International Journal of Humanities, Education, and Social Sciences

e-ISSN : 3026-0892
p-ISSN : 3026-1422

Index: Harvard, Boston, Sydney, Dimensions, Lens, Scilit, Semantic, Google, etc

https://doi.org/10.58578/IJHESS.v4i3.12541

# Generative AI Use in Indonesian Language Learning: A Descriptive Survey of University Students

**Jafar**
Universitas Negeri Makassar, Indonesia
jafar@unm.ac.id



| Submitted: | Revised: | Accepted: | Published: |
|---|---|---|---|
| Sep 22, 2026 | Oct 21, 2026 | Nov 1, 2026 | Nov 6, 2026 |



**Abstract**

Despite considerable research attention to generative artificial intelligence (GenAI) in higher education, empirical evidence on its use in Indonesian language courses remains limited, particularly studies integrating usage, perceptions, opportunities, challenges, AI literacy, and ethics. This study aimed to describe these six dimensions among first-year university students. A quantitative descriptive survey involved 69 students from two parallel classes. Data were collected using a 30-item questionnaire (Cronbach's α = .752–.873) and analyzed using descriptive statistics, categorical interpretation, and supplementary nonparametric comparisons. The findings showed that 87.0% of students had used GenAI, primarily Gemini and ChatGPT. Perceptions (M = 3.79), opportunities (M = 3.70), and ethical awareness (M = 3.82) were high, whereas challenges (M = 3.37), AI literacy (M = 3.27), and usage intensity (M = 2.88) were moderate. Users scored higher than non-users on perceptions, opportunities, and AI literacy; however, among users, AI literacy was not associated with frequency or duration of use. These findings suggest that students' acceptance of GenAI exceeds their AI literacy, identifying a potential acceptance–competence gap in GenAI adoption. The study contributes a conceptual perspective on this gap and highlights the need for lecturers and institutions to integrate structured AI literacy and ethics instruction into Indonesian language courses. Further research should examine this gap using larger samples and performance-based measures.


IJHESS Journal is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 4.0 International License

**Keywords:** Generative Artificial Intelligence; Indonesian Language Learning; AI Literacy; AI Ethics; Student Perceptions

## INTRODUCTION

The rapid spread of generative artificial intelligence (GenAI), particularly large language model applications such as ChatGPT and Gemini, has changed the way students search for information, draft texts, and complete academic assignments. Because these tools produce fluent, human-like text on demand, they directly touch the core activities of language learning: reading, writing, revising, and responding to feedback (Kasneci et al., 2023; Kohnke et al., 2023). Reviews of the early literature describe GenAI as a technology that offers personalized support and immediate feedback while also raising serious concerns about accuracy, over-reliance, and academic integrity (Lo, 2023; Farrokhnia et al., 2024). International guidance has therefore called on higher education institutions to move beyond prohibition toward the responsible and pedagogically grounded use of GenAI (Miao & Holmes, 2023; Chan, 2023).

In Indonesia, the Indonesian language (Bahasa Indonesia) course is a compulsory general course that aims to develop students' academic literacy, especially their ability to write coherent, accurate, and ethically sourced academic texts. GenAI is particularly relevant to this course because its main outputs, written texts, are exactly the products students are assessed on. Recent Indonesian studies report that students already use AI to understand course material, improve assignments, and obtain faster feedback, while lecturers face problems of dependency and limited guidance (Tristianto et al., 2025; Nurfitri, 2026). The question is therefore no longer whether students use GenAI, but how they use it and whether they have the literacy and ethical awareness needed to use it well.

This study argues that the educational value of GenAI in language learning depends less on access to the technology than on students' AI literacy, namely the competencies that enable them to evaluate AI critically, communicate and collaborate with it effectively, and use it ethically (Long & Magerko, 2020; Ng et al., 2021). Technology acceptance theories explain that perceived usefulness and ease of use drive adoption (Davis, 1989; Venkatesh et al., 2003); however, acceptance does not guarantee competent use. If students accept GenAI enthusiastically but lack the literacy to evaluate its outputs, the technology may weaken rather

than strengthen the writing and critical thinking skills that the Indonesian language course is meant to develop (Barrot, 2023; Cotton et al., 2024). A descriptive mapping of usage, perceptions, and literacy is thus a necessary first step before designing instructional interventions.

Previous studies have examined students' views of GenAI from several angles. Large-scale surveys in higher education have documented positive perceptions alongside concerns about accuracy, privacy, and integrity (Chan & Hu, 2023), while acceptance studies have identified habit, performance expectancy, and hedonic motivation as predictors of use (Strzelecki, 2024). In language education, studies have focused mainly on English as a foreign language, showing that ChatGPT can support second language writing and argumentation but requires guided use (Barrot, 2023; Su et al., 2023; Yan, 2023; Puteri et al., 2025). Indonesian studies have reported high readiness and acceptance among university students (Martha et al., 2025; Amin et al., 2026; Saputra & Zuhriyah, 2026), yet most concern English learning, thesis writing, or general learning. Studies situated in the Indonesian language course are still largely conceptual or limited to general perceptions (Tristianto et al., 2025; Nurfitri, 2026). Few studies have profiled usage, perception, opportunities, challenges, AI literacy, and ethics simultaneously within the same group of students, which makes it difficult to see whether positive attitudes are matched by adequate competence.

The novelty of this study lies in three aspects: its context, the Indonesian language course as a first-language academic literacy course rather than a foreign language course; its multidimensional profile, which combines six dimensions of GenAI use in one instrument; and its analytical focus on the relationship between acceptance and competence, examined by comparing users with non-users and relating usage intensity to AI literacy. The study draws on the technology acceptance perspective (Davis, 1989; Venkatesh et al., 2003) to interpret perceptions and opportunities, and on the AI literacy framework (Long & Magerko, 2020; Ng et al., 2021) to interpret students' competence and ethical awareness.

Accordingly, this study aims to describe the use of GenAI in Indonesian language learning among first-year university students, specifically their usage patterns, perceptions, perceived opportunities, perceived challenges, AI literacy, and ethical awareness. It also aims to describe how these dimensions differ between students who have and have not used GenAI and how they vary with frequency and duration of use, so that the results can inform the design of GenAI-integrated Indonesian language instruction.

## METHODS

This study employed a quantitative approach with a descriptive survey design. A descriptive survey was chosen because the aim of the study was to portray the current state of a phenomenon, namely students' use of GenAI, in numerical form rather than to test causal relationships (Creswell & Creswell, 2018). The study was cross-sectional, with data collected at a single point in time.

The design consisted of four stages: (1) developing a questionnaire based on the literature on GenAI in education, technology acceptance, and AI literacy; (2) administering the questionnaire to students in the Indonesian language course; (3) testing the validity and reliability of the instrument using the collected data; and (4) analyzing the data descriptively, followed by supporting comparisons between subgroups. This sequence ensured alignment between the research objective (to describe GenAI use), the instrument (a multidimensional questionnaire), and the type of data collected (ordinal scale scores).

The participants were 69 first-year students enrolled in two parallel classes (D1 and D2) of the Indonesian language course at Universitas Negeri Makassar (UNM), Indonesia. Students from the two intact classes were invited to participate, and all 69 who completed the questionnaire were included in the analysis; no responses were excluded because there were no missing data. Of the participants, 36 (52.2%) were in class D1 and 33 (47.8%) in class D2; 43 (62.3%) were female, 25 (36.2%) male, and one (1.4%) preferred not to state their gender. Their ages ranged from 17 to 21 years (M = 18.54, SD = 0.83). Table 1 summarizes the profile of the respondents.

**Table 1. Profile of Respondents (N = 69)**

| Characteristic | Category | n | % |
|---|---|---|---|
| Class | D1 | 36 | 52.2 |
| | D2 | 33 | 47.8 |
| Gender | Female | 43 | 62.3 |
| | Male | 25 | 36.2 |
| | Prefer not to say | 1 | 1.4 |
| Age (years) | 17 | 6 | 8.7 |
| | 18 | 28 | 40.6 |
| | 19 | 28 | 40.6 |
| | 20–21 | 7 | 10.1 |

| | | | |
|---|---|---|---|
| Ever used GenAI | Yes | 60 | 87.0 |
| | No | 9 | 13.0 |
| Frequency of use | Never | 9 | 13.0 |
| | Rarely | 16 | 23.2 |
| | Several times a month | 11 | 15.9 |
| | Several times a week | 26 | 37.7 |
| | Daily | 7 | 10.1 |
| Duration of use | Never used | 9 | 13.0 |
| | Less than 3 months | 8 | 11.6 |
| | 3–6 months | 22 | 31.9 |
| | 7–12 months | 17 | 24.6 |
| | More than 1 year | 13 | 18.8 |
| Application used* | Gemini | 47 | 78.3 |
| | ChatGPT | 36 | 60.0 |
| | Other applications | 14 | 23.3 |
| | More than one application | 30 | 50.0 |

**Instruments and Data Collection**

Data were collected using a researcher-developed questionnaire consisting of two parts. The first part recorded respondents' characteristics and GenAI background (class, age, gender, experience of use, applications used, frequency, and duration of use). The second part contained 30 items grouped into six dimensions of five items each: usage pattern (P1–P5), perception (PS1–PS5), opportunities (PL1–PL5), challenges (T1–T5), AI literacy (LA1–LA5), and AI ethics (E1–E5). The usage pattern dimension measured the intensity of GenAI use for learning activities on a five-point frequency scale (1 = never, 2 = rarely, 3 = sometimes, 4 = often, 5 = always). The remaining five dimensions measured students' views of GenAI, its potential benefits for Indonesian language learning, the difficulties and risks it poses, students' self-assessed ability to understand and critically evaluate AI outputs, and their awareness of responsible use, on a five-point agreement scale (1 = strongly disagree to 5 = strongly agree). The dimensions were adapted from the literature on students' GenAI perceptions (Chan & Hu, 2023), technology acceptance (Venkatesh et al., 2003), AI literacy (Ng et al., 2021), and academic integrity (Cotton et al., 2024). The questionnaire was administered to students in both classes during regular sessions of the Indonesian language course at Universitas Negeri Makassar.

Item validity was examined using the corrected item–total correlation within each dimension, and reliability using Cronbach's alpha (Tavakol & Dennick, 2011). As shown in Table 2, all corrected item–total correlations ranged from .455 to .736, exceeding the commonly used criterion of .30, and all uncorrected item–total correlations (.649–.849) exceeded the critical value of $r = .237$ ($N = 69$, $\alpha = .05$). Cronbach's alpha ranged from .752 to .873, indicating acceptable to good internal consistency for every dimension. Removing any item did not increase the alpha of its dimension, so all 30 items were retained.

**Table 2. Validity and Reliability of the Instrument**

| Dimension | Items | Scale | Corrected item–total r | Cronbach's α |
|---|---|---|---|---|
| Usage pattern | 5 | Frequency | .686–.736 | .873 |
| Perception | 5 | Agreement | .537–.672 | .815 |
| Opportunities | 5 | Agreement | .546–.631 | .803 |
| Challenges | 5 | Agreement | .536–.734 | .818 |
| AI literacy | 5 | Agreement | .499–.657 | .802 |
| AI ethics | 5 | Agreement | .455–.646 | .752 |

Participation was voluntary. Students were informed of the purpose of the study, that their responses would not affect their course grades, and that the data would be reported anonymously. Informed consent was obtained before the questionnaire was completed, and students could decline or withdraw at any time without consequence. The survey was anonymous and posed minimal risk; no names or student identification numbers were collected, and respondents were identified only by codes (R001–R069).

## Data Analysis Technique

The data were analyzed using descriptive statistics, including frequencies, percentages, means (M), standard deviations (SD), and minimum and maximum scores, for each item and dimension. A dimension score for each respondent was calculated as the mean of the five items in that dimension. Mean scores were interpreted using five categories of equal width, with the interval width calculated using Equation (1):

$$i = (\text{highest score} - \text{lowest score}) / \text{number of categories} = (5 - 1) / 5 = 0.80 \quad (1)$$

The resulting categories were very low (1.00–1.80), low (1.81–2.60), moderate (2.61–3.40), high (3.41–4.20), and very high (4.21–5.00). The respondent achievement level (RAL) for each dimension was calculated as the percentage of the maximum possible score using Equation (2):

$$RAL = (M / 5) \times 100\% \quad (2)$$

To provide supporting evidence, dimension scores were compared between GenAI users (n = 60) and non-users (n = 9) using the Mann–Whitney U test, chosen because of the unequal and small subgroup sizes, with effect size $r = |z|/\sqrt{N}$. Differences between classes and between genders were examined using independent samples t-tests with Cohen's d, and associations between frequency or duration of use and dimension scores were examined using Spearman's rank correlation (ϱ). Significance was set at $\alpha = .05$. All analyses were conducted in Python 3 using the pandas and SciPy libraries.

## RESULTS

**Overall profile of GenAI use.** Most respondents (60 of 69; 87.0%) had used GenAI. Among users, Gemini was the most frequently used application (78.3%), followed by ChatGPT (60.0%), and half of the users (50.0%) used more than one application. The most common frequency of use was several times a week (37.7%), followed by rarely (23.2%), while only 10.1% used GenAI daily. Most users had used GenAI for three to twelve months (56.5% of all respondents), indicating that GenAI use among these students was widespread but relatively recent (Table 1).

**Profile across the six dimensions.** Table 3 and Figure 1 present the descriptive statistics for each dimension. Three dimensions fell into the high category: AI ethics (M = 3.82, SD = 0.61; RAL = 76.5%), perception (M = 3.79, SD = 0.64; RAL = 75.7%), and opportunities (M = 3.70, SD = 0.67; RAL = 74.0%). The other three fell into the moderate category: challenges (M = 3.37, SD = 0.70; RAL = 67.3%), AI literacy (M = 3.27, SD = 0.69; RAL = 65.3%), and usage pattern (M = 2.88, SD = 0.90; RAL = 57.6%). The proportion of respondents whose individual scores reached the high or very high category was 72.5% for AI ethics, 68.1% for perception, and 60.9% for opportunities, but only 37.7% for AI literacy and challenges and 29.0% for usage pattern.

**Table 3. Descriptive Statistics of GenAI Use Dimensions (N = 69)**

| Dimension | M | SD | Min–Max | RAL (%) | High–very high (%) | Category |
|---|---|---|---|---|---|---|
| AI ethics | 3.82 | 0.61 | 2.2–4.8 | 76.5 | 72.5 | High |
| Perception | 3.79 | 0.64 | 2.4–5.0 | 75.7 | 68.1 | High |
| Opportunities | 3.70 | 0.67 | 2.4–5.0 | 74.0 | 60.9 | High |
| Challenges | 3.37 | 0.70 | 2.0–4.8 | 67.3 | 37.7 | Moderate |
| AI literacy | 3.27 | 0.69 | 1.8–5.0 | 65.3 | 37.7 | Moderate |
| Usage pattern | 2.88 | 0.90 | 1.0–4.4 | 57.6 | 29.0 | Moderate |

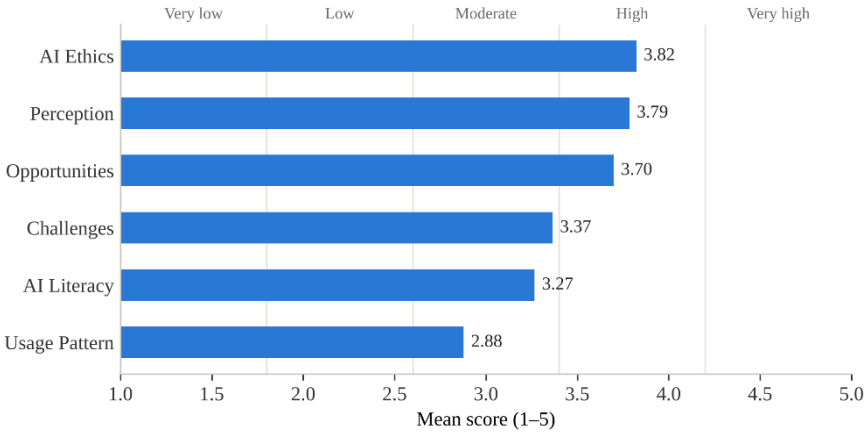



**Figure 1. Mean Scores of the Six Dimensions of GenAI Use**

**Item-level results.** Table 4 presents the item means. In the perception dimension, PS5 obtained the highest mean of all 30 items (M = 4.03), with 78.3% of respondents agreeing or strongly agreeing. In the ethics dimension, E4 (M = 3.99) and E5 (M = 3.96) were the highest, and 76.8% and 73.9% of respondents, respectively, agreed with them. In contrast, all five AI literacy items fell into the moderate category (M = 3.16–3.38); for LA1, the lowest literacy item, only 34.7% of respondents agreed, while 43.5% chose the neutral option. The lowest item overall was P1 in the usage pattern dimension (M = 2.59), which 52.2% of respondents reported never or rarely doing. In the challenges dimension, T3 (M = 3.46), T4 (M = 3.41), and T5 (M = 3.51) reached the high category, whereas T1 (M = 3.17) and T2 (M = 3.28) remained moderate.

**Table 4. Item Means and Standard Deviations by Dimension**

| Dimension | Item 1 | Item 2 | Item 3 | Item 4 | Item 5 |
|---|---|---|---|---|---|
| Usage pattern | 2.59 (1.14) | 2.83 (1.15) | 2.91 (1.01) | 2.97 (1.07) | 3.09 (1.16) |
| Perception | 3.71 (0.75) | 3.62 (0.99) | 3.71 (0.86) | 3.86 (0.90) | 4.03 (0.73) |
| Opportunities | 3.61 (0.96) | 3.59 (0.93) | 3.78 (0.87) | 3.64 (0.84) | 3.87 (0.86) |
| Challenges | 3.17 (0.94) | 3.28 (0.97) | 3.46 (0.85) | 3.41 (0.88) | 3.51 (0.95) |
| AI literacy | 3.16 (0.95) | 3.22 (0.92) | 3.26 (0.90) | 3.32 (0.80) | 3.38 (1.04) |
| AI ethics | 3.65 (0.85) | 3.70 (0.86) | 3.83 (0.94) | 3.99 (0.80) | 3.96 (0.85) |

**Users and non-users.** As shown in Table 5 and Figure 2, GenAI users scored significantly higher than non-users on usage pattern ($U = 539.0$, $p < .001$, $r = .58$), opportunities ($U = 453.5$, $p = .001$, $r = .39$), perception ($U = 422.0$, $p = .007$, $r = .33$), and AI literacy ($U = 400.0$, $p = .020$, $r = .28$). The difference in AI ethics (M = 3.87 vs. 3.51) was

not significant (p = .079). Non-users rated challenges higher than users (M = 3.62 vs. 3.33), but this difference was also not significant (p = .355).

**Table 5. Comparison of Dimension Scores between GenAI Users and Non-users**

| Dimension | Users M (SD) | Non-users M (SD) | U | z | p | r |
|---|---|---|---|---|---|---|
| Usage pattern | 3.11 (0.71) | 1.31 (0.25) | 539.0 | 4.79 | < .001 | .58 |
| Perception | 3.86 (0.64) | 3.27 (0.45) | 422.0 | 2.71 | .007 | .33 |
| Opportunities | 3.80 (0.64) | 3.02 (0.42) | 453.5 | 3.27 | .001 | .39 |
| Challenges | 3.33 (0.69) | 3.62 (0.72) | 218.0 | −0.93 | .355 | .11 |
| AI literacy | 3.34 (0.69) | 2.80 (0.48) | 400.0 | 2.32 | .020 | .28 |
| AI ethics | 3.87 (0.61) | 3.51 (0.50) | 368.5 | 1.76 | .079 | .21 |

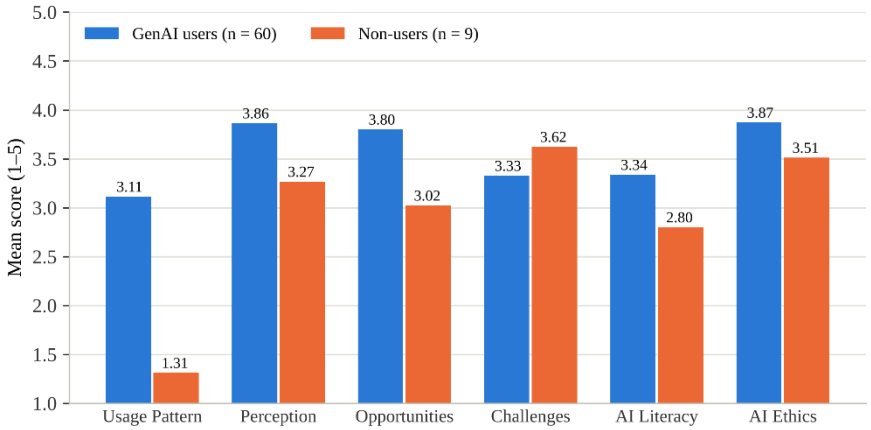



**Figure 2. Mean Dimension Scores of GenAI Users and Non-users**

**Frequency and duration of use.** Across all respondents, frequency of use correlated positively with usage pattern (ϱ = .34, p = .004), perception (ϱ = .38, p = .001), and opportunities (ϱ = .31, p = .009), but not with AI literacy (ϱ = .03, p = .788). When the analysis was restricted to users, these associations largely disappeared: frequency was no longer related to opportunities (ϱ = .09, p = .493) or AI literacy (ϱ = −.19, p = .137), its association with perception was borderline (ϱ = .25, p = .050), and it was negatively related to AI ethics (ϱ = −.27, p = .040). Duration of use was not significantly related to any dimension among users (all $|\varrho| \leq .16$, p > .21).

**Class, gender, and relations between dimensions.** No significant differences were found between classes D1 and D2 on any dimension (all p > .27). Male students reported a higher usage pattern score than female students (M = 3.19 vs. 2.70; t(66) = 2.21, p = .031, d = 0.55), but no gender differences were found on the other five dimensions (all p > .26). Correlations between dimensions were generally weak; the only significant

correlations were between usage pattern and opportunities ($r = .37$, $p = .002$) and between AI literacy and opportunities ($r = .24$, $p = .048$). Perception was not correlated with AI literacy ($r = .05$).

Three findings did not follow the general pattern. First, six of the nine students who reported never having used GenAI gave usage pattern ratings above the minimum on at least one item (dimension $M = 1.31$). Because this may reflect response inconsistency or indirect exposure, the usage pattern dimension was also computed for users only ($M = 3.11$, $SD = 0.71$), which still falls into the moderate category. Second, the gender difference in usage pattern was not robust: eight of the nine non-users were female, and when only users were analyzed, the difference between male and female students was not significant ($M = 3.28$ vs. $3.01$, $p = .149$). Third, daily users ($n = 7$) had the highest perception score ($M = 4.20$) but the lowest AI ethics score of all frequency groups ($M = 3.43$), compared with 3.86–4.06 among other user groups.

## DISCUSSION

### 1. Analysis and Interpretation of Findings

The main finding of this study is an imbalance between students' acceptance of GenAI and their competence in using it. Students' perceptions of GenAI, the opportunities they see in it, and their stated ethical awareness were high, whereas their AI literacy and actual intensity of use were only moderate. This pattern suggests that, in the Indonesian language course, GenAI is already accepted as a legitimate learning aid, but students do not yet feel fully capable of understanding how it works or of evaluating its outputs critically. We refer to this pattern as an acceptance–competence gap. The weak correlation between perception and AI literacy ($r = .05$) supports this interpretation: students who view GenAI positively are not necessarily those who feel more literate in using it.

The supporting findings clarify the source of this gap. Experience with GenAI made a clear difference: users had more positive perceptions, saw more opportunities, and reported higher AI literacy than non-users. However, among users, neither how often nor how long they had used GenAI was associated with higher literacy. In other words, the initial step from non-use to use appears to raise students' confidence, but further use does not by itself deepen their literacy. The negative association between frequency of use and ethical awareness

among users, together with the lowest ethics score among daily users, indicates that intensive use without guidance may be accompanied by weaker attention to responsible use.

**2. Comparison with Theories and Previous Literature**

The high perception and opportunity scores are consistent with technology acceptance theory, which holds that perceived usefulness is a central driver of technology adoption (Davis, 1989; Venkatesh et al., 2003), and with previous surveys reporting positive student attitudes toward GenAI in higher education (Chan & Hu, 2023; Strzelecki, 2024). They also agree with Indonesian studies showing high readiness and acceptance of ChatGPT among university students (Martha et al., 2025; Amin et al., 2026) and positive views of its role in generating ideas and organizing writing (Saputra & Zuhriyah, 2026). The dominance of Gemini over ChatGPT in this study, however, differs from most earlier studies, which focused almost exclusively on ChatGPT; this difference may reflect the easy availability of Gemini on the Android devices and Google accounts widely used by Indonesian students.

The moderate AI literacy scores, and the absence of a relationship between usage intensity and literacy, support the view that AI literacy is a set of competencies that must be learned rather than a by-product of exposure (Long & Magerko, 2020; Ng et al., 2021). This is in line with O'Dea et al. (2024), who found that university students' GenAI literacy is shaped by factors beyond frequency of use. It is also consistent with language education studies showing that GenAI benefits writing only when students are guided to evaluate and revise its outputs (Barrot, 2023; Su et al., 2023; Yan, 2023). Puteri et al. (2025), for example, found that Indonesian EFL students perceived ChatGPT positively, yet their essay quality did not improve significantly more than that of students taught conventionally. The present findings extend this evidence to the Indonesian language course: positive perceptions do not automatically translate into literate use.

The moderate challenge scores suggest that students recognize the risks of GenAI, such as inaccurate outputs and over-reliance, but do not regard them as a major barrier. This is consistent with reviews that describe GenAI as offering strengths and opportunities alongside weaknesses and threats (Farrokhnia et al., 2024; Lo, 2023; Tlili et al., 2023) and with reports of dependency as a key challenge in Indonesian language classes (Tristianto et al., 2025). The high self-reported ethical awareness contrasts with concerns raised in the literature about academic integrity in the GenAI era (Cotton et al., 2024; Rudolph et al., 2023). Because ethics was measured as stated awareness rather than observed behavior, and

because the most frequent users reported the lowest ethical awareness, the high ethics score should be read as a willingness to act responsibly that still needs to be supported by clear rules and practice (Chan, 2023; Kasneci et al., 2023).

### 3. Research Implications

Theoretically, this study contributes the concept of an acceptance–competence gap to the literature on GenAI adoption in language education. It shows that acceptance models explain why students adopt GenAI, but that a separate literacy perspective is needed to explain whether they use it well. Practically, the findings indicate that Indonesian language lecturers should not assume that students who use GenAI frequently are competent users. Lecturers can integrate explicit AI literacy instruction into writing tasks, for example by asking students to compare AI-generated drafts with their own, to verify facts and sources, and to document how AI was used. Process-based assessment, in which drafts and revision notes are evaluated, can reduce the risk of uncritical copying. At the institutional level, the findings support the development of clear GenAI guidelines that define acceptable use, disclosure, and assessment rules for general courses (Chan, 2023; Miao & Holmes, 2023; Zawacki-Richter et al., 2019).

### 4. Research Limitations and Directions for Future Research

This study has several limitations. First, the sample was small (N = 69) and drawn from two classes at one institution, which limits the generalizability of the findings. Second, the non-user group was very small (n = 9), so the comparisons between users and non-users should be interpreted with caution. Third, all dimensions, including AI literacy and ethics, were measured by self-report, which may differ from actual competence and behavior. Fourth, the instrument was developed for this study, and although its reliability was acceptable, its factor structure has not been confirmed. Finally, the cross-sectional design does not allow causal conclusions. Future studies are encouraged to use larger multi-institutional samples, to confirm the instrument through confirmatory factor analysis, to measure AI literacy with performance-based tasks, and to combine surveys with interviews or analyses of students' writing to explain how GenAI is used in Indonesian language learning.

## CONCLUSION

This study concludes that GenAI has become a common part of Indonesian language learning among first-year university students, but that students' acceptance of the technology has outpaced their competence in using it. Students hold positive perceptions of GenAI, recognize its opportunities, and express ethical awareness, yet their AI literacy and intensity of use remain moderate. Using GenAI distinguishes students from non-users in perception, opportunities, and AI literacy, but more frequent or longer use is not accompanied by higher literacy and is associated with lower ethical awareness among the most frequent users.

The study contributes to knowledge by providing a multidimensional profile of GenAI use in a first-language academic literacy course and by identifying an acceptance–competence gap that technology acceptance models alone do not capture. This contribution is limited to the context of the participating classes, but it offers a basis for designing GenAI-integrated instruction in Indonesian language courses. Lecturers are recommended to embed structured AI literacy and ethics activities into writing instruction, and institutions are recommended to provide clear guidelines on GenAI use and disclosure. Future research should examine whether such instruction narrows the acceptance–competence gap, using experimental or longitudinal designs, larger samples, and performance-based measures of AI literacy.

## REFERENCES

Amin, M., Mustaqim, B., Pratama, W., & Sibuea, A. M. (2026). Assessing student perspectives on ChatGPT in higher education: A quantitative analysis. *IAES International Journal of Artificial Intelligence, 15*(2), 1062–1070. https://doi.org/10.11591/ijai.v15.i2.pp1062-1070

Barrot, J. S. (2023). Using ChatGPT for second language writing: Pitfalls and potentials. *Assessing Writing, 57*, Article 100745. https://doi.org/10.1016/j.asw.2023.100745

Chan, C. K. Y. (2023). A comprehensive AI policy education framework for university teaching and learning. *International Journal of Educational Technology in Higher Education, 20*, Article 38. https://doi.org/10.1186/s41239-023-00408-3

Chan, C. K. Y., & Hu, W. (2023). Students' voices on generative AI: Perceptions, benefits, and challenges in higher education. *International Journal of Educational Technology in Higher Education, 20*, Article 43. https://doi.org/10.1186/s41239-023-00411-8

Cotton, D. R. E., Cotton, P. A., & Shipway, J. R. (2024). Chatting and cheating: Ensuring academic integrity in the era of ChatGPT. *Innovations in Education and Teaching International, 61*(2), 228–239. https://doi.org/10.1080/14703297.2023.2190148

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications. https://www.sagepub.com/shop/buy-a-book/research-design-5-255675

Davis, F. D. (1989). Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS Quarterly, 13*(3), 319–340. https://doi.org/10.2307/249008

Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International, 61*(3), 460–474. https://doi.org/10.1080/14703297.2023.2195846

Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günnemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., … Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences, 103*, Article 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kohnke, L., Moorhouse, B. L., & Zou, D. (2023). ChatGPT for language teaching and learning. *RELC Journal, 54*(2), 537–550. https://doi.org/10.1177/00336882231162868

Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences, 13*(4), Article 410. https://doi.org/10.3390/educsci13040410

Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. In *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems* (pp. 1–16). Association for Computing Machinery. https://doi.org/10.1145/3313831.3376727

Martha, A. S. D., Widowati, S., & Rahayu, D. P. (2025). Assessing student readiness and perceptions of ChatGPT in learning: A case study in Indonesian higher education. *Journal of Information Technology Education: Research, 24*, Article 23. https://doi.org/10.28945/5599

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO. https://doi.org/10.54675/EWZM9535

Ng, D. T. K., Leung, J. K. L., Chu, S. K. W., & Qiao, M. S. (2021). Conceptualizing AI literacy: An exploratory review. *Computers and Education: Artificial Intelligence, 2*, Article 100041. https://doi.org/10.1016/j.caeai.2021.100041

Nurfitri, N. (2026). Pembelajaran Bahasa Indonesia Berbasis Kecerdasan Buatan untuk Meningkatkan Literasi Akademik Mahasiswa [Artificial intelligence-based Indonesian language learning to improve students' academic literacy]. *Ameena Journal, 4*(3), 624–633. https://doi.org/10.63732/aij.v4i3.303

O'Dea, X., Ng, D. T. K., O'Dea, M., & Shkuratskyy, V. (2026). Factors affecting university students' generative AI literacy: Evidence and evaluation in the UK and Hong Kong contexts. *Policy Futures in Education, 24*(1), 13–34. https://doi.org/10.1177/14782103241287401

Puteri, C. G., Cahyono, B. Y., Widiati, U., & Suryati, N. (2025). Using ChatGPT as a writing assistant: A study on essay quality development among Indonesian EFL university

students. *Indonesian Journal of Applied Linguistics, 15*(2), 332–346. https://doi.org/10.17509/3ebmy158

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), 342–363. https://doi.org/10.37074/jalt.2023.6.1.9

Saputra, N. R. A., & Zuhriyah, M. (2026). Students' perception of ChatGPT use in thesis proposals writing at a pesantren-based university. *Forum Ilmu Sosial, 53*(1), 45–52. https://doi.org/10.15294/fis.v53i01.48787

Strzelecki, A. (2024). To use or not to use ChatGPT in higher education? A study of students' acceptance and use of technology. *Interactive Learning Environments, 32*(9), 5142–5155. https://doi.org/10.1080/10494820.2023.2209881

Su, Y., Lin, Y., & Lai, C. (2023). Collaborating with ChatGPT in argumentative writing classrooms. *Assessing Writing, 57*, Article 100752. https://doi.org/10.1016/j.asw.2023.100752

Tavakol, M., & Dennick, R. (2011). Making sense of Cronbach's alpha. *International Journal of Medical Education, 2*, 53–55. https://doi.org/10.5116/ijme.4dfb.8dfd

Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments, 10*, Article 15. https://doi.org/10.1186/s40561-023-00237-x

Tristianto, M. R., Nugraha, A. S., Ramdani, A., & Nababan, A. S. (2025). Pengaruh Artificial Intelligence (AI) dalam Evaluasi Pembelajaran Bahasa Indonesia pada Kalangan Mahasiswa [The influence of artificial intelligence (AI) on the evaluation of Indonesian language learning among university students]. *Jurnal Nakula: Pusat Ilmu Pendidikan, Bahasa dan Ilmu Sosial, 3*(3), 47–65. https://doi.org/10.61132/nakula.v3i3.1766

Venkatesh, V., Morris, M. G., Davis, G. B., & Davis, F. D. (2003). User acceptance of information technology: Toward a unified view. *MIS Quarterly, 27*(3), 425–478. https://doi.org/10.2307/30036540

Yan, D. (2023). Impact of ChatGPT on learners in a L2 writing practicum: An exploratory investigation. *Education and Information Technologies, 28*(11), 13943–13967. https://doi.org/10.1007/s10639-023-11742-4

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education—Where are the educators? *International Journal of Educational Technology in Higher Education, 16*, Article 39. https://doi.org/10.1186/s41239-019-0171-0