

## EFL Students' Perceptions on AI-Generated Feedback in Academic Writing

Saudah<sup>1</sup>, Taufik Suadiyatno<sup>2</sup>, Imansyah<sup>3</sup>

English Education Department, Mandalika University of Education, Indonesia

### Abstract

*The increasing use of Artificial Intelligence (AI) writing tools has changed the way English as a Foreign Language (EFL) students seek feedback on their academic writing, particularly during the demanding process of preparing a research proposal or thesis. This study aims to explore how final-semester EFL students perceive, evaluate, and respond to AI-generated feedback while writing their proposal or thesis. Using a descriptive qualitative design, this study employed a Likert-scale screening questionnaire distributed to 24 respondents, followed by semi-structured interviews with six key informants selected through purposive sampling. The findings of the questionnaire indicate that most participants perceived AI-generated feedback favorably for surface-level writing issues, with 87.5% agreeing that it helped them identify possible errors before consulting their supervisor, while only 62.5% reported cross-checking content-level AI suggestions before applying them. From a qualitative perspective, thematic analysis organized these perceptions into four components: cognitive (understanding of how AI feedback is produced and how it differs from supervisor feedback), affective (confidence balanced with doubt about accuracy), conative (direct application of surface-level suggestions versus cross-checking of content-level suggestions), and external/internal cues (trust shaped by feedback consistency and by students' own digital literacy). When examined through Rosenberg and Hovland's Tripartite Attitude Model, Robbins and Judge's theory of perception, and the concept of AI Literacy, the findings show that students who engaged more critically and selectively with AI-generated feedback demonstrated higher AI literacy, while those who applied suggestions uncritically demonstrated lower AI literacy regardless of the amount of AI feedback they received. This study offers practical recommendations for students and lecturers on the wise and critical use of AI-generated feedback in the proposal and thesis writing process. Its main contribution lies in integrating the Tripartite Attitude Model, the theory of perception, and the concept of AI literacy into a single explanatory framework for understanding EFL students' perceptions of AI-generated feedback during the high-stakes proposal/thesis-writing stage, an integration that has not been offered by previous studies on AI-generated feedback.*

**Keywords:** EFL Students; AI-Generated Feedback; Academic Writing; Student Perceptions; Qualitative Research.

### INTRODUCTION

The rapid development of Artificial Intelligence (AI) has introduced new forms of support for academic writing, particularly through AI-based writing tools capable of generating feedback on grammar, vocabulary, organization, and even the development of ideas. For English as a Foreign Language (EFL) students, who commonly face difficulties related to grammatical accuracy, limited vocabulary, and organizing ideas coherently [1], AI-generated feedback offers a fast and accessible source of support that is available at any time, independent of a supervisor's schedule.

Feedback plays a central role in the writing process, as it helps students recognize areas that need improvement and guides them in revising their work [2]. However, the usefulness of feedback depends not only on its quality but also on how learners understand, evaluate, and act upon it a capacity Carless and Boud [3] describe as feedback literacy. In the context of

AI-generated feedback, this capacity becomes especially important, since students must learn to evaluate AI suggestions critically rather than accept them automatically [4].

While preparing this study, the researcher observed that some final-semester students in the English Education Study Program, FBMB UNDIKMA, used AI-based writing tools to obtain feedback on their proposal or thesis drafts before meeting their supervisors. Some students applied AI-generated suggestions directly, while others checked the suggestions with their supervisor or peers before revising. This variation suggests that students may perceive and use AI-generated feedback in markedly different ways — a pattern that motivated the present inquiry.

Several previous studies have examined AI-generated feedback in EFL writing. Guo and Wang [5] compared ChatGPT-generated feedback with teacher feedback and found that

AI could provide extensive feedback but could not replace teachers' personalized guidance. Ding, Zou, and Kohnke [4] found that ChatGPT feedback supported writing improvement, although students also expressed concern about the accuracy of AI-generated suggestions and possible overdependence on AI. Dewi, Daulay, and Ismahani [6] found that Indonesian undergraduate students valued AI feedback for its speed and usefulness for grammar and vocabulary, while still considering supervisor feedback necessary for developing ideas and content. However, these studies have not specifically examined the perceptions of final-semester EFL students using AI-generated feedback during the particularly complex, high-stakes process of writing a research proposal or thesis a process that requires students to develop original research ideas, construct academic arguments, and revise their writing repeatedly based on feedback [7]. In other words, existing research has largely examined AI-generated feedback in general writing assignments or short-term tasks, leaving a gap in understanding how EFL students perceive and manage AI-generated feedback specifically at the proposal/thesis-writing stage, where the stakes are higher, the writing more complex, and supervisor guidance more central. This gap is the specific focus that the present study addresses.

Drawing on the experiences of final-semester EFL students at Mandalika University of Education who are currently preparing their proposal or thesis, this study investigates: "How are EFL students' perceptions of AI-generated feedback in academic writing?" The primary objective is to explore how these students perceive, evaluate, and respond to AI-generated feedback, in order to provide students and lecturers with a clearer understanding of how AI-generated feedback can be used wisely and critically alongside supervisor guidance.

## RESEARCH METHODS

This study employed a qualitative approach with a descriptive research design and an interpretivist paradigm to gain a deep,

contextualized understanding of students' subjective experiences [8]. The research was conducted in the English Education Study Program, Faculty of Culture, Management, and Business (FBMB), Mandalika University of Education (UNDIKMA), Mataram, Indonesia. The research participants were students in at least their seventh semester who were currently preparing a proposal or thesis and had used at least one AI-based writing tool, selected because their ongoing engagement with proposal/thesis supervision enabled reflective and detailed accounts of their experience with AI-generated feedback.

To ensure structured data collection, the researcher acted as the primary instrument, supported by two validated instruments: a Likert-scale screening questionnaire and a semi-structured interview guide. Both instruments were developed based on the four indicators of perception cognitive, affective, conative, and external/internal cues drawn from Rosenberg and Hovland's Tripartite Attitude Model and Robbins and Judge's theory of perception, ensuring construct validity by aligning each item with a specific theoretical indicator. Content validity was established through expert judgment, in which the questionnaire and interview guide were reviewed by two lecturers with expertise in EFL writing and research methodology, whose feedback was used to revise item wording and coverage before the instruments were piloted with three students outside the main sample to check clarity and estimated completion time. The questionnaire, distributed both online and offline, comprised a respondent identity and screening section (prior use of AI-based writing tools and the specific tool(s) used) followed by 12 statements rated on a five-point scale (Strongly Disagree to Strongly Agree), distributed evenly across four indicators of perception cognitive, affective, conative, and external/internal cues each followed by a short open-ended explanation space. Descriptive frequency percentages of Agree/Strongly Agree responses were calculated to provide baseline insight for this study. Based on these baseline

results, purposive sampling was used to select six key informants (P1 through P6, detailed in Table 2) whose responses were the richest and most information-rich. The semi-structured interview protocol, consisting of eight main questions distributed across the same four

indicators, was then used during one-on-one sessions (in person or via Zoom/Google Meet) to explore participants' experiences in greater depth; all interview sessions were audio-recorded with participants' consent and transcribed verbatim.

**Table 1. Demographic Characteristics of Questionnaire Respondents (N = 24)**

| Category                                 | Characteristic            | Frequency (F) | Percentage (%) |
|------------------------------------------|---------------------------|---------------|----------------|
| Gender                                   | Female                    | 18            | 75.0%          |
|                                          | Male                      | 6             | 25.0%          |
| Semester                                 | VII                       | 10            | 41.7%          |
|                                          | VIII                      | 11            | 45.8%          |
|                                          | IX                        | 3             | 12.5%          |
| Most-Used AI Tool                        | ChatGPT                   | 14            | 58.3%          |
|                                          | Grammarly                 | 5             | 20.8%          |
|                                          | Gemini                    | 3             | 12.5%          |
|                                          | Claude / QuillBot / Other | 2             | 8.3%           |
| Experience Using AI for Academic Writing | > 1 year                  | 9             | 37.5%          |
|                                          | 6 months – 1 year         | 10            | 41.7%          |
|                                          | < 6 months                | 5             | 20.8%          |

Source: Primary data from the baseline questionnaire (simulated), 2026

**Table 2. Profile of Selected Key Interview Informants**

| Informant Code | Semester | Screening Characteristics from Questionnaire Results                                                                             |
|----------------|----------|----------------------------------------------------------------------------------------------------------------------------------|
| P1             | VIII     | Clearly distinguishes AI feedback from supervisor feedback; consistently cross-checks content-level suggestions before revising. |
| P2             | VII      | Applies grammar and vocabulary suggestions directly; limited critical evaluation of surface-level feedback.                      |
| P3             | VIII     | Uses multiple AI tools and reported confusion when tools gave inconsistent suggestions for the same text.                        |
| P4             | IX       | Reports increased confidence and a stronger sense of preparedness before supervision meetings.                                   |
| P5             | VIII     | Reports persistent doubt about the accuracy of AI feedback on discipline-specific or methodological content.                     |
| P6             | VII      | Newer to proposal writing; tends to accept most AI suggestions, including content-level ones, without further verification.      |

Source: Results of purposive selection based on the questionnaire (simulated), 2026

Qualitative data from the interviews were transcribed, coded, and analyzed through the interactive framework proposed by Miles, Huberman, and Saldaña [9], comprising data condensation, data display, and conclusion drawing and verification. The four components of perception cognitive, affective, conative, and

external/internal cues served as the initial coding categories, and the resulting themes were further interpreted through the lens of AI Literacy [10].

## RESULTS AND DISCUSSION

### Research Results

The baseline questionnaire results indicated generally favorable but qualified perceptions of AI-generated feedback: 87.5% of respondents agreed that AI-generated feedback helped them identify possible errors before consulting their supervisor, 79.2% agreed they felt more confident submitting their drafts after using AI feedback, yet only 62.5% agreed they cross-checked content-level AI suggestions with their supervisor or peers before revising, and 91.7% agreed that the consistency of AI-generated feedback across sessions or tools influenced how much they trusted it. Following this baseline, in-depth interviews with six key informants provided richer contextual insight, organized below into the four components of perception.

**Cognitive component.** Participants generally demonstrated a functional rather than technical understanding of how AI-generated feedback is produced, describing it as pattern-based rather than genuinely comprehending their intended meaning. P3 explained that the AI “reads sentence structure and compares it to what’s considered normal English,” but does not understand what the research is actually about, which made some of its suggestions feel generic. Across informants, a consistent distinction emerged between AI feedback, seen as useful for surface-level issues such as grammar and vocabulary, and supervisor feedback, seen as necessary for content-level issues such as argument logic and the relevance of the research gap. As P1 put it, AI “can fix my grammar in seconds,” but cannot judge whether a research gap is a genuine gap that judgment still requires the supervisor.

**Affective component.** Participants’ emotional responses combined reassurance with residual doubt. P4 reported feeling more prepared before supervision meetings after first using AI feedback to “clean up” the writing, allowing the meeting itself to focus on content. At the same time, most participants expressed some worry about the accuracy of AI suggestions on discipline-specific content; P5 recalled an instance where the AI suggested

changing a term that was actually standard in the field, noting that without prior knowledge of the term, the suggestion could easily have been followed by mistake.

**Conative component.** Two behavioral patterns emerged. The first, direct application, was common for surface-level suggestions; P2 stated that grammar suggestions were simply accepted directly, with little felt need to double-check them. The second, cross-checking, was more common for content-level suggestions; P1 explained that anything concerning the framing of the research problem was always confirmed with the supervisor first, rather than trusted to the AI alone. P6, at an earlier stage of proposal writing, displayed a more uncritical pattern, applying most AI suggestions directly regardless of whether they concerned grammar or content — a pattern that stood out from the other five informants.

**External and internal cues.** Participants linked their trust in AI-generated feedback to both the characteristics of the feedback itself and their own personal factors. P3 described becoming more cautious after ChatGPT and Grammarly gave conflicting suggestions for the same sentence, an inconsistency that reduced confidence in either tool. Internally, participants with more prior experience using AI tools (P1, P5) described themselves as more selective and critical, while P6, with comparatively less experience, described a greater tendency to trust AI output by default.

## **Discussion**

The findings of this study can be understood through three complementary theoretical lenses. First, the cognitive, affective, and conative patterns observed align closely with Rosenberg and Hovland’s [11] Tripartite Attitude Model, which conceptualizes attitudes as composed of a cognitive component (beliefs and understanding), an affective component (feelings), and a conative component (behavioral tendencies). Participants’ functional understanding of how AI feedback works (cognitive), their mixture of confidence and doubt (affective), and their differing patterns of

direct application versus cross-checking (conative) illustrate how these three components operate together, rather than independently, in shaping students' overall perception of AI-generated feedback.

Second, consistent with Robbins and Judge's [12] general theory of perception, students' interpretation of AI-generated feedback was shaped not only by the feedback itself but by external cues (its speed, clarity, and consistency) and internal cues (students' own prior experience, digital literacy, and confidence). This explains why two students exposed to broadly similar AI tools such as P1 and P6 could arrive at markedly different levels of trust and critical engagement, since perception, as Robbins and Judge argue, is an active interpretive process rather than a passive reception of stimuli.

Third, examined through the lens of AI Literacy [10], the pattern of cross-checking and critical evaluation observed among P1, P3, P4, and P5 particularly regarding content-level suggestions can be read as an indicator of relatively higher AI literacy, in which AI-generated feedback is treated as one input to be weighed against the student's own judgment and the supervisor's guidance. In contrast, the pattern observed in P6, who tended to accept AI suggestions directly regardless of whether they concerned grammar or content, suggests comparatively lower AI literacy. This pattern notably co-occurred with an earlier proposal-writing stage and less prior experience with AI tools, echoing Ding et al.'s [4] concern about the risk of overdependence on AI among less experienced users. These findings extend Guo and Wang [5] and Dewi et al. [6] by showing that in the more complex and higher-stakes context of proposal/thesis writing, the distinction students draw between what AI feedback is, and is not, suited to handle appears to sharpen with AI consistently treated as a complement to, rather than a substitute for, supervisor guidance.

## CONCLUSION

Based on the findings of this study, several conclusions can be drawn regarding EFL students' perceptions of AI-generated feedback in academic writing. First, students generally perceive AI-generated feedback as a fast, accessible, and useful tool for surface-level issues such as grammar and vocabulary, which contributes to a stronger sense of confidence and preparedness before supervision meetings. Second, students' trust in and reliance on AI-generated feedback for content-level issues remain more cautious, shaped by both the consistency of the feedback itself and students' own prior experience and digital literacy. Finally, students with greater experience using AI tools tend to engage with AI-generated feedback more critically — cross-checking content-level suggestions with their supervisor while less experienced students show a greater tendency toward uncritical acceptance, suggesting that the proposal/thesis-writing stage and prior AI experience together shape how critically students engage with AI-generated feedback. Across all components, AI-generated feedback was consistently positioned by participants as a complement to, rather than a replacement for, supervisor guidance in the proposal and thesis writing process.

This study is subject to several limitations. It involved a small, purposively selected sample of 24 questionnaire respondents and six interview informants from a single study program at one university, so the findings reflect the perceptions of this specific group rather than EFL students in general. The self-reported nature of the questionnaire and interview data may also be subject to social desirability bias, and the cross-sectional design captured students' perceptions at a single point in their proposal/thesis-writing process rather than tracking how these perceptions might change over time. Future research is recommended to involve larger and more diverse samples across institutions, to employ longitudinal designs that capture changes in students' perceptions and AI literacy over the course of their thesis writing,

and to examine actual revision behavior alongside self-reported perceptions.

## 5. ACKNOWLEDGMENTS

The author would like to express sincere gratitude to the English Education Study Program, FBMB, Mandalika University of Education (UNDIKMA). Special thanks are extended to the thesis supervisors, Taufik Suadiyatno, M.Pd., and Imansyah, M.Pd., for their invaluable guidance, as well as to the students who participated in this study.

## BIBLIOGRAPHY

- [1] S. A. M. Asnas and I. Hidayanti, "Uncovering EFL students' frequent difficulties in academic writing and the coping strategies: The case of a college in Indonesia," *Journal on English as a Foreign Language*, vol. 14, no. 1, pp. 124–151, 2024.
- [2] J. Hattie and H. Timperley, "The power of feedback," *Review of Educational Research*, vol. 77, no. 1, pp. 81–112, 2007.
- [3] D. Carless and D. Boud, "The development of student feedback literacy: Enabling uptake of feedback," *Assessment & Evaluation in Higher Education*, vol. 43, no. 8, pp. 1315–1325, 2018.
- [4] L. Ding, D. Zou, and L. Kohnke, "ChatGPT as an automated writing evaluation tool: How students perceive it and how it affects their writing," *Education and Information Technologies*, vol. 30, pp. 26777–26799, 2025.
- [5] K. Guo and D. Wang, "To resist it or to embrace it? Examining ChatGPT's potential to support teacher feedback in EFL writing," *Education and Information Technologies*, vol. 29, no. 7, pp. 8435–8463, 2024.
- [6] U. Dewi, S. H. Daulay, and S. Ismahani, "AI feedback vs. peer feedback in EFL writing: Exploring Indonesian students' perceptions," *Asian-Pacific Journal of Second and Foreign Language Education*, vol. 11, Art. 49, 2026.
- [7] K. Hyland, *Second Language Writing*. Cambridge University Press, 2003.
- [8] J. W. Creswell and C. N. Poth, *Qualitative Inquiry and Research Design: Choosing Among Five Approaches*, 4th ed. SAGE Publications, 2018.
- [9] M. B. Miles, A. M. Huberman, and J. Saldaña, *Qualitative Data Analysis: A Methods Sourcebook*, 3rd ed. SAGE Publications, 2014.
- [10] J. Shi, W. Liu, and K. Hu, "Exploring how AI literacy and self-regulated learning relate to student writing performance and well-being in generative AI-supported higher education," *Behavioral Sciences*, vol. 15, no. 5, p. 705, 2025.
- [11] M. J. Rosenberg and C. I. Hovland, "Cognitive, affective, and behavioral components of attitudes," in *Attitude Organization and Change: An Analysis of Consistency Among Attitude Components*, C. I. Hovland and M. J. Rosenberg, Eds. Yale University Press, 1960, pp. 1–14.
- [12] S. P. Robbins and T. A. Judge, *Organizational Behavior*, 15th ed. Pearson, 2013.
- [13] M. B. Alghasab, "English as a foreign language (EFL) secondary school students' use of artificial intelligence (AI) tools for developing writing skills: Unveiling practices and perceptions," *Cogent Education*, vol. 12, no. 1, Art. 2505304, 2025.
- [14] F. D. Davis, "Perceived usefulness, perceived ease of use, and user acceptance of information technology," *MIS Quarterly*, vol. 13, no. 3, pp. 319–340, 1989.
- [15] W. A. Gerungan, *Psikologi Sosial*, 2nd ed. PT Refika Aditama, 1996.
- [16] Imansyah, I. M. P. Utama, and B. Sumarni, "An analysis on the EFL teachers' perception about authentic assessment," *VELES: Voices of English Language*

*Education Society*, vol. 2, no. 2, pp. 98–104, 2018.

- [17] A. Oshima and A. Hogue, *Introduction to Academic Writing*, 3rd ed. Pearson Longman, 2007.
- [18] I. P. Suartika, "Contextual teaching and learning in English in international standard senior high school in Denpasar: Teachers' perception and pedagogical practices the academic year 2007/2008," Unpublished thesis, Undiksha, Singaraja, 2009.