed. Beyond being told what needs to change, students want to know why it matters and how they can go about making those changes. Grammar and language usage also appeared as important points, with students really valuing grammar corrections and asking for even more focus on those details. Another area students appreciated was feedback on organization and structure.

In terms of AI, the comments were mostly positive. One student pointed out, "I think AI feedback is more delicate, so it is useful." Another shared, "I believe using AI in assignments is appropriate given the current state of affairs... Instead of seeing AI as just a bad thing, we should focus on teaching students how to use it properly as a new tool." Interestingly, one student did not even realize some feedback was from AI until it was mentioned. They said, "Actually before I see your mention, I didn't recognize. For me, AI feedback is not bad," which shows that some students cannot really tell where the feedback is coming from and do not view AI-generated input as less valuable.

V. DISCUSSION

1. The Indistinguishability and Complementary Value of Blended Feedback

One interesting finding is that most students could not consistently tell the difference between feedback generated by AI and that from their instructors. This shows that today's large language models can produce comments that similar to what a human teacher would say, which is a big improvement from older, formulaic automated writing evaluation systems. However, the question of transparency remains: should students be informed about which comments come from AI and which are from their teachers?

Another key takeaway is how much students prefer getting feedback from both AI and their teachers, with 67.7% of them feeling this way. It seems that they value the different perspectives each source provides. Getting input from various sources can give students a fuller picture of their writing strengths and weaknesses since different sources might draw attention to different aspects or explain similar issues in various ways. This overlap can actually be helpful, not just redundant; it reinforces certain points when both sources highlight the same problems, while also possibly encouraging deeper thinking when there is a discrepancy.

This preference for mixed feedback has some serious implications for how we think about AI's role in writing instruction. Instead of seeing AI as something that could replace teachers, these findings suggest it might work better as a complementary tool that boosts what teachers can offer. Accordingly, AI should be viewed as an assistant to human instruction, not a replacement. However, if we want to implement this blended approach effectively, we need to be mindful of avoiding too much repetition while still giving thorough comments and corrections and helping students make sense of feedback from different sources.

2. Grammar Focus and High Engagement

Grammar corrections topped the list as the most helpful type of feedback at 54.8%. For EFL learners, grammatical accuracy is a consistent challenge that can strongly affect how well they communicate. Korean students, in particular, find article usage, keeping verb tenses consistent, and making sure subjects agree with verbs to be difficult. Their strong recognition of grammar corrections as useful probably shows they have a realistic view of their own needs rather than just focusing too much on details. However, teachers using AI-generated feedback need to ensure that stressing grammar does not take the focus away from bigger issues in writing. A large number of students (48.4%) also found organizational comments helpful, 41.9% appreciated examples, and 29.0% appreciated content-related feedback. This suggests that a mixed approach works well by covering various aspects of writing.

It is also significant that 87.1% of students read all or most of the feedback they got. Several reasons might explain this high engagement: the quality and usefulness of the feedback itself, the fact that it was not too overwhelming for learners, maybe even their curiosity about where the feedback came from, and course requirements pushing them to learn from the feedback they received. These encouraging engagement rates could give educators confidence about integrating AI feedback into their teaching since it seems like thoughtful AI commentary does connect with students meaningfully.

3. Limitations and Future Directions

There are a few limitations of this study that should be acknowledged. For one, the small sample size (31 students) from a single institution in Korea makes it hard to generalize the findings. It would be beneficial to see future research take on multiple schools with larger and more diverse groups of students. Additionally, relying solely on self-reported survey data does not really provide a clear picture of how AI feedback actually affects writing quality. Augmenting with methods that examine both perceptions and actual writing samples would be useful for future studies.

Another issue is that investigating students' perceptions just once during a course may not really show how their attitudes might shift over time. A longer-term study following students could reveal whether those positive feelings lead to lasting improvements in their writing. Furthermore, not labeling the feedback means we cannot really compare how students view AI comments versus teacher feedback. Future research could benefit from clearly marking from whence each type of feedback originated for a deeper analysis.

Finally, without comparison groups that only received teacher feedback or just AI feedback, it is challenging to make strong claims about how effective this blended approach really is. Using experimental designs that test out different feedback setups could definitely help solidify those conclusions.

Nonetheless, even with these limitations, this study adds to the growing conversation about using AI in writing instruction. It provides some solid evidence about how EFL students interact with both AI and teacher feedback, which can guide educators thinking about integrating these tools while pointing out key areas for future research.

VI. CONCLUSION

This study investigated how Korean university students feel about AI-generated feedback in an English as a Foreign Language (EFL) writing course in which they got feedback from both their instructors and an AI learning model. The overwhelming majority at 83.9% found the AI feedback to be helpful. Interestingly, 67.7% of the students admitted they could not reliably distinguish between the AI feedback and what their instructor provided. This demonstrates that today's generative AI tools can provide feedback that is seen as being comparable to that of human teachers.

One of the main findings of this study is that Korean EFL students actually prefer getting feedback from both AI and teachers rather than just one of them. This indicates that they see value in having different perspectives rather than treating AI and human input as rivals. Instead, it paints a picture of AI as a tool that can enhance what teachers do rather than taking their place.

In terms of the kind of feedback students found most useful, grammar correction was the highest at 54.8%. Organizational comments came in next at 48.4%, with examples following closely behind at 41.9%. This highlights some ongoing challenges students face in mastering grammar while also recognizing how crucial it is to have clear organization in their writing. Moreover, with 87.1% of students reading all or most of the feedback they received, it is evident that well-designed AI feedback resonates with learners.

For educators pondering incorporating AI feedback into their teaching, there are some important considerations. Carefully designed AI-generated feedback can effectively complement traditional teacher input. Students appreciate having multiple sources of feedback, and they find blended approaches more valuable than sticking with just one type. It is vital for the AI feedback to attend to various aspects of writing, addressing both surface-level issues and deeper concerns as well. The AI feedback needs to fit into a structured revision process that encourages students to engage with it seriously, meaning they should read it, understand it, and actually use those suggestions.

As generative AI tech continues to evolve, its role in writing instruction could grow even more significant, but these tools should enhance rather than replace the essential connections, conversations, and understanding that come from human teachers. The aim should be to harness AI in a way that makes writing instruction more effective and responsive to individual student needs while still keeping human elements front and center.

This study shows that when instructors thoughtfully combine AI-generated feedback with traditional teacher insights, student writing development can be supported within EFL settings. Korean university students find this blend helpful and engage with it meaningfully, clearly favoring approaches that mix the strengths of both human and AI feedback. These findings should encourage educators who are considering bringing AI into their classrooms but also remind them how important it is to implement these tools thoughtfully while keeping student needs and essential teaching elements in mind.

REFERENCES

[1] Strobl, C., Ailhaud, E., Benetos, K., Devitt, A., Kruse, O., Proske, A., & Rapp, C. (2019). Digital support for academic writing: A review of technologies and pedagogies. Computers & Education, 131, 33-48.

[2] Fu, Q. K., Zou, D., Xie, H., & Cheng, G. (2022). A review of AWE feedback: Types, learning outcomes, and implications. Computer Assisted Language Learning, 37(1-2), 179-221. https://doi.org/10.1080/09588221.2022.2033787

[3] Fleckenstein, J., Liebenow, L. W., & Meyer, J. (2023). Automated feedback and writing: A multi-level meta-analysis of effects on students' performance. Frontiers in Artificial Intelligence, 6, 1162454.

[4] Escalante, J., Pack, A., & Barrett, A. (2023). AI-generated feedback on writing: Insights into efficacy and ENL student preference. International Journal of Educational Technology in Higher Education, 20(1), Article 57. https://doi.org/10.1186/s41239-023-00425-2

[5] Song, C., & Song, Y. (2023). Enhancing academic writing skills and motivation: Assessing the efficacy of ChatGPT in AI-assisted language learning for EFL students. Frontiers in Psychology, 14, 1260843. https://doi.org/10.3389/fpsyg.2023.1260843

[6] Dai, W., Lin, J., Jin, H., Li, T., Tsai, Y. S., Gašević, D., & Chen, G. (2023). Can large language models provide feedback to students? A case study on ChatGPT. IEEE International Conference on Advanced Learning Technologies, 323-325.

[7] Shi, H. and Aryadoust, V. (2024). A systematic review of AI-based automated written feedback research. ReCALL, 36(2), 187-209.

[8] Utami, S. P. T., Andayani, Winarni, R., & Sumarwati. (2023). Utilization of artificial intelligence technology in an academic writing class: How do Indonesian students perceive? Contemporary Educational Technology, 15(4) ep450. https://doi.org/10.30935/cedtech/13419

[9] Asadi, M., Ebadi, S., & Mohammadi, L. (2025). The impact of integrating ChatGPT with teachers' feedback on EFL writing skills. Thinking Skills and Creativity, 56, 101766.

[10] Wiboolyasarin, W., Wiboolyasarin, K., Suwanwihok, K., Jinowat, N., & Muenjanchoey, R. (2024). Synergizing collaborative writing and AI feedback: An investigation into enhancing L2 writing proficiency in wiki-based environments. Heliyon, 10(8), e29234.