LinguaEducare: Journal of English and Linguistic Studies
ISSN: 3064-0547 (e)
2026, Vol. 3, No. 1, page 59-72
https://journal.ciptapustaka.com/index.php/LEC

OPEN ACCESS

Ethical Awareness and Algorithmic Bias in AI-IDLE: Perspectives of Sharia Faculty Students at UIN Raden Intan Lampung

Kartika S[1*]
[1]*Sharia Faculty, UIN Raden Intan Lampung, Indonesia*

ABSTRACT

Generative AI tools are increasingly used for AI-mediated Informal Digital Learning of English (AI-IDLE), yet research has paid limited attention to how learners negotiate ethical responsibility and algorithmic bias during everyday, out-of-class use, particularly in value-oriented faculties. This qualitative study examined ethical awareness and perceived algorithmic bias in AI-IDLE among Sharia Faculty students at UIN Raden Intan Lampung. Using a qualitative descriptive case study design, twelve undergraduates were recruited through purposive maximum-variation sampling. Data were generated from semi-structured interviews and a curated set of anonymized AI interaction artifacts (e.g., prompts and outputs for explanation, drafting, and translation), supported by brief artifact walkthroughs when feasible. Data were analyzed via reflexive thematic analysis with iterative coding, memoing, and cross-source comparison to strengthen interpretive transparency. Findings showed that ethical awareness operated as context-sensitive boundary-making: students differentiated AI as a learning scaffold versus a substitute for intellectual labor, tightened limits in assessment-adjacent tasks, and reclaimed authorship through substantive rewriting to preserve voice and stance. Ethical reasoning also included privacy stewardship through data minimization in prompts, though awareness varied. Regarding bias, students most often noted Western-centric framing in value-laden examples and language standardization that diluted pragmatic or culturally embedded meanings. Fluent outputs produced an authority effect, but many participants reported mitigation routines such as verification, requesting multiple viewpoints, and re-prompting for Indonesian/Islamic contextualization. The study implies that AI-IDLE guidance should integrate ethical boundary-setting, privacy-aware prompting, and critical AI literacy practices to support responsible and culturally responsive informal language learning..

ARTICLE INFO

Keywords:
AI-IDLE;
algorithmic bias;
critical AI literacy;
ethical awareness;
informal digital learning of English

Article History:
Received: 26 February 2026
Revised: 9 March 2026
Accepted: 24 March 2026
Published: 25 March 2026

How to Cite in APA Style:
S, K. (2026). Ethical Awareness and Algorithmic Bias in AI-IDLE: Perspectives of Sharia Faculty Students at UIN Raden Intan Lampung. *LinguaEducare: Journal of English and Linguistic Studies, 3*(1), 59-72.
https://doi.org/10.63324/lec.3v.1i.139

Introduction

Generative AI—especially large language models (LLMs) such as ChatGPT—has rapidly become part of university students' everyday academic and language-learning routines, not only inside classrooms but also through self-directed, "always-on" interactions. In language learning contexts, these tools can expand access to

*Corresponding author: Kartika S kartika@radenintan.ac.id

explanations, models, and conversational practice, while also accelerating drafting and revision work that used to require more time and human support (Guo, 2024). At the same time, the same probabilistic fluency that makes LLM outputs look convincing can obscure inaccuracies, shallow reasoning, and overconfident generalizations—risks that matter when students treat AI outputs as trustworthy learning input (Sawalha et al., 2024). As a result, recent scholarship increasingly frames GenAI use as a literacy issue: learners need skills for evaluating, questioning, and ethically integrating AI support rather than simply consuming outputs (Long et al., 2023). This shift is particularly relevant in EFL settings, where AI-generated language can shape learners' vocabulary, pragmatics, and cultural understandings.

Within language education research, Informal Digital Learning of English (IDLE) provides a useful lens for understanding how learners develop English through digital activities outside formal instruction, such as social media engagement, online videos, games, and interest-driven communities. IDLE emphasizes that learning is often incidental, socially situated, and motivated by personal goals rather than curricular requirements (Zadorozhnyy & Lee, 2025). However, outcomes are not determined by "time online" alone; the quality and diversity of activities, the depth of engagement, and learners' strategic choices all influence what is actually learned. Empirical work has shown that affective factors and patterns of participation can shape IDLE outcomes, suggesting that informal learning is both opportunity-rich and uneven across learners. Meanwhile, synthesis work indicates IDLE research has grown quickly but still needs sharper attention to digital literacies and learner agency in shaping meaningful language development (Liu et al., 2024). Therefore, IDLE research highlights why "where" and "how" learners interact digitally may matter as much as "how often."

A key development is that IDLE is now increasingly mediated by AI, as learners use chatbots and LLM-based tools for quick explanations, vocabulary support, translation, and simulated conversation—creating what recent work terms AI-mediated IDLE (AI-IDLE). In AI-IDLE, learners can generate tailored prompts, request feedback, and iterate instantly, which can strengthen autonomy and persistence in self-learning (Liu et al., 2025). Yet AI-IDLE also changes the nature of "input" and "interaction," because the language and cultural content learners receive is generated, not simply encountered, and it can be optimized for smoothness rather than pedagogical value. Studies of learners' adoption of AI in informal English learning suggest that perceived usefulness and convenience can quickly normalize AI as a default learning companion, especially for self-study (Liu & Ma, 2023). This normalization is important because it can quietly shift learners' standards for evidence, authorship, and critical evaluation of content. Therefore, understanding AI-IDLE requires attention not only to learning benefits but also to the ethical and epistemic conditions under which informal AI support becomes "acceptable" and "trusted."

As AI support moves into informal learning spaces, ethical awareness becomes more complex because boundaries between legitimate assistance and problematic outsourcing are less visible than in classroom assignments. Students may use AI for "practice" but later reuse the same outputs for graded tasks, creating gray zones around originality, citation, and disclosure (Celik, 2023). At the same time, AI systems raise

privacy concerns, because prompts may include personal data, academic work, or sensitive beliefs—issues that are rarely addressed explicitly when students use tools casually outside class (Jin et al., 2025). Higher education research on ChatGPT and academic integrity shows that both students and faculty recognize productivity and learning benefits, while also voicing concerns about plagiarism, fabricated information, and the need for clearer institutional guidelines for responsible use. From a broader responsible-AI perspective, ethics is also intertwined with fairness, accountability, and transparency—meaning that ethical use is not only about "not cheating," but also about understanding system limits and responsibilities when AI tools influence learning decisions (Rapanta, 2025). In short, AI-IDLE makes ethics a daily practice issue, not merely a policy statement, because students continually decide what to trust, what to submit, and what to acknowledge.

Alongside ethics, algorithmic bias is a central concern because LLM outputs are shaped by training data and design choices that can privilege certain languages, cultural narratives, and value assumptions. For EFL learners, bias matters not only as a social-justice issue but also as a learning issue: biased outputs can become repeated input, examples, and "models" that shape learners' worldview and communicative norms (Kundi et al., 2022). Recent evidence suggests that LLM responses can reflect Western-leaning cultural perspectives, which becomes particularly consequential when learners use AI to interpret social issues, cultural practices, or moral questions in English (Tao et al., 2024). Bias is also not limited to overt stereotypes; it can appear as subtle representational imbalance—whose histories are centered, whose norms are treated as default, and which perspectives are framed as "reasonable." Work on evaluating cultural bias and historical representativeness in LLM outputs underscores that bias can be systematic and measurable, not merely anecdotal. Therefore, investigating bias in AI-IDLE means examining how learners notice, interpret, and respond to culturally skewed content during everyday self-learning interactions.

These issues are especially salient for Sharia Faculty students at UIN Raden Intan Lampung because their academic formation foregrounds ethical reasoning, normative evaluation, and moral accountability—dimensions that directly intersect with debates on responsible AI use. From an Islamic ethics perspective, global AI norm-setting can be overly Eurocentric, and there is increasing scholarly emphasis on pluralist ethical benchmarking that brings non-Western moral traditions into conversations about fairness, transparency, and the "good" in AI deployment (Mandasari et al., 2025). In practice, this suggests Sharia students may approach AI use not only through pragmatic concerns (speed, convenience) but also through value-based judgments (permissibility, integrity, harm/benefit, and responsibility). At the same time, Indonesian EFL-related research shows that learners and future teachers often view ChatGPT as helpful for writing and learning support, indicating high adoption potential in local contexts even while ethical concerns remain (Marzuki et al., 2023). Despite this relevance, AI-IDLE research still rarely centers students from faith-based faculties, whose ethical frameworks may shape how bias is identified and how responsible use is defined. This gap motivates a context-sensitive qualitative inquiry that treats ethics and bias as lived, interpreted experiences rather than abstract risks.

Accordingly, this qualitative study explores how Sharia Faculty students at UIN Raden Intan Lampung make sense of ethical awareness and algorithmic bias while engaging in AI-IDLE for English learning beyond the classroom. Rather than assuming AI use is uniformly beneficial or uniformly harmful, the study positions students as active interpreters who negotiate norms of authorship, trust, and cultural fit in real usage situations. The goal is to clarify what "responsible" AI-IDLE looks like from students' perspectives, including the criteria they use to judge acceptability, the moments that trigger ethical doubt, and the practical strategies they use to evaluate questionable outputs. By foregrounding learner voices, the study also aims to contribute to culturally responsive guidance for AI-supported informal learning in Indonesian higher education contexts. Ultimately, this inquiry treats ethical awareness and bias recognition as part of students' developing critical AI literacy within language learning.

1. How do Sharia Faculty students at UIN Raden Intan Lampung perceive and enact ethical awareness when using AI tools for informal digital learning of English (AI-IDLE)?
2. How do these students perceive and respond to algorithmic bias in AI outputs encountered during AI-IDLE, and what strategies do they use to evaluate and manage it?

Method

This study adopted a qualitative descriptive case study (Yin, 2018) design to examine Sharia Faculty students' ethical awareness and perceptions of algorithmic bias in AI-IDLE at UIN Raden Intan Lampung. The case is bounded by a single institutional context and focuses on how students interpret and negotiate ethical responsibility and bias in everyday AI-IDLE practices. Participants were selected through purposive maximum-variation sampling to capture diversity in year of study, self-reported English proficiency, intensity of AI use, and types of AI tools used. The primary instrument was a semi-structured interview guide aligned with the two research questions, supplemented by an artifact checklist to help participants identify relevant examples of AI-IDLE use. Data were collected through individual interviews (45–75 minutes) conducted online or face-to-face, with audio recording and verbatim transcription; participants also provided limited, anonymized examples of AI interactions to contextualize their accounts. Data were analyzed using reflexive thematic analysis, beginning with familiarization and initial coding, followed by iterative theme development and refinement that explicitly mapped themes to RQ1 and RQ2.

Participants consisted of undergraduate students from the Faculty of Sharia, UIN Raden Intan Lampung, who had used at least one AI tool (e.g., ChatGPT-like chatbots, grammar assistants, translation tools) for informal English learning within the last three to six months. Recruitment occurred via faculty networks and student groups, and sampling continued until analytic sufficiency was reached, with an anticipated range of 12–20 participants depending on the richness of accounts and cross-case variation. The main instrument remained the semi-structured interview protocol, while a short demographic form captured background variables relevant to interpretation (semester,

frequency of AI use, and typical AI-IDLE purposes). Data collection prioritized participants' naturally occurring practices by inviting them to describe concrete episodes of AI-IDLE use rather than responding to hypothetical scenarios. Interviews were complemented by a brief artifact walkthrough when feasible, where participants explained the context, intent, and perceived issues in one or two anonymized prompts/outputs. The analysis process incorporated these contextual details as interpretive anchors, ensuring that themes were grounded in specific practices rather than abstract attitudes.

To strengthen the alignment between methods and ethical/bias constructs, the interview instrument included prompts on academic integrity (authorship, disclosure, acceptable assistance), privacy (data sharing in prompts), and responsible use (verification, dependence, and decision-making). For algorithmic bias, prompts explored perceived cultural/ideological skew, stereotyping, linguistic variety preferences, and perceived authority effects of fluent AI outputs. Data collection followed a consistent protocol: consent and confidentiality briefing, interview, optional artifact submission, and a short follow-up for clarification when needed. Artifacts were limited to small excerpts (e.g., screenshots) and were anonymized by removing names, identifiers, and any sensitive personal details before storage. During analysis, artifact excerpts were used to triangulate and enrich interview claims (e.g., illustrating the type of bias perceived) without treating artifacts as frequency-based evidence. Codes were developed inductively, then organized into broader categories corresponding to ethical awareness and bias engagement, with memos documenting interpretive decisions.

Data analysis followed an iterative reflexive thematic analysis procedure (Braun & Clarke, 2021): (1) repeated reading of transcripts and artifacts, (2) initial open coding across the full dataset, (3) clustering codes into candidate themes, (4) reviewing themes for internal coherence and distinctiveness, and (5) defining and naming themes with clear links to each research question. Throughout coding, special attention was given to how participants justified their decisions (e.g., what criteria they used to label something as unethical or biased) and to the strategies they reported for managing risks (e.g., cross-checking sources, rewriting, prompting tactics). Credibility was supported through peer debriefing with a colleague familiar with qualitative analysis, and through an audit trail that recorded coding iterations, analytic memos, and theme revisions. To avoid a simplistic "validation" logic, selective member checking was conducted by sharing a brief theme summary with a small subset of participants (e.g., 3–4 students) and treating their feedback as an additional interpretive resource. Findings were reported with thick description and anonymized excerpts to preserve participant voice.

Ethical considerations were embedded across the design, participant handling, instruments, and procedures. Participants provided written informed consent and were informed they could withdraw at any stage without penalty; confidentiality was maintained through pseudonyms and removal of identifying details from transcripts and artifacts. Because the topic involves academic integrity and tool use, the study avoided eliciting confessions of misconduct; instead, it focused on perceptions, boundary-making, and reasoning processes in AI-IDLE. Data collection minimized risk by discouraging the sharing of sensitive personal data or third-party information in

screenshots and by storing all files in encrypted, password-protected folders accessible only to the research team. The analysis treated ethical awareness and bias not as fixed traits but as contextual judgments shaped by academic norms, tool affordances, and cultural values, which aligns with the interpretive aims of the study. Overall, the methodological choices were intended to produce a credible, context-sensitive account of how Sharia Faculty students navigate ethics and algorithmic bias in AI-IDLE.

Results

Twelve undergraduate students (P01–P12) from the Faculty of Sharia participated in semi-structured interviews. All participants reported routine engagement with AI tools (ChatGPT-like systems, grammar assistants, and translation tools) for AI-mediated Informal Digital Learning of English (AI-IDLE). Twenty-four anonymized artifacts (A01–A24) were analyzed to triangulate interview findings. The analysis generated three themes for RQ1 (ethical awareness) and three themes for RQ2 (algorithmic bias).

RQ1: Ethical Awareness in AI-IDLE

Table 1. Thematic Findings for RQ1: Ethical Awareness in AI-IDLE

Theme	Conceptual Definition	Subthemes	n (out of 12)	Verbatim Evidence	Artifact Anchors
T1. Boundary-Making: Scaffold vs. Substitution	Ethical AI use is defined by whether AI functions as cognitive scaffolding or replaces intellectual labor. Boundaries intensify in assessment-related contexts.	(a) Practice vs graded tasks (b) Brainstorming vs final drafting (c) Cognitive engagement monitoring	11	P03: 'If AI explains grammar, I learn. But if it writes my paragraph and I submit it, that feels dishonest.' P07: 'For practice, I use it freely. For assignments, I rewrite everything myself.'	A04 (AI outline reconstructed manually) A09 (Grammar feedback-only usage)
T2. Authorship and Voice Preservation	Ethical reasoning centers on maintaining epistemic ownership through revision, restructuring, and integration of personal reasoning.	(a) Rewriting AI output (b) Personal stance integration (c) Structural modification	9	P01: 'Even if AI gives ideas, I must change them so they reflect my thinking.' P08: 'If I don't revise it, it's not really mine.'	A11 (65% lexical modification in revision) A13 (Argument paragraph restructured)
T3. Privacy Stewardship and Prompt Regulation	Ethical awareness extends to management of sensitive or identifiable information shared with AI tools.	(a) Data minimization (b) Redaction of case details (c) Risk uncertainty	7	P09: 'I avoid typing real case details. I generalize the situation.' P12: 'I assume	A18 (Redacted legal scenario prompt) A21 (Modified

	it's private, but maybe it's not.'	hypothetical case input)

Table 1 illustrates that ethical awareness in AI-mediated Informal Digital Learning of English (AI-IDLE) is constructed as a contextual and graduated form of moral reasoning rather than a binary classification of acceptable versus unacceptable AI use. The dominant theme, Boundary-Making: Scaffold vs. Substitution (n = 11), indicates that participants consistently differentiated between AI as a cognitive support tool and AI as a replacement for intellectual effort. Ethical concern intensified when AI use approached formal assessment contexts, suggesting that task stakes functioned as a regulatory mechanism. Students were more permissive when AI was used for clarification, vocabulary expansion, or outlining, but expressed discomfort when AI-generated text was directly transferable to graded submissions. This finding suggests that ethical positioning is closely tied to perceived authorship responsibility and academic legitimacy.

The second theme, Authorship and Voice Preservation (n = 9), further reinforces the importance of epistemic ownership in ethical reasoning. Participants emphasized the need to restructure, revise, and integrate personal perspectives into AI-generated drafts in order to maintain intellectual agency. Artifact analysis supports this interpretation, as before–after comparisons revealed substantial lexical and structural transformation rather than superficial editing. Ethical AI use, therefore, was not conceptualized merely as avoiding plagiarism detection but as actively reclaiming authorial voice. This indicates that students' moral evaluations extend beyond compliance toward deeper notions of authenticity and accountability in knowledge production.

The third theme, Privacy Stewardship and Prompt Regulation (n = 7), expands ethical awareness beyond academic integrity into the domain of digital responsibility. Several participants reported deliberately minimizing identifiable information when inputting prompts, particularly in discussions involving legal or religious case scenarios. However, the data also reveal uneven awareness regarding data retention and AI system transparency. Some students assumed conversational privacy without fully understanding backend data governance implications. This variability suggests that ethical AI-IDLE engagement involves not only moral judgment about authorship but also differential levels of digital risk literacy.

Therefore, the findings indicate that ethical awareness in AI-IDLE is enacted through boundary regulation, authorship negotiation, and selective data management. Rather than rejecting AI tools, students positioned themselves as active moral agents who continuously calibrated their engagement depending on context, task demands, and perceived consequences. These patterns suggest that ethical AI use in informal learning environments is best understood as negotiated agency rather than simple compliance with external policy constraints.

RQ2: Algorithmic Bias in AI-IDLE

Table 2. Thematic Findings for RQ1: Ethical Awareness in AI-IDLE

Theme	Conceptual Definition	Subthemes	n (out of 12)	Verbatim Evidence	Artifact Anchors
T4. Western-Centric Framing	AI outputs frequently default to Western socio-cultural examples and normative assumptions presented as universal.	(a) Western case references (b) Value universalization (c) Cultural mismatch	10	P05: 'When I asked about social issues, the examples were always from Western society.' P10: 'It feels like Western values are the default answer.'	A06 (US legal case framing in moral discussion) A17 (Gender role example from Western context)
T5. Language Standardization and Pragmatic Dilution	AI prioritizes standardized English norms, sometimes flattening culturally specific nuance or stance.	(a) Tone shift (b) Directness increase (c) Cultural semantic loss	8	P02: 'It corrected my sentence but changed the meaning slightly.' P11: 'My original tone was softer, but AI made it more direct.'	A15 (Translation shift reducing nuance) A19 (Rephrased culturally embedded phrase)
T6. Authority Effect and Mitigation Strategies	Fluent AI responses create perceived epistemic authority; students apply verification and contextual re-prompting.	(a) Initial trust (b) Cross-checking (c) Re-prompt for local context	9	P04: 'It sounds confident, so I trust it at first.' P06: 'I usually cross-check with other sources before believing it.'	A20 (Re-prompt requesting Indonesian context) A23 (Request for alternative viewpoints)

Table 2 indicates that perceptions of algorithmic bias in AI-mediated Informal Digital Learning of English (AI-IDLE) were neither abstract nor theoretical; rather, they emerged through recurring patterns in AI outputs encountered during everyday use. The most prevalent theme, Western-Centric Framing (n = 10), reveals that participants frequently perceived AI-generated responses as defaulting to Western socio-cultural contexts and normative assumptions. This pattern was particularly salient in morally or socially sensitive discussions, where examples drawn from Western legal systems or social practices were presented as implicitly universal. Participants described this framing not as overt discrimination but as subtle cultural privileging, suggesting that bias manifested through representational imbalance rather than explicit stereotyping. The prominence of this theme underscores that cultural alignment becomes especially visible when AI is used to interpret value-laden or context-specific topics.

The second theme, Language Standardization and Pragmatic Dilution (n = 8), highlights a more nuanced form of bias embedded in linguistic optimization processes.

While participants acknowledged improvements in grammatical accuracy and structural clarity, several reported shifts in tone, stance, and cultural nuance after AI revision. In particular, translations or paraphrases involving culturally embedded expressions sometimes resulted in reduced contextual depth or altered pragmatic force. This suggests that algorithmic preference for standardized English norms may inadvertently flatten localized meaning. Rather than overt error, bias here operated through subtle semantic recalibration, where linguistic fluency masked shifts in intended communicative positioning.

The third theme, Authority Effect and Mitigation Strategies (n = 9), introduces a cognitive dimension of algorithmic bias. Participants described fluent and confident AI responses as initially persuasive, increasing the likelihood of acceptance without immediate scrutiny. However, most participants reported engaging in verification practices, including cross-checking external sources, requesting alternative viewpoints, and re-prompting for Indonesian or Islamic contextualization. These mitigation strategies indicate that perceived authority does not necessarily lead to passive consumption; instead, students demonstrated emergent evaluative competencies. The coexistence of initial trust and subsequent verification suggests that algorithmic authority functions as both a risk factor and a catalyst for critical engagement.

Collectively, the findings suggest that algorithmic bias in AI-IDLE operates at multiple levels: cultural framing, linguistic normalization, and epistemic authority. Importantly, participants did not merely identify bias but actively negotiated it through adaptive prompting and verification practices. This indicates that AI-IDLE environments simultaneously expose learners to representational asymmetries while fostering developing forms of critical AI literacy.

Discussion

The RQ1 findings suggest that ethical awareness in AI-IDLE is enacted through boundary-making that calibrates AI use according to task stakes (practice vs. assessment-adjacent work) and the extent to which AI functions as scaffolding versus substitution. This pattern aligns with AI-IDLE research showing that learners adopt LLM tools because they provide fast, interactive support for informal learning, but their engagement depends on how learners frame “legitimate” use and how they regulate their own learning processes (Liu et al., 2024). In mixed-method AI-IDLE work, GenAI can support language outcomes (oral proficiency), yet sustained extramural engagement is not automatic—implying that self-regulation and learner agency remain central even when AI tools are readily available (Guan et al., 2025).

The prominence of authorship and voice preservation extends the integrity discussion beyond plagiarism-as-detection toward epistemic ownership—who controls ideas, stance, and reasoning in the final product. This resonates with higher-education scholarship on ChatGPT that highlights the tension between productivity gains and risks of outsourcing intellectual work, especially in writing-intensive tasks (Casal & Kessler, 2023). Academic integrity discussions increasingly emphasize assessment redesign and transparent guidance, because the “copy–paste” model does not fully capture how students actually use LLMs (iterative drafting, rewriting, and paraphrasing) (Hastomo,

et al., 2025). In this sense, the findings help specify what integrity looks like in AI-IDLE: not merely avoiding direct submission, but actively reasserting authorial responsibility through revision and decision-making (Cotton et al., 2024).

The theme of privacy stewardship and prompt regulation indicates that ethical awareness in AI-IDLE includes digital responsibility, but unevenly. This is consistent with broader reviews noting that ChatGPT research in higher education is still early and that limitations—including reliability, transparency, and governance issues—remain under-addressed relative to rapid adoption (Hastomo et al., 2025). From an AI-IDLE perspective, this matters because informal learning often occurs outside institutional oversight, making "prompt hygiene" (data minimization, anonymization, and avoiding sensitive details) a practical ethics competency rather than a policy footnote (Baig & Yadegaridehkordi, 2024). Therefore, the present findings point to the need for explicit instruction on privacy-aware prompting within language learning contexts, especially when students discuss case-like or value-laden scenarios.

Under RQ2, the dominance of Western-centric framing is strongly compatible with recent large-scale evaluations showing that widely used LLM outputs can reflect cultural values disproportionately aligned with English-speaking and Protestant European contexts. Such cultural skew becomes consequential in AI-IDLE because learners may treat AI-generated examples and explanations as neutral "models" for reasoning in English, including on moral or socio-legal topics where local norms are essential for meaning accuracy (Tao et al., 2024). Importantly, the same evaluation literature suggests that cultural prompting can partially improve alignment (G. L. Liu et al., 2025), which conceptually matches students' reported mitigation strategies (requesting Indonesian context, multiple viewpoints, or culturally grounded examples).

Finally, the themes of language standardization/pragmatic dilution and the authority effect refine how bias operates in AI-IDLE: not only through overt stereotyping, but through subtle shifts in tone, stance, and what counts as "good English," amplified by the persuasive fluency of AI outputs. This dovetails with evidence that GenAI in informal language learning contexts shows measurable benefits for proficiency and self-regulation, yet also raises the pedagogical challenge of ensuring learners remain critically engaged rather than passively accepting optimized text (Guan et al., 2024). Theoretically, curent findings support positioning critical AI literacy in AI-IDLE as a bundle of practices—verification, constraint-based prompting, triangulation, and stance-preserving revision—rather than as a general attitude of skepticism.

Conclusion

This study indicates that Sharia Faculty students' ethical awareness in AI-IDLE is enacted as context-sensitive regulation rather than a fixed "allowed vs. not allowed" rule. Participants primarily managed their AI use through boundary-making (using AI as a learning scaffold but avoiding direct substitution of intellectual work), reinforced by authorship and voice preservation through substantive rewriting and stance ownership. Ethical awareness also extended to privacy stewardship, with several students reporting data-minimization practices when prompts involved sensitive or

case-like information. Alongside these ethical practices, participants perceived algorithmic bias mainly through Western-centric framing in examples and value-laden explanations, and through language standardization that sometimes diluted pragmatic or culturally embedded nuance; however, many also described mitigation routines such as verification, re-prompting for local contextualization, and requesting alternative viewpoints.

Several limitations should be acknowledged. As a qualitative case study bounded to one institutional context with purposive sampling, the findings are transferable through contextual fit rather than statistically generalizable, and they rely on self-reported experiences and selected artifacts that may be shaped by recall and social-desirability effects. Despite these limitations, the study implies that guidance for AI-IDLE should move beyond "don't copy" rules toward teachable practices of ethical boundary-setting, authorship-preserving revision, and privacy-aware prompting, while strengthening critical AI literacy routines (triangulation, constraint-based prompting, and multi-perspective requests) to reduce cultural framing bias and authority effects. Future research should extend this work through multi-site and longitudinal designs, include richer interaction traces (prompt–revision logs), and evaluate whether explicit instruction in ethical prompting and culturally responsive prompting improves both learning quality and responsible use over time.

References

Baig, M. I., & Yadegaridehkordi, E. (2024). ChatGPT in the higher education: A systematic literature review and research challenges. *International Journal of Educational Research*, *127*, 102411. https://doi.org/10.1016/j.ijer.2024.102411

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

Casal, J. E., & Kessler, M. (2023). Can linguists distinguish between ChatGPT/AI and human writing?: A study of research ethics and academic publishing. *Research Methods in Applied Linguistics*, *2*(3), 100068. https://doi.org/10.1016/j.rmal.2023.100068

Celik, I. (2023). Towards Intelligent-TPACK: An empirical study on teachers' professional knowledge to ethically integrate artificial intelligence (AI)-based tools into education. *Computers in Human Behavior*, *138*, 107468. https://doi.org/10.1016/J.CHB.2022.107468

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

Guan, L., Li, S., & Gu, M. M. (2024). AI in informal digital English learning: A meta-analysis of its effectiveness on proficiency, motivation, and self-regulation. *Computers and Education: Artificial Intelligence*, *7*, 100323. https://doi.org/10.1016/j.caeai.2024.100323

Guan, L., Zhang, E. Y., & Gu, M. M. (2025). Examining generative AI–mediated informal digital learning of English practices with social cognitive theory: a mixed-

methods study. *ReCALL*, *37*(3), 315–331. https://doi.org/10.1017/S0958344024000259

Guo, X. (2024). Facilitator or thinking inhibitor: understanding the role of ChatGPT-generated written corrective feedback in language learning. *Interactive Learning Environments*, *32*, 1–19. https://doi.org/10.1080/10494820.2024.2445177

Hastomo, T., Widiati, U., Ivone, F. M., & Zen, E. L. (2025). EFL teachers' self-efficacy in using AI tools: A comparative study in Indonesia. *Advanced Education*, *19*(27), 103–123. https://doi.org/10.20535/2410-8286.335502

Hastomo, T., Widiati, U., Ivone, F. M., Zen, E. L., Hasbi, M., & Khulel, B. (2025). AI-powered conversational agents and intercultural learning: Insights from Indonesian EFL students. *Intercultural Communication Education*, *8*(1), 103217. https://doi.org/10.29140/ice.v8n1.103127

Jin, Y., Yan, L., Echeverria, V., Gašević, D., & Martinez-Maldonado, R. (2025). Generative AI in higher education: A global perspective of institutional adoption policies and guidelines. *Computers and Education: Artificial Intelligence*, *8*, 100348. https://doi.org/10.1016/j.caeai.2024.100348

Kundi, B., Morr, C. El, Gorman, R., & Dua, E. (2022). Artificial Intelligence and Bias: A Scoping Review. In *AI and Society* (1st ed.). Chapman and Hall/CRC.

Liu, G. L., Darvin, R., & Ma, C. (2024). Unpacking the role of motivation and enjoyment in AI-mediated informal digital learning of English (AI-IDLE): A mixed-method investigation in the Chinese context. *Computers in Human Behavior*, *160*, 108362. https://doi.org/10.1016/j.chb.2024.108362

Liu, G. L., Darvin, R., & Ma, C. (2025). Exploring AI-mediated informal digital learning of English (AI-IDLE): a mixed-method investigation of Chinese EFL learners' AI adoption and experiences. *Computer Assisted Language Learning*, *38*(7), 1632–1660. https://doi.org/10.1080/09588221.2024.2310288

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*. https://doi.org/10.1080/17501229.2023.2240316

Long, D., Roberts, J., Magerko, B., Holstein, K., DiPaola, D., & Martin, F. (2023). AI literacy: Finding common threads between education, design, policy, and explainability. *Extended Abstracts of the 2023 CHI Conference on Human Factors in Computing Systems*, 1–6. https://doi.org/10.1145/3544549.3573808

Mandasari, B., Basthomi, Y., Hastomo, T., Afrianto, Hamzah, I., & Aminatun, D. (2025). The snapshots of Indonesian pre-service English teachers' perspectives on integrating technology-based tools to rural schools. *Voices of English Language Education Society*, *9*(1), 42–57. https://doi.org/10.29408/veles.v9i1.27965

Marzuki, Widiati, U., Rusdin, D., Darwin, & Indrawati, I. (2023). The impact of AI writing tools on the content and organization of students' writing: EFL teachers' perspective. *Cogent Education*, *10*(2), 1–17. https://doi.org/10.1080/2331186X.2023.2236469

Rapanta, C. (2025). Postdigital citizenship education and argumentation: Can AI support students' critical literacy? In J. Zajda & A. Rapoport (Eds.), *Discourses of*

Globalisation and Citizenship Education. Globalisation, Comparative Education and Policy Research (Vol. 46, pp. 69–86). Springer. https://doi.org/10.1007/978-3-031-86145-1_5

Sawalha, G., Taj, I., & Shoufan, A. (2024). Analyzing student prompts and their effect on ChatGPT's performance. *Cogent Education*, *11*(1). https://doi.org/10.1080/2331186X.2024.2397200

Tao, Y., Viberg, O., Baker, R. S., & Kizilcec, R. F. (2024). Cultural bias and cultural alignment of large language models. *PNAS Nexus*, *3*(9). https://doi.org/10.1093/pnasnexus/pgae346

Yin, R. K. (2018). *Case study research and applications: Design and methods*. Sage Publications.

Zadorozhnyy, A., & Lee, J. S. (2025). Comparing classroom and digital settings: the role of basic psychological needs on EFL students' IDLE engagement. *ReCALL*, 1–18. https://doi.org/10.1017/S0958344025000059