p-ISSN 2252-7427, e-ISSN 2541-3635
https://journal.uniku.ac.id/index.php/IEFLJ/index

p-ISSN 2252-7427 e-ISSN 2541-3635
Indonesian EFL Journal
AISEE
The Association of Indonesia Scholars of English Education

SCAFFOLD OR CRUTCH? EFL LEARNERS' MULTIMODAL GENAI EXPERIENCES IN INDONESIA

Yuliana Friska[1]*, Angga Hidayat[2], Ani Kusumaningsih[3], Moh. Thoyib Syafi'i[4], Nurmalia[5]
[1,3,4,5]Universitas Pamulang, Banten, Indonesia
[2]Florida Gulf Coast University, Fort Myers, USA

Corresponding author's email: dosen01381@unpam.ac.id

APA Citation: Friska, Y., Hidayat, A., Kusumaningsih, A., Syafi'I, M. T., & Nurmalia. (2026). Scaffold or crutch? EFL learners' multimodal GenAI experiences in Indonesia. *Indonesian EFL Journal*, *12*(3), 769-780. https://doi.org/10.25134/ieflj.v12i3.150

Received: 04-03-2026 Accepted: 18-05-2026 Published: 27-08-2026

Abstract: Despite the rapid integration of generative AI (GenAI) into EFL learning, most existing research has relied on single-method survey designs that capture student attitudes without explaining how learners actually engage with these tools. This study employed an explanatory sequential mixed-methods design to investigate EFL students' perceptions of multimodal GenAI's pedagogical value and explain those perceptions through their learning experiences. Guided by Sociocultural Theory, Multimodal Learning Theory, and Self-Determination Theory, a five-point Likert-scale survey was administered to 183 undergraduate EFL students across five Indonesian higher education institutions, followed by semi-structured interviews with 13 purposefully selected participants representing three distinct perception profiles. Quantitative data were analysed using descriptive statistics and exploratory factor analysis in R, while qualitative data were analysed through reflexive thematic analysis. Findings showed moderately positive perceptions of GenAI's pedagogical value (M = 3.62, SD = 0.75), with critical thinking concern as the most salient risk (M = 3.90, SD = 1.15) and strong demand for lecturer guidance on ethical AI use (M = 3.72, SD = 1.20). Four themes were identified: GenAI as a language empowerment tool, balancing support and overreliance, enhancing engagement through multimodality, and shaping teacher roles. These findings point to the need for AI literacy curricula that develop students' critical, ethical, and self-regulatory capacities in Indonesian higher education contexts.
Keywords: *EFL learning; generative AI; multimodal learning; learner agency; mixed-methods*

INTRODUCTION

Generative artificial intelligence (GenAI) has substantially changed how EFL students engage with English since ChatGPT's public release in late 2022. Unlike earlier text-only tools, modern GenAI platforms such as ChatGPT, Google Gemini, and Microsoft Copilot generate text, images, audio, and spoken language within a single interface (Godwin-Jones, 2023), creating learning conditions earlier technology-enhanced language learning research did not anticipate. A richer theoretical and empirical account is therefore needed, addressing not only how learners benefit from these tools but how they navigate the accompanying risks.

Ci and Jiang (2025) argued that the integration of GenAI into language teaching is no longer peripheral but integral to preparing learners for contemporary communicative demands, its multimodal affordances are now constitutive of what literacy practice requires.

Chen et al. (2020) noted that educational technology research has consistently shifted toward learner-centred outcomes, a trend evident in the growing body of GenAI studies in EFL contexts. Liu and Ma (2023) found that among 405 EFL learners, perceived ease of use significantly predicted perceived usefulness, which shaped students' intention to use ChatGPT beyond the classroom, while Vo and Nguyen (2024) and Alsaedi (2024) reported positive learner attitudes, though Alsaedi (2024) also documented concerns about academic dishonesty and reduced interpersonal skills. Across these studies, learners generally viewed GenAI as beneficial for writing development, confidence building, and

engagement (Almusharraf et al., 2025; Baidoo-Anu & Ansah, 2023; Chan, 2023; Guo & Li, 2024; Huang & Mizumoto, 2024; Kohnke et al., 2023). What remains less examined is how learners actually navigate these tools and whether positive attitudes translate into productive, self-regulated use.

Despite these perceived benefits, several concerns persist. The most theoretically consequential is overreliance. Chiu (2024) observed that uncritical adoption risks undermining independent thinking and higher-order skills, since students who depend excessively on GenAI forgo opportunities to develop their own linguistic and cognitive resources (Lantolf et al., 2020; Selwyn, 2024). Academic integrity presents an equally pressing challenge, as many students remain uncertain about acceptable boundaries (Chan, 2023; Yeo, 2023). Yan et al. (2024), analysing metaphors used by 281 Chinese EFL learners, found students simultaneously viewed GenAI as a transformative resource and a source of intellectual dependency. In Indonesian context, Darwin et al. (2024) documented similarly mixed views among EFL students. A review by Wang et al. (2025) confirmed that although 41.9% of GenAI studies in language education examined user perceptions, the tools' multimodal capabilities and learners' lived experiences remained substantially underexplored.

Several gaps organise this synthesis. Methodologically, survey-based designs document what students think but not how they use these tools, a limitation Lee et al. (2025) confirmed across 49 studies (2023–2024), most relying on self-report alone. Contextually, Indonesia remains underresearched, characterised by high linguistic diversity, uneven digital access, and limited institutional policy, distinguishing it from the East Asian settings dominating existing literature (Darwin et al., 2024). Arifin et al. (2024) similarly found that affordances and constraints were experienced unevenly across Indonesian institutional contexts, a pattern Cong-Lem's (2025) equity-focused review situates within broader disparities in GenAI access. In terms of skill coverage, most studies concentrate on writing, while speaking, listening, and reading remain comparatively neglected (Li et al., 2025; Lee, 2022; Sim et al., 2025), and multimodal capabilities beyond text, this study's focus, remain even less examined despite growing theoretical interest (Zapata, 2026). Finally, no study has combined quantitative and qualitative data within a theoretically grounded design to capture both the scope and depth of EFL learners' experiences with multimodal GenAI, a gap persisting even as reviews of teacher-facing use (Nazim & Alzubi, 2025) and general attitudes (Wu et al., 2025) continue to expand. Jeon (2022) substantiated this premise empirically, showing that learners who exercised greater control over AI-supported tasks engaged more deeply and persisted longer.

Three research questions guided the study: (RQ1) What are EFL students' perceptions of multimodal GenAI's pedagogical value? (RQ2) How do EFL students describe their multimodal GenAI learning practices? (RQ3) How do EFL students negotiate the benefits and risks of multimodal GenAI use?

METHOD

This study employed an explanatory sequential mixed-methods design (Creswell, 2022), quantitative data established general patterns in students' perceptions of multimodal GenAI, and qualitative interviews then explained those patterns through students' learning experiences and decision-making (Tashakkori et al., 2021). Integration occurred at the interpretation stage. Interview probes were derived from quantitative patterns, and qualitative findings elaborated the survey results (Fetters & Tajima, 2022).

The quantitative phase involved 183 undergraduate EFL students from five Indonesian higher education institutions, selected to represent variation in institutional type and location (Atmojo & Nugroho, 2020), with eligibility limited to students with prior GenAI experience for English learning. For the qualitative phase, 13 participants were selected through purposeful maximum variation sampling (Nyimbili & Nyimbili, 2024). Three perception profiles, positive, mixed/neutral, and critical/sceptical, were based on composite scores on perceived pedagogical value and risk/ethics items, following a three-band Likert classification (Low: 1.00–2.33; Moderate: 2.34–3.67; High: 3.68–5.00). Five from the positive group, four from each remaining group. Five of the 13 were concurrently pre-service English teachers, whose perspectives are reported separately.

Table 1. *Participant characteristics*

Variable	Category	n	%
Gender	Female	129	70.5
	Male	54	29.5
Prior GenAI Experience	Yes	148	80.9
	No	35	19.1

p-ISSN 2252-7427, e-ISSN 2541-3635
https://journal.uniku.ac.id/index.php/IEFLJ/index

Weekly GenAI Use	< 1 hour	89	48.6
	1-3 hours	62	33.9
	>= 4 hours	32	17.5
Primary Tool	ChatGPT only	74	40.4
	ChatGPT + others	74	40.4
	Other tools only	35	19.1

The quantitative instrument was a researcher-developed five-point Likert-scale questionnaire with ten items across three groupings: Perceived Pedagogical Value (4 items), Perceived Risks (3 items), and Ethics and Guidance Needs (3 items), adapted from existing GenAI perception instruments (Liu & Ma, 2023; Vo & Nguyen, 2024; Chan, 2023; Yeo, 2023) and reduced from 18 to 10 items following expert review, pilot testing, and cognitive interviews confirming item clarity. Only the Pedagogical Value subscale achieved acceptable reliability (α = .80; Hayes & Coutts, 2020); the other groupings yielded lower coefficients (α = .28; α = .49) due to short scale length and item heterogeneity (Sijtsma & Pfadt, 2021) and are reported at item level.

The qualitative instrument was a semi-structured interview guide covering actual GenAI practices, perceived benefits, risk management, and ethical reflections, conducted in Bahasa Indonesia to avoid language proficiency confounding responses (McKenna, 2022), audio-recorded, transcribed, and translated into English with back-translation verified by two lecturers.

Quantitative data were analyzed in R (R Core Team, 2023) using descriptive statistics, Cronbach's alpha, and EFA, with interview probes derived from salient quantitative patterns to explain rather than replicate survey findings. Qualitative data followed Braun and Clarke's (2021) reflexive thematic analysis, with the first author's positionality treated as an interpretive resource and reflexivity maintained through a research journal and peer debriefing. Rigour was ensured through audit trails and member-checking with all 13 participants. Integration was achieved through a joint display mapping quantitative findings, qualitative themes, and theoretical lenses (Fetters & Tajima, 2022).

RESULTS AND DISCUSSION

Instrument validity and reliability

The instrument's psychometric properties are reported before substantive findings for interpretive transparency. Composite scores are reported only for the Perceived Instructional Benefit factor (F1), which demonstrated adequate internal consistency. Items in the remaining groupings are presented at item level only.

Table 2. *Descriptive statistics of perception items and factor reliability*

Item	Grouping	M	SD	95% CI	FL
Improves English writing	Ped. Value	3.68	0.92	[3.55, 3.81]	.71
Helps understand feedback	Ped. Value	3.66	0.97	[3.52, 3.80]	.68
Increases writing confidence	Ped. Value	3.58	0.95	[3.44, 3.72]	.65
Makes learning more enjoyable	Ped. Value	3.57	0.95	[3.43, 3.71]	.62
Reduces critical thinking[1]	Risk	3.90	1.15	[3.73, 4.07]	.58
Too dependent on GenAI[1]	Risk	2.62	1.01	[2.47, 2.77]	.61
Used GenAI without permission[1]	Risk	2.75	1.14	[2.58, 2.92]	.55
Unsure about ethical acceptability[1]	Ethics	3.46	0.88	[3.33, 3.59]	.52
Need guidance from lecturers[1]	Ethics	3.72	1.20	[3.55, 3.89]	.49
GenAI as a replacement for teacher feedback[2]	Cross-load	2.33	1.26	[2.15, 2.51]	n/a
Perceived Instructional Benefit (α = .80)		3.62	0.75	[3.51, 3.73]	

Construct validity was examined through EFA. Ninety-five percent confidence intervals in Table 2 were calculated as $M \pm 1.96(SD/\sqrt{n})$, with n = 183 for all items, as the raw response file contained no missing data. The KMO index indicated good sampling adequacy (.77), and Bartlett's test of sphericity was significant ($\chi^2(45) = 410.39$, $p < .001$), confirming appropriateness for factor analysis (Goretzko et al., 2021). Parallel analysis and scree plot inspection supported a three-factor solution accounting for 41.8% of total variance, consistent with expectations for exploratory instruments (Goretzko et al., 2021). Table 3 presents the full pattern matrix; loadings below .30 are suppressed.

Table 3. *Exploratory factor analysis pattern matrix*

Item	F1	F2	F3
Improves English writing	.71		
Helps understand feedback	.68		
Increases writing confidence	.65		
Makes learning more enjoyable	.62		
Unsure about ethical acceptability		.61	

Reduces critical thinking		.58	
Need guidance from lecturers		.52	
Too dependent on GenAI			.61
Used GenAI without permission			.55
GenAI as a replacement for teacher feedback	.37	.33	.35
Eigenvalue	2.41	1.53	1.22
% Variance	17.8	13.4	10.6

EFA identified three factors departing in part from the a priori groupings. Factor 1 captured four items converging on direct instructional benefit (writing, feedback comprehension, confidence, enjoyment), labelled Perceived Instructional Benefit. Factor 2 grouped items reflecting ethical uncertainty, critical-thinking concern, and the need for lecturer guidance, suggesting an Ethical and Cognitive Anxiety cluster. Factor 3 comprised two behavioural items on self-reported dependency and unauthorised use, pointing to Behavioural Overreliance. The item on GenAI as a teacher-feedback replacement cross-loaded across all three factors (F1 = .37, F2 = .33, F3 = .35) and is excluded from factor-based interpretation, with its item-level mean reported separately. Alpha for F2 and F3 was markedly lower (α = .28 and α = .49), predictable given short item count and cross-domain content, further justifying the qualitative phase as the primary vehicle for examining risk.

EFL students' perceptions of multimodal GenAI (RQ1)

Descriptive statistics revealed a moderately positive orientation toward the pedagogical value of multimodal GenAI, alongside moderate risk awareness and notable ethical uncertainty. Students agreed that GenAI improved their English writing (M = 3.68, SD = 0.92), helped them understand feedback (M = 3.66, SD = 0.97), increased writing confidence (M = 3.58, SD = 0.95), and made learning more enjoyable (M = 3.57, SD = 0.95). GenAI as a replacement for teacher feedback returned the instrument's lowest mean (M = 2.33, SD = 1.26), suggesting students viewed it as supplementary support rather than a substitute for human instruction.

Item-level analysis showed critical thinking concern received the highest mean of all ten items (M = 3.90, SD = 1.15), identifying cognitive dependency as the most prominent risk, followed by the demand for instructor guidance on ethical GenAI use (M = 3.72, SD = 1.20). Prior GenAI experience was widespread (80.9%), with ChatGPT the primary tool, used most for translating sentences (63.4%), checking grammar (49.2%), and generating writing ideas (37.2%), mostly for fewer than one hour weekly.

One discrepancy merits close attention. Students reported strong concern about GenAI reducing their critical thinking (M = 3.90) while simultaneously reporting relatively low self-perceived overreliance (M = 2.62). This gap suggests learners may recognise cognitive offloading as a general risk without identifying themselves as its subjects, consistent with the optimistic self-assessment bias documented in the self-regulation literature (Hofer et al., 2022). The qualitative phase was designed to examine whether this discrepancy holds at the level of individual experience.

Learning experiences and practices (RQ2-RQ3)

Reflexive thematic analysis of the 13 semi-structured interviews yielded four overarching themes; Table 4 summarises themes, sub-themes, and representative quotes, with demographic characteristics and perception profiles for each participant cited (P1–P13) reported in Appendix B.

Table 4. *Qualitative findings from semi-structured interviews*

Theme	Sub-theme	Representative Quote
1. GenAI as Language Empowerment Tool	Multimodal scaffolding	"If there is a difficult word, I can ask ChatGPT for its synonym, and then I also look at the context in the image created by AI." (P3)
	Mediated skill development	"I asked ChatGPT for a script, then recorded my voice to check pronunciation." (P7)
2. Support vs. Overreliance	Automation complacency	"If I use ChatGPT too often, I feel lazy to think. All the answers are already there." (P2)
	Strategic self-regulation	"Usually, I ask for an example text, but I change the sentences so they match my style." (P9)
3. Engagement Through Multimodality	Affective engagement	"If there are pictures and videos, I feel more excited to learn. It's like playing a game." (P5)

	Competence satisfaction	"If I write a description and then see an image exactly like what I imagined, it's very motivating." (P11)
	Reduced speaking anxiety	"If there's a video, I can practice intonation and expression. I feel more confident." (P8)
4. Shaping Teacher Roles	GenAI as co-designer	"I can ask ChatGPT to make a complete lesson plan, then I modify it according to the class's needs." (P1)
	Cultural and content concerns	"AI helps a lot, but sometimes I worry if the materials it creates are wrong or culturally inappropriate." (P4)

Theme 1: Multimodal GenAI as a language empowerment tool

Participants described multimodal GenAI as a resource extending language learning well beyond conventional classroom boundaries, deliberately combining textual outputs with AI-generated images, dialogue scripts, and visual contextual cues rather than relying on text alone:

> "If there is a difficult word, I can ask ChatGPT for its synonym, and then I also look at the context in the image created by AI." (P3)

Theme 2: Navigating the balance between support and overreliance

While participants valued GenAI's immediacy and accessibility, many articulated genuine concern about excessive cognitive dependency, elaborating the quantitative discrepancy between abstract concern about critical thinking (M = 3.90) and low self-reported overreliance (M = 2.62). In the interviews, the boundary between productive use and dependency emerged as a lived tension:

> "If I use ChatGPT too often, I feel lazy to think. All the answers are already there." (P2)

P2's reflection articulates what Grinschgl and Neubauer (2022) term cognitive offloading, the gradual transfer of cognitive effort to an external tool, a pattern Gerlich (2025) similarly links to declining critical thinking with habitual AI use. This participant reported only moderate dependency in the survey, consistent with the self-assessment bias discussed earlier. Other participants demonstrated more sophisticated self-regulatory strategies:

> "Usually, I ask for an example text, but I change the sentences so they match my style." (P9)

P9's approach reflects the self-monitoring and evaluative regulation Liu et al. (2025) associate with strategic AI use, evaluating output against personal standards and restructuring it accordingly. This adaptive engagement was more characteristic of positive and mixed-profile participants; sceptical-profile participants tended toward either avoidance or unreflective acceptance, with less evidence of strategic modification.

Theme 3: Enhancing engagement and motivation through multimodality

Participants consistently characterised multimodal GenAI experiences as more engaging and emotionally rewarding than conventional tasks, with AI-generated images, scenario simulations, and dialogue scripts transforming practice into experiences students described as immersive:

> "If there are pictures and videos, I feel more excited to learn. It's like playing a game." (P5)

Learners also derived intrinsic motivation from seeing AI-generated visuals that matched their own written descriptions:

> "If I write a description and then see an image exactly like what I imagined, it's very motivating." (P11)

P11's account reflects the competence need SDT identifies as a primary driver of intrinsic motivation (Ryan & Deci, 2020): the visible correspondence between intention and outcome generates a sense of capability that reinforces continued engagement. Reduced speaking anxiety through low-stakes AI-mediated practice was similarly prominent:

> "If there's a video, I can practice intonation and expression. I feel more confident." (P8)

These findings converge with the quantitative results for enjoyment (M = 3.57) and writing confidence (M = 3.58), aligning with Teng's (2025) finding of sustained motivational gains among EFL writers using GenAI. Taken together, multimodal GenAI appears to create conditions in which learners feel safe enough to try, and affective safety, as much as linguistic input, may drive sustained engagement.

Theme 4: Shaping teacher roles and pedagogical practices

This theme emerged from five participants who were simultaneously undergraduate students and

pre-service teachers. Their reflections positioned multimodal GenAI as a co-designer in lesson development, a starting point for material generation subject to professional judgment:

> "I can ask ChatGPT to make a complete lesson plan, then I modify it according to the class's needs." (P1)

Rather than a shortcut, GenAI served these teacher-participants as an efficiency tool preserving space for professional decision-making, though they also raised substantive concerns about the quality and cultural appropriateness of AI-generated content:

> "AI helps a lot, but sometimes I worry if the materials it creates are wrong or culturally inappropriate." (P4)

P4's concern points to the gap between GenAI's generative capacity and the contextual sensitivity effective EFL instruction requires, a gap Darwin et al. (2024) documented at the learner level too. These findings converge with the strong rejection of GenAI as a teacher replacement (M = 2.33), suggesting the teacher's role is transformed rather than displaced.

Table 5 presents the joint display integrating quantitative findings, qualitative themes, theoretical lenses, and integration types, following Fetters and Tajima's (2022) framework.

Table 5. *Joint display*

Quantitative Finding	Qualitative Theme	Theoretical Lens	Integration Type
High ped. value (M=3.62); writing (M=3.68); confidence (M=3.58)	Theme 1: Language Empowerment Tool	CTML (Mayer, 2021); SCT-ZPD (Poehner & Lantolf, 2024)	Elaboration
High critical thinking concern (M=3.90) vs. low overreliance (M=2.62)	Theme 2: Support vs. Overreliance	SCT overmediation; SRL (Liu et al., 2025); Optimistic bias (Hofer et al., 2022)	Elaboration + Expansion
High enjoyment (M=3.57) and confidence (M=3.58)	Theme 3: Engagement Through Multimodality	SDT competence need (Ryan & Deci, 2020); Multiliteracies (Zapata, 2026)	Elaboration
High guidance need (M=3.72); low AI-as-teacher (M=2.33)	Theme 4: Shaping Teacher Roles	SCT mediation (Lantolf et al., 2020); Situated agency (Ushioda, 2020)	Expansion

GenAI as a learning scaffold

Across both phases, students perceived multimodal GenAI as a tool that extended their language learning beyond what they could achieve independently, consistent with the ZPD (Poehner & Lantolf, 2024) and the broader sociocultural account of mediated language development (Lantolf et al., 2020). The composite mean for Perceived Instructional Benefit (M = 3.62), alongside item-level means for writing improvement (M = 3.68) and feedback comprehension (M = 3.66), reflects substantive pedagogical gain, corroborating Jeon's (2024) finding that AI chatbot affordances support accessible, learner-responsive feedback. The qualitative data added what the survey could not: rather than accepting GenAI outputs passively, students consistently described modifying and evaluating them, indicating active cognitive engagement.

The risk of overmediation, where the tool displaces rather than scaffolds development, was nonetheless evident: the highest survey item mean (M = 3.90) shows students are aware of this risk, yet the discrepancy with the moderate self-reported overreliance (M = 2.62) is more instructive. Interview data confirm students can name the risk of dependency without recognising themselves as its subjects, consistent with the Dunning-Kruger pattern documented by Hofer et al. (2022), a direct implication for educators relying on self-report surveys, which may underestimate dependent use.

Multimodal affordances and cognitive integration

In Themes 1 and 3, participants described purposefully combining text, images, and audio, aligning directly with Mayer's (2021) CTML, when both representational channels are simultaneously active, understanding deepens and retention strengthens. The motivational resonance students reported when AI-generated visuals matched their written descriptions illustrates Mayer's coherence effect, where aligning verbal and imagistic representations reinforces comprehension and sustains engagement.

p-ISSN 2252-7427, e-ISSN 2541-3635
https://journal.uniku.ac.id/index.php/IEFLJ/index

The multiliteracies framework (Zapata, 2026) extends this reading, participants operated within the multimodal semiotic environment the framework envisions, combining text, image, and audio to construct meaning no single mode could convey alone, since GenAI's transduction capabilities create transformative opportunities provided learners develop the critical AI literacy to evaluate and guide those transductions. That participants actively modified and evaluated outputs suggests emerging multimodal literacy practices even without explicit scaffolding. The strong rejection of GenAI as a teacher replacement (M = 2.33) is theoretically significant here, reflecting students' intuitive recognition that contextual sensitivity, cultural knowledge, and relational attunement remain central to effective instruction in ways GenAI cannot replicate.

Learner agency, self-determination, and ethical uncertainty

SDT (Ryan & Deci, 2020) offers the most productive lens for RQ3: when GenAI enabled students to complete tasks in visible, tangible ways, it met the competence need SDT identifies as driving intrinsic motivation, converging with enjoyment (M = 3.57) and writing confidence (M = 3.58) and aligning with Teng's (2025) finding of sustained motivational gains among EFL writers using GenAI. In contrast, using GenAI for tasks expected to be completed independently (M = 2.75), amid uncertainty about its acceptability (M = 3.46), reflects a convenience-oriented, extrinsically driven pattern unlikely to support long-term autonomous motivation.

Students who adapted GenAI outputs to their own voice demonstrated the self-monitoring and evaluative regulation Liu et al. (2025) identify as central to strategic AI use, evaluating output against personal standards and adjusting accordingly. The strong demand for lecturer guidance on ethical GenAI use (M = 3.72) is particularly significant where institutional policy remains nascent: Ushioda (2020) argues situated agency develops within social and institutional structures, and without clear guidelines students face conditions that foster anxiety and inconsistent practice. The qualitative data confirm this across all three perception profiles, with participants expressing a desire not for permission or prohibition but for pedagogically informed guidance.

An alternative explanation deserves consideration: differences in item framing between the survey's abstract wording and the interview's concrete prompts, rather than genuine self-assessment bias alone, may partly account for the gap, a possibility future within-subject designs could test. Findings are best read as indicative of these five institutions and comparable Indonesian contexts rather than broadly generalisable nationwide, given the variation in resourcing and institutional GenAI policy across the country.

Pedagogical implications for Indonesian EFL Contexts

Several concrete directions follow. Language lecturers can integrate GenAI literacy into existing course structures without specialised infrastructure, despite large classes, limited contact hours, and uneven digital access—for instance, structured prompting tasks requiring students to compare AI-generated text with their own drafts, evaluate cultural accuracy, and reflect on which cognitive steps they delegate to the tool, consistent with Zapata's (2026) multiliteracies pedagogy.

For teacher educators, Theme 4 points to a professional development need, teacher-participants described productive but unguided GenAI use in material design, without frameworks for evaluating accuracy or cultural appropriateness. Professional development in Indonesian LPTK settings should include practical protocols for assessing AI-generated materials for cultural representation and reliability before classroom adoption.

Theoretical contributions to AI-mediated EFL research

This study advances AI-mediated EFL research in three ways. First, SCT, Multimodal Learning Theory, and SDT function as complementary rather than competing frameworks. SCT explains how learners direct GenAI mediation, Multimodal Learning Theory accounts for why combining textual and non-textual outputs deepens engagement, and SDT clarifies why visible competence gains sustain motivation despite ethical uncertainty. This contrasts with Wang et al.'s (2025) observation that most GenAI-in-language-education studies apply a single lens, typically SDT or TAM, in isolation, the joint display (Table 5) instead shows these dimensions co-occurring within the same learner accounts. Second, the study evidences that Indonesian EFL learners actively and purposefully engage GenAI's multimodal dimension, extending a literature still dominated by East Asian settings (Li et al., 2025) and writing-focused designs (Huang & Mizumoto, 2024): where Li et al. (2025) found multimodality

discussed mainly as a technical feature, participants here treated combining text, image, and audio as a deliberate strategy rather than an incidental affordance. Third, documenting the discrepancy between abstract concern about critical thinking (M = 3.90) and self-assessed dependency (M = 2.62) advances understanding of how learner self-perception may under-represent actual GenAI-mediated cognitive practice, with direct implications for self-report-only designs such as several reviewed by Wang et al. (2025).

CONCLUSION

This study examined EFL students' perceptions and experiences of multimodal GenAI across five Indonesian higher education institutions using an explanatory sequential mixed-methods design. Three findings emerge from the integrated analysis. First, students consistently valued GenAI as a supplementary learning tool while firmly rejecting it as a substitute for human teachers, confirmed by both the survey (AI-as-teacher M = 2.33) and interview data. Second, abstract risk awareness and self-assessed practice do not necessarily align. Students recognised cognitive dependency as a general concern (M = 3.90) without perceiving their own use as dependent (M = 2.62), a discrepancy with methodological implications for EFL-GenAI research relying solely on self-report. Third, the strong demand for lecturer guidance on ethical GenAI use (M = 3.72) points to an institutional gap particularly pressing where formal AI policy remains underdeveloped.

In terms of practice, the findings support integrating explicit AI literacy components into EFL course design, specifically activities requiring learners to evaluate, adapt, and critically reflect on GenAI outputs rather than accept them at face value. For researchers, the study demonstrates the value of mixed-methods designs that move beyond attitude surveys to explain how learners actually engage with GenAI. Future work should extend inquiry to language skills beyond writing, adopt longitudinal designs, and engage more diverse participant groups to strengthen transferability across Indonesian EFL contexts, as the pedagogical frameworks guiding GenAI use must keep pace with the tools themselves.

For Indonesian EFL educators specifically, three concrete steps follow. Lecturers can build short, scaffolded GenAI tasks into existing assignments, requiring both an AI-generated draft and a justified revised version. So that, critical evaluation becomes a graded task, not an unassessed expectation. Programme coordinators can treat the ethical uncertainty documented here (M = 3.46) as a starting point for department-level guidelines on acceptable GenAI use by assignment type, rather than leaving students to infer boundaries alone. Institutionally, universities should treat GenAI policy as a curriculum matter rather than only an integrity matter. Students here asked mainly for guidance on using GenAI well, not rules about when it is prohibited, so embedding AI literacy within existing programmes will reach students better than standalone policy documents.

This study has several limitations. Convenience sampling limits transferability, and future studies should include participants from more varied proficiency levels, programme types, and institutional contexts. The cross-sectional design precludes conclusions about how perceptions and practices change over time, longitudinal designs are needed. Self-report measures of sensitive behaviours, such as unauthorised GenAI use, are susceptible to social desirability effects, which complementary data from learning logs or screen recordings could address.

ACKNOWLEDGEMENT

The authors thank the students and institutions across five Indonesian universities for their participation, and the ELT scholars and colleagues who assisted with instrument review and peer debriefing.

DISCLOSURE OF AI USE

In preparing this manuscript, the authors made selective use of generative AI assistance for language refinement, grammar checking, and formatting support. The conceptual framing, research design, data analysis, interpretation, and conclusions reported in this study remain entirely the authors' own intellectual work, and the authors take full responsibility for the accuracy, originality, and academic integrity of the manuscript. No AI tool was involved in data collection, statistical analysis, or the generation of findings.

REFERENCES

Almusharraf, A., Bailey, D., Almusharraf, N., & Alotaibi, T. (2025). Students' perceptions of generative AI in EFL writing: Strategies, self-efficacy, satisfaction and behavioural intention. *Australasian Journal of Educational Technology*, *41*(5), 18–36. https://doi.org/10.14742/ajet.10045

Alsaedi, N. (2024). ChatGPT and EFL/ESL writing: A systematic review of advantages and challenges.

English Language Teaching, 17(5), 41–50. https://doi.org/10.5539/elt.v17n5p41

Arifin, M. A., Rahman, A. A., Balla, A., Susanto, A. K., & Pratiwi, A. C. (2024). ChatGPT affordances and Indonesian EFL students' perceptions in L2 writing: A collaborative reflexive thematic analysis. *Changing English: Studies in Culture and Education, 32 (2),* 195–211. https://doi.org/10.1080/1358684X.2024.2418132

Atmojo, A. E. P., & Nugroho, A. (2020). EFL classes must go online! Teaching activities and challenges during COVID-19 pandemic in Indonesia. *Register Journal, 13*(1), 49–76. https://doi.org/10.18326/rgt.v13i1.49-76

Baidoo-Anu, D., & Owusu Ansah, L. (2023). Education in the era of generative artificial intelligence (AI): Understanding the potential benefits of ChatGPT in promoting teaching and learning. *Journal of AI, 7*(1), 52–62. https://doi.org/10.61969/jai.1337500

Braun, V., & Clarke, V. (2021). *Thematic analysis: A practical guide.* Sage.

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*(1), 38. https://doi.org/10.1186/s41239-023-00408-3

Chen, X., Zou, D., & Xie, H. (2020). Fifty years of British Journal of Educational Technology: A topic modeling based bibliometric perspective. *British Journal of Educational Technology, 51*(3), 692–708. https://doi.org/10.1111/bjet.12907

Chiu, T. K. F. (2024). Future research recommendations for transforming higher education with generative AI. *Computers and Education: Artificial Intelligence, 6*, Article 100197. https://doi.org/10.1016/j.caeai.2023.100197

Ci, F., & Jiang, L. (2025). The integration of generative artificial intelligence into digital multimodal composing in second language classrooms: A scoping review from the perspective of tasks. *Digital Applied Linguistics, 2*, 102939. https://doi.org/10.29140/dal.v2.102939

Cong-Lem, N. (2025). Bridging or breaking? A systematic review of how generative AI shapes equity in foreign language education. *International Journal of Applied Linguistics. 36*(2), 1-12. https://doi.org/10.1111/ijal.70052

Creswell, J. W. (2022). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). Sage.

Darwin, Rusdin, D., Mukminatien, N., Suryati, N., Laksmi, E. D., & Marzuki. (2024). Critical thinking in the AI era: An exploration of EFL students' perceptions, benefits, and limitations. *Cogent Education, 11*(1), Article 2290342. https://doi.org/10.1080/2331186X.2023.2290342

Fetters, M. D., & Tajima, C. (2022). Joint displays of integrated data collection in mixed methods research. *International Journal of Qualitative Methods, 21*. https://doi.org/10.1177/16094069221104564

Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies, 15*(1), Article 6. https://doi.org/10.3390/soc15010006

Godwin-Jones, R. (2023). Emerging spaces for language learning: AI bots, ambient intelligence, and the metaverse. *Language Learning & Technology, 27*(2), 6–27. https://doi.org/10.64152/10125/73501

Goretzko, D., Pham, T. T. H., & Bühner, M. (2021). Exploratory factor analysis: Current use, methodological developments and recommendations for good practice. *Current Psychology, 40*, 3510–3521. https://doi.org/10.1007/s12144-019-00300-2

Grinschgl, S., & Neubauer, A. C. (2022). Supporting cognition with modern technology: Distributed cognition today and in an AI-enhanced future. *Frontiers in Artificial Intelligence, 5*, Article 908261. https://doi.org/10.3389/frai.2022.908261

Guo, K., & Li, D. (2024). Understanding EFL students' use of self-made AI chatbots as personalized writing assistance tools: A mixed methods study. *System, 124*, Article 103362. https://doi.org/10.1016/j.system.2024.103362

Hayes, A. F., & Coutts, J. J. (2020). Use omega rather than Cronbach's alpha for estimating reliability. But… *Communication Methods and Measures, 14*(1), 1–24. https://doi.org/10.1080/19312458.2020.1718629

Hofer, G., Mraulak, V., Grinschgl, S., & Neubauer, A. C. (2022). Less intelligent and unaware? Accuracy and Dunning-Kruger effects for self-estimates of different aspects of intelligence. *Journal of Intelligence, 10*(1), Article 10. https://doi.org/10.3390/jintelligence10010010

Huang, J., & Mizumoto, A. (2024). Examining the effect of generative AI on students' motivation and writing self-efficacy. *Digital Applied Linguistics, 1*, 102324. https://doi.org/10.29140/dal.v1.102324

Jeon, J. (2022). Exploring AI chatbot affordances in the EFL classroom: Young learners' experiences and perspectives. *Computer Assisted Language Learning, 37*(1–2), 1–26. https://doi.org/10.1080/09588221.2021.2021241

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

Lantolf, J. P., Poehner, M. E., & Thorne, S. L. (2020). Sociocultural theory and L2 development. In B.

VanPatten, G. D. Keating, & S. Wulff (Eds.), *Theories in second language acquisition: An introduction* (3rd ed., pp. 223–247). Routledge.

Lee, S.-M. (2022). An investigation of machine translation output quality and the influencing factors of source texts. *ReCALL, 34*(1), 81–94. https://doi.org/10.1017/S0958344021000124

Lee, S., Choe, H., Zou, D., & Jeon, J. (2025). Generative AI (GenAI) in the language classroom: A systematic review. *Interactive Learning Environments, 34*(1), 335–359. https://doi.org/10.1080/10494820.2025.2498537

Li, B., Tan, Y. L., Wang, C., & Lowell, V. (2025). Two years of innovation: A systematic review of empirical generative AI research in language learning and teaching. *Computers and Education: Artificial Intelligence, 9, Article 100226.* https://doi.org/10.1016/j.caeai.2025.100445

Liu, G., & Ma, C. (2023). Measuring EFL learners' use of ChatGPT in informal digital learning of English based on the technology acceptance model. *Innovation in Language Learning and Teaching, 18*(2), 125–138. https://doi.org/10.1080/17501229.2023.2240316

Liu, X., Xiao, Y., & Li, D. (2025). Assessing strategic use of artificial intelligence in self-regulated learning: Instrument development and evidence from Chinese university students. *International Journal of Educational Technology in Higher Education, 22*, Article 69. https://doi.org/10.1186/s41239-025-00567-5

Mayer, R. E. (2021). *Multimedia learning* (3rd ed.). Cambridge University Press.

McKenna, L. (2022). Translation of research interviews: Do we have a problem with qualitative rigor? *Nurse Author & Editor, 32*(1), 1–3. https://doi.org/10.1111/nae2.31

Nazim, M., & Alzubi, A. A. F. (2025). Empowering EFL teachers' perceptions of generative AI-mediated self-professionalism. *PLOS ONE, 20*(6), e0326735. https://doi.org/10.1371/journal.pone.0326735

Nyimbili, F., & Nyimbili, L. (2024). Types of purposive sampling techniques with their examples and application in qualitative research studies. *British Journal of Multidisciplinary and Advanced Studies, 5*(1), 90–99. https://doi.org/10.37745/bjmas.2022.0419

Poehner, M. E., & Lantolf, J. P. (2024). *Sociocultural theory and second language developmental education.* Cambridge University Press.

R Core Team. (2023). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing. https://www.R-project.org/

Ryan, R. M., & Deci, E. L. (2020). Intrinsic and extrinsic motivation from a self-determination theory perspective. *Contemporary Educational Psychology, 61*, Article 101860. https://doi.org/10.1016/j.cedpsych.2020.101860

Selwyn, N. (2024). Digital degrowth: Toward radically sustainable education technology. *Learning, Media and Technology, 49*(2), 186–199. https://doi.org/10.1080/17439884.2022.2159978

Sijtsma, K., & Pfadt, J. M. (2021). Part II: On the use, the misuse, and the very limited usefulness of Cronbach's alpha: Discussing lower bounds and correlated errors. *Psychometrika, 86*(4), 843–860. https://doi.org/10.1007/s11336-021-09789-8

Sim, S., Kim, Y., & Ku, K. (2025). Exploring generative AI as a roleplay interlocutor in L2 task-based pragmatics learning: Comparing ChatGPT-learner and technology-mediated peer interactions. *TESOL Quarterly, 59*(S1), S86–S116. https://doi.org/10.1002/tesq.70010

Tashakkori, A., Johnson, R. B., & Teddlie, C. (2021). *Foundations of mixed methods research* (2nd ed.). Sage.

Teng, M. F. (2025). Examining longitudinal development of writing motivation in the GenAI context: A self-determination theory perspective. *Learning and Motivation, 91*, Article 102157. https://doi.org/10.1016/j.lmot.2025.102157

Ushioda, E. (2020). *Language learning motivation: An ethical agenda for research.* Oxford University Press.

Vo, T. K. A., & Nguyen, H. (2024). Generative artificial intelligence and ChatGPT in language learning: EFL students' perceptions of technology acceptance. *Journal of University Teaching and Learning Practice, 21*(6), 199–218. https://doi.org/10.53761/fr1rkj58

Wang, Y., Zhang, T., Yao, L., & Seedhouse, P. (2025). A scoping review of empirical studies on generative artificial intelligence in language education. *Innovation in Language Learning and Teaching. 19*(1), 1-28. https://doi.org/10.1080/17501229.2025.2509759

Wu, F., Dang, Y., & Li, M. (2025). A systematic review of responses, attitudes, and utilization behaviors on generative AI for teaching and learning in higher education. *Behavioral Sciences, 15*(4), Article 467. https://doi.org/10.3390/bs15040467

Yan, Y., Sun, W., & Zhao, X. (2024). Metaphorical conceptualizations of generative artificial intelligence use by Chinese university EFL learners. *Frontiers in Education, 9*, Article 1430494. https://doi.org/10.3389/feduc.2024.1430494

Yeo, J. (2023). Academic integrity in the age of AI: Student perceptions and institutional responses. *Studies in Higher Education, 48*(10), 1567–1581. https://doi.org/10.1002/tesj.716

Zapata, G. C. (Ed.). (2026). *Generative AI technologies, multiliteracies, and language education.* Routledge.