Attractive : Innovative Education Journal
Vol. 8 No. 4, December 2026

# Teacher Metacognitive Scaffolding, Productive and Substitutive Generative AI Offloading, and Self-Regulated Learning in Madrasah Aliyah

**Fatimah [1]*, Rika Ariyani [2], Fithri Azni [1], Iswandi [1], Muhammad Hendriyanto [3], Miftah Hur Rahman Zh [4], Zahid Zufar At Thaariq [5]**
[1] *Universitas Merangin, Indonesia*
[2] *Universitas Islam Mambaul Ulum Jambi, Indonesia*
[3] *Universitas Negeri Surabaya, Indonesia*
[4] *Universitas Darunnajah, Indonesia*
[5] *Cukurova University, Adana, Turkiye*

fatimah.as17@gmail.com*

**ARTICLE INFO**
*Article history:*
Received
August 12, 2026
Revised
October 24, 2026
Accepted
November 17, 2026

**ABSTRACT**

Secondary school students increasingly use generative artificial intelligence (GenAI), and evidence shows it can either support or weaken their self-regulated learning (SRL). This study examined whether teacher metacognitive scaffolding (TMS) and AI literacy (AIL) are related to SRL through productive cognitive offloading (PCO) and substitutive cognitive offloading (SCO), and whether peer collaborative AI use (PCAU) strengthens the link between scaffolding and productive offloading. We conducted a cross-sectional survey of 327 students from 24 classes in eight Indonesian madrasah aliyah (Islamic senior high schools) selected through stratified multistage cluster sampling. We analysed the data using partial least squares structural equation modelling with 10,000 bootstrap subsamples. The measurement model met the reliability and validity criteria. TMS was positively related to PCO ($\beta = 0.378$) and negatively related to SCO ($\beta = -0.351$), and AIL showed the same pattern with smaller coefficients. PCO was positively related to SRL ($\beta = 0.291$), whereas SCO was negatively related to it ($\beta = -0.266$). Both forms of offloading partially mediated the relationships of TMS and AIL with SRL. PCAU strengthened the TMS–PCO path ($\beta = 0.257$), and scaffolding was not significantly related to productive offloading when peer collaboration was low. The model explained 45.8% of the variance in SRL. The findings indicate that teacher guidance is associated with SRL both directly and, for nearly half of its total effect, through the way students offload their thinking to GenAI, particularly when students verify AI outputs together.

**Keywords**: cognitive offloading; generative artificial intelligence; metacognitive scaffolding; self-regulated learning

Journal Homepage https://attractivejournal.com/index.php/aj/
This is an open access article under the CC BY SA license
https://creativecommons.org/licenses/by-sa/4.0/
@ 2026 by the authors
Published by CV. Creative Tugu Pena

## INTRODUCTION

Generative artificial intelligence (GenAI) tools such as ChatGPT and Gemini have become part of students' study routines, and their spread to secondary schools raises concerns about students' intellectual agency and the loss of productive struggle (Tao et al., 2026). Researchers have argued that the educational value of large language models depends on whether learners develop critical thinking and fact-checking strategies (Kasneci et al., 2023), and that a single metaphor cannot capture the relationship between learners and GenAI (Lodge et al., 2023). For secondary school students, whose self-regulatory skills are still developing, the pressing

question is how GenAI use affects their capacity to plan, monitor, and evaluate their own learning.

Recent evidence points to two contrasting patterns. A scoping review of 123 studies found that agency-supportive uses of GenAI were associated with self-regulated learning (SRL), self-efficacy, and reflective engagement. In contrast, agency-eroding uses were associated with cognitive offloading, over-reliance, and uncritical acceptance of AI outputs (G. Wang et al., 2026). University students who wrote with ChatGPT improved their essays without gaining more knowledge than other students, a pattern described as metacognitive laziness (Fan et al., 2025). In a field experiment with about 1,000 high school students, unrestricted access to GPT-4 raised practice performance. However, performance declined once access was removed, whereas a tutor version with learning safeguards largely avoided this harm (Bastani et al., 2025). Frequent AI use has also been linked with greater cognitive offloading and weaker critical thinking (Gerlich, 2025; Zhai et al., 2024), with reliance on AI rather than learning from it (Darvishi et al., 2024), and with procrastination and reduced independent thinking (Abbas et al., 2024; S. Zhang et al., 2024).

A second strand of research suggests that the direction of these effects depends on the support that surrounds GenAI use. Explicit metacognitive support improved university students' task strategies and self-evaluation (Xu et al., 2025), self-regulated scaffolding produced larger gains in self-efficacy than GenAI use alone (Xia et al., 2026), and a meta-analysis showed that metacognitive prompts improve both SRL activities and learning outcomes (Guo, 2022). A ChatGPT-based chatbot designed for SRL also improved secondary students' engagement and motivation (Ng et al., 2024). However, a randomised controlled trial with 371 German secondary students found that GenAI-based interventions did not outperform a standard ChatGPT condition in effort, domain knowledge, or elaboration strategy use (Fütterer et al., 2026). Intervention studies of this kind were not designed to show which cognitive processes GenAI support reaches, or fails to reach, in students' self-regulation.

Cognitive offloading offers a promising explanation. It refers to using external tools to reduce a task's cognitive demands, and it can improve immediate performance while weakening memory for the offloaded information (Grinschgl et al., 2021; Risko & Gilbert, 2016). Offloading to GenAI, however, is not a single behaviour. Gao and Zhang (2026) distinguished offloading in which learners use AI for task entry and then verify and reconstruct its outputs from offloading in which AI replaces their reasoning. Zhu et al. (2026) separated autonomous offloading, in which AI serves as a scaffold, from dependent offloading, in which core thinking is delegated. These distinctions were developed with doctoral students, university students, and knowledge workers, and they have not been tested as mechanisms linking teacher support with students' SRL. Quantitative studies have also tended to treat offloading or AI dependence as a single construct (Acosta-Enriquez et al., 2025; Iqbal et al., 2025; Liu, 2026), which cannot show whether one teaching practice strengthens one form of offloading while weakening the other. This leaves a conceptual gap, because it remains unknown whether teacher scaffolding relates differently to productive and substitutive forms of AI-mediated offloading.

A second gap concerns the mechanism. Studies have linked teacher support and AI literacy with students' motivation, procrastination, and SRL (Chiu et al., 2024; Hermita et al., 2026; Sardi et al., 2025), but they have not explained how these relationships operate through the way students offload cognitive work to GenAI. Productive and substitutive offloading may therefore be the cognitive mechanisms that connect educational support to SRL. The evidence is also limited in context. A systematic review concluded that most studies report a positive role of GenAI in SRL but warned that over-reliance may weaken independent thinking (Sardi et al., 2025), and most of the evidence discussed above comes from higher education or from Western and East Asian schools. Findings on AI literacy are also inconsistent, since AI literacy among Indonesian university students was found to reduce academic procrastination but to aggravate it under certain conditions (Hermita et al., 2026). Indonesian research on GenAI has concentrated on university students' acceptance and perceptions (Faizin et al., 2025; Rinekso et al., 2026; Sarwanti et al., 2024), including their concerns about over-dependence (Rahiem, 2026), whereas studies in madrasah remain largely qualitative (Hermawan et al., 2025) or focus on

digital learning media (Zh et al., 2024, 2026). Madrasah aliyah students follow a curriculum that combines general and Islamic subjects, often in pesantren-based settings where learning is closely tied to teacher guidance and communal study. However, researchers have not examined quantitatively how these conditions shape their GenAI use.

A third gap concerns boundary conditions, because the social side of GenAI use is underexplored. Humans and AI can jointly support group regulation (Järvelä et al., 2023), and student teams prompting ChatGPT have been observed debating and revising many of its responses (Perifanou & Economides, 2025), although a chatbot can also crowd out peer interaction (Feng, 2025). In learning cultures that emphasise relational bonds (Bong et al., 2026), checking AI outputs with peers may help translate teacher guidance into students' own verification habits, but this possibility has not been tested. Even if teacher scaffolding promotes productive AI use, its effect may therefore depend on whether students verify AI outputs together with their peers.

This study therefore tests a dual-pathway model in which teacher metacognitive scaffolding (TMS) and AI literacy (AIL) are related to SRL through productive cognitive offloading (PCO) and substitutive cognitive offloading (SCO), with peer collaborative AI use (PCAU) moderating the link between scaffolding and productive offloading. It addresses four research questions: (1) How are TMS and AIL related to PCO and SCO? (2) How are PCO and SCO related to SRL? (3) Do PCO and SCO mediate the relationships of TMS and AIL with SRL? (4) Does PCAU strengthen the relationship between TMS and PCO? The first research question addresses the conceptual gap, the second and third address the mechanism gap, and the fourth addresses the boundary-condition gap. The research context is Indonesian madrasah aliyah, and the unit of analysis is the individual student who has used GenAI for schoolwork in classes where teachers permit its use. The study also provides preliminary evidence for the reliability and construct validity of short scales that distinguish productive from substitutive offloading.

The next section develops the hypotheses, followed by the method, the results, and a discussion that relates the findings to prior research. The final section presents the conclusions and directions for future research.

## THEORETICAL FRAMEWORK AND HYPOTHESIS DEVELOPMENT

### *2.1 Self-Regulated Learning in GenAI-Supported Learning*

SRL refers to the cyclical process through which learners set goals, apply strategies, monitor their understanding, and evaluate their progress (Schunk & Greene, 2018). In phase-based models, forethought, performance, and self-reflection form a loop, which researchers have used to derive design principles for AI chatbots that support goal setting, feedback, and personalisation (Chang et al., 2023). This study focuses on the metacognitive component of SRL, namely planning, monitoring, and regulating, because it is most at stake when thinking is handed over to AI (Fan et al., 2025). Molenaar (2022) argued that AI can either take over regulation from learners or share it with them in a way that gradually returns control to the learner. The present model treats students' cognitive offloading as the point at which regulation is either retained or surrendered.

### *2.2 Teacher Metacognitive Scaffolding and Cognitive Offloading*

Scaffolding is temporary, contingent support that enables learners to perform beyond their current capability and is withdrawn as they become able to regulate their learning independently (Molenaar, 2022; Wood et al., 1976). Metacognitive scaffolding directs this support towards planning, monitoring, and evaluation. In GenAI-supported classrooms, teachers can ask students to set a learning goal before using AI, check whether they understand AI responses, compare AI outputs with textbooks, and explain in their own words what they have learned. Such prompts should encourage students to use AI as a starting point while keeping verification and reconstruction under their own control, which is the defining feature of productive offloading (Gao & Zhang, 2026). They should also discourage uncritical copying of AI answers. Teacher support has been shown to shape Grade 10 students' motivation and competence in learning with an AI chatbot (Chiu et al., 2024), and guidance on adapting chatbot responses led German secondary students to use more than twice as many adaptation prompts as unguided peers (Klar, 2025).

**H1:** *TMS is positively related to PCO.*

**H2:** *TMS is negatively related to SCO.*

### *2.3 AI Literacy and Cognitive Offloading*

AI literacy encompasses the ability to use and apply AI, to understand how it works, to detect AI-generated content, and to consider the ethical implications of its use (Carolus et al., 2023; Ng et al., 2021), and recent scales treat the evaluation of AI outputs as one of its core dimensions (Syamsudduha et al., 2026; B. Wang et al., 2023). Students who know that GenAI can produce fluent but inaccurate answers should be more likely to verify its outputs and less likely to accept them uncritically. The benefit of structured chatbot scaffolding also depends on students' AI literacy (Q. Zhang & Liang, 2026). However, AI literacy includes prompting skills that can make delegation easier, which is consistent with what Hermita et al. (2026) called the Janus face of AI literacy. Although a negative relationship between AIL and SCO therefore cannot be taken for granted, the dominant argument in the literature favours a protective role.

**H3:** *AIL is positively related to PCO.*

**H4:** *AIL is negatively related to SCO.*

### *2.4 Cognitive Offloading and Self-Regulated Learning*

Productive offloading preserves the monitoring and evaluation that constitute SRL, because students who verify AI claims against other sources or rewrite AI explanations in their own words continue to judge their own understanding. Substitutive offloading, in contrast, transfers these activities to the AI. Dependent offloading has been associated with a transfer of cognitive agency to AI and with poorer perceived cognitive outcomes. In contrast, autonomous offloading has been associated with higher intrinsic motivation and better perceived outcomes (Zhu et al., 2026). SRL ability negatively predicted AI dependence among college students (Liu, 2026), and students who worked with a large language model reported lower cognitive load but produced weaker reasoning than students who used a search engine (Stadler et al., 2024). Teacher scaffolding may also relate directly to SRL, because students can adopt the planning and monitoring routines that teachers prompt and model (Guo, 2022; Schunk & Greene, 2018).

**H5:** *PCO is positively related to SRL.*

**H6:** *SCO is negatively related to SRL.*

**H7:** *TMS is positively related to SRL.*

### *2.5 The Mediating Roles of Productive and Substitutive Offloading*

If teacher scaffolding and AI literacy shape the way students offload, and the two forms of offloading in turn relate to SRL, then part of the association between these antecedents and SRL should operate through offloading. This reasoning addresses the question left open by Fütterer et al. (2026) about why support for GenAI use does not always reach students' strategy use. The indirect effect through SCO is expected to be positive, but this does not imply that substitutive offloading is beneficial. Stronger scaffolding should be associated with less SCO, and less SCO should be associated with higher SRL, so the positive indirect effect reflects the benefit of reducing harmful offloading. The corresponding indirect effect of AIL through SCO was examined as an exploratory effect.

**H8:** *PCO mediates the relationship between TMS and SRL.*

**H9:** *SCO mediates the relationship between TMS and SRL.*

**H10:** *PCO mediates the relationship between AIL and SRL.*

### *2.6 The Moderating Role of Peer Collaborative AI Use*

Self-determination theory proposes that satisfying the need for relatedness supports the internalisation of regulation (Ryan & Deci, 2017), and autonomous motivation is the strongest predictor of adaptive learning outcomes (Howard et al., 2021). In GenAI-supported learning, teacher support has been found to satisfy students' need for relatedness more consistently than their need for autonomy (Chiu et al., 2024; Fang et al., 2026). When students compare AI answers with those of classmates, check them together, or discuss how to prompt AI, the teacher's instructions to verify and reconstruct become shared classroom practices. Learning activities that combine self-determination theory with socially shared regulation have been proposed for ChatGPT-mediated mathematical modelling (Kaya & Yıldız Durak, 2026), and Indonesian master's students have reported peer consultation as one of their strategies for

evaluating ChatGPT outputs (Werdiningsih et al., 2024). In a collectivist learning culture, such peer practices are likely to amplify the effect of teacher guidance.

**H11:** *PCAU strengthens the positive relationship between TMS and PCO.*

## METHOD

### Research Design

This study employed a quantitative, cross-sectional survey design to test the model in Figure 1. Partial least squares structural equation modelling (PLS-SEM) was selected because the study pursued both explanatory and predictive aims, the model contains two mediators and an interaction term, and four of the six scales were newly developed (Hair et al., 2022; Hair & Alamer, 2022). Because the data were collected at a single point in time, the relationships are interpreted as predictive associations rather than causal effects.

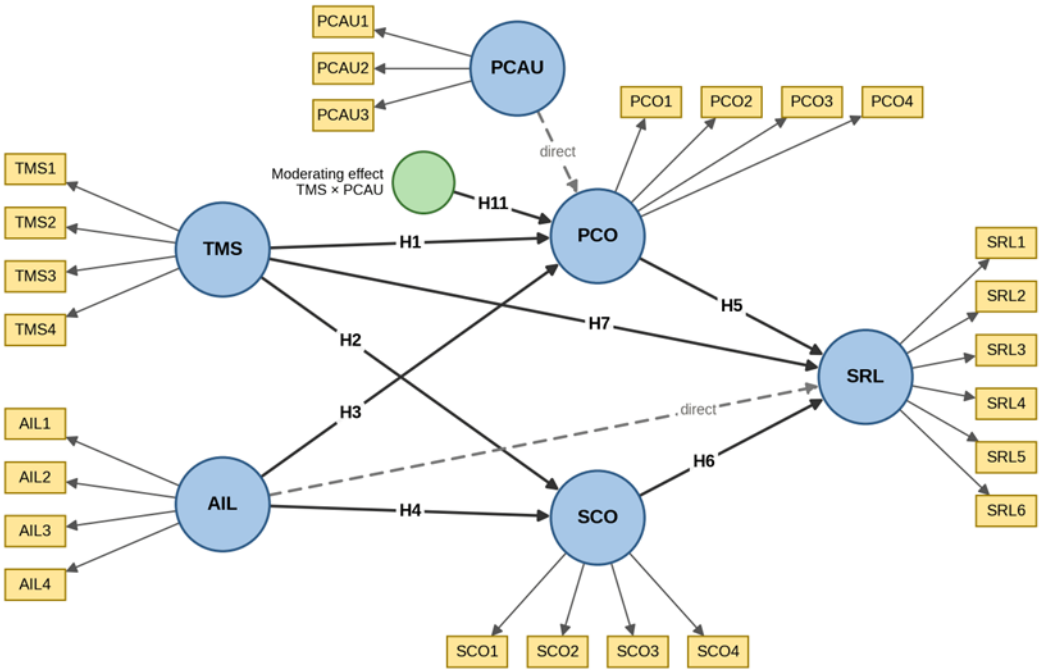



**Figure 1.** Research model and hypotheses

Figure 1 shows the six reflective constructs, their indicators, and the hypothesised paths. TMS and AIL are the exogenous constructs, PCO and SCO are the mediators, and SRL is the final endogenous construct, while PCAU moderates the TMS–PCO path through an interaction term. The paths from PCAU to PCO and from AIL to SRL were not hypothesised but were estimated so that the moderation and mediation effects could be interpreted correctly (Hair et al., 2021).

### Population and Sampling

The population comprised students in grades X to XII of madrasah aliyah who had used GenAI for schoolwork. Stratified multistage cluster random sampling was applied because a student-level sampling frame was not available, whereas lists of madrasah were available from the Education Management Information System (EMIS) of the Ministry of Religious Affairs. Madrasah were stratified by status (state and private) and type (pesantren-based and non-pesantren), and two madrasah were randomly drawn from each of the four strata. One class per grade was then randomly selected in each madrasah, and all students in the selected classes who met the inclusion criteria were invited to participate. These criteria were having used GenAI for school assignments during the previous semester, attending a class in which at least one teacher permitted or directed AI use, and providing assent together with parental consent. The inverse square root method indicated that at least 275 observations were required to detect a path coefficient of 0.15 at the 5% significance level (Hair et al., 2022).

**Participants**

A total of 347 students responded. Thirteen were excluded at screening because they had not used GenAI for schoolwork ($n$ = 8) or reported that no teacher permitted AI use ($n$ = 5), and seven failed the attention-check item. The final sample consisted of 327 students from 24 classes in eight madrasah (Table 1).

**Table 1.** Profile of respondents ($n$ = 327)

| Characteristic | Category | *n* | % |
|---|---|---|---|
| Gender | Male | 156 | 47.7 |
| | Female | 171 | 52.3 |
| Grade | X | 102 | 31.2 |
| | XI | 109 | 33.3 |
| | XII | 116 | 35.5 |
| Madrasah status | State | 177 | 54.1 |
| | Private | 150 | 45.9 |
| Madrasah type | Pesantren-based | 160 | 48.9 |
| | Non-pesantren | 167 | 51.1 |
| Residence | Boarding | 139 | 42.5 |
| | Non-boarding | 188 | 57.5 |
| GenAI use for schoolwork in the past month | Almost never | 48 | 14.7 |
| | 1–3 times a month | 89 | 27.2 |
| | 1–2 times a week | 84 | 25.7 |
| | 3–5 times a week | 64 | 19.6 |
| | Almost every day | 42 | 12.8 |
| GenAI tools used (multiple responses) | ChatGPT | 298 | 91.1 |
| | Gemini | 207 | 63.3 |
| | Meta AI | 92 | 28.1 |
| | Copilot | 47 | 14.4 |
| | Others | 10 | 3.1 |

*Note.* Percentages for GenAI tools do not sum to 100 because multiple responses were allowed.

Table 1 shows that the sample was balanced across gender, grade, madrasah status, and madrasah type. ChatGPT was the most widely used tool, followed by Gemini, and most students used GenAI for schoolwork at least once a month. This profile resembles reports that Indonesian students mainly use ChatGPT for assignments, self-learning, and language support (Sarwanti et al., 2024).

**Instruments**

All constructs were measured reflectively on a five-point scale ranging from 1 (very untrue of me) to 5 (very true of me). The TMS items were developed from the planning, monitoring, evaluation, and reflection functions of metacognitive scaffolding (Guo, 2022; Xu et al., 2025). The AIL items were adapted from the four facets of the Meta AI Literacy Scale (Carolus et al., 2023) and simplified for senior high school students. The PCO and SCO items were developed from the distinction between accountable reconstruction and replacement of reasoning (Gao & Zhang, 2026) and from studies of over-reliance on AI (Gerlich, 2025; Zhai et al., 2024). SRL was measured with six items adapted from the metacognitive self-regulation subscale of the Motivated Strategies for Learning Questionnaire (MSLQ; Pintrich et al., 1991), covering planning, monitoring, and regulating. Short forms of the MSLQ have shown acceptable validity among senior secondary students (F. Wang et al., 2023), although this subscale's reliability has varied across samples (de Araujo et al., 2023). The PCAU items were developed from the relatedness principle of self-determination theory and from research on

socially shared regulation (Järvelä et al., 2023; Ryan & Deci, 2017). Adapted items were translated into Indonesian and back-translated. Two experts, whose combined expertise covered educational psychology, educational technology, and madrasah teaching, reviewed all items for relevance, clarity, and fit with the construct definitions, and three students tested the wording through cognitive interviews. The final questionnaire contained 25 substantive items and one attention-check item (Appendix A).

The four new scales (TMS, PCO, SCO, and PCAU) were developed in four steps. First, each construct was given a conceptual definition based on the literature cited above (Table A2 in Appendix A). Second, items were written to cover the content of each definition, and the PCO and SCO items described the same situation of AI use for schoolwork so that the two scales differed only in whether the student retained or surrendered verification and reconstruction. Third, the expert review and cognitive interviews were used to check content relevance and comprehension. Fourth, the scales were evaluated in the main sample through indicator loadings, internal consistency, convergent validity, discriminant validity based on the HTMT ratio, and correlations with an external criterion. Because the expert panel was small and the scales were developed and evaluated in the same sample, these steps provide preliminary rather than conclusive evidence of construct validity, and cross-validation in an independent sample remains necessary (Hair et al., 2022).

**Data Collection**

The questionnaire was administered online during class hours under researcher supervision from 21 January 2026 to 30 August 2026. We applied procedural remedies for common method bias (Podsakoff et al., 2024): responses were anonymous, construct labels were not displayed, the SRL items preceded the predictor items, and the PCO and SCO items were interleaved.

**Data Analysis**

The analysis followed the two-stage PLS-SEM procedure (Hair et al., 2022; Sarstedt et al., 2022). The model was estimated in SmartPLS 4.1 (Ringle et al., 2024) using path weighting, standardised data, and a stop criterion of $10^{-7}$. The measurement model was assessed through loadings (≥ 0.708), Cronbach's alpha, $\rho_A$ and $\rho_c$ (0.70–0.95), average variance extracted (AVE ≥ 0.50), and the heterotrait–monotrait ratio (HTMT < 0.85) with bootstrap confidence intervals, which was preferred over the Fornell–Larcker criterion (Guenther et al., 2023; Hair & Alamer, 2022). The structural model was assessed through inner VIF (< 3), path coefficients tested through bootstrapping with 10,000 subsamples (bootstrap t-tests and percentile confidence intervals), $R^2$, $f^2$, and PLSpredict with 10 folds and 10 repetitions (Hair, 2021). We tested specific indirect effects using the same bootstrap procedure and classified the mediation type based on the significance and sign of the direct and indirect effects (Hair et al., 2022). Moderation was tested using a two-stage approach, followed by simple slopes at ±1 SD of PCAU and an index of moderated mediation (Becker et al., 2023; Cheah et al., 2021). Gender, grade, and GenAI use frequency served as control variables for SRL. We also examined common method bias using full collinearity VIFs against a threshold of 3.3 (F. Kock et al., 2021; N. Kock & Dow, 2025), and computed intraclass correlations and design effects because students were nested in classes (Hox et al., 2017; McNeish, 2026). Finally, we compared pesantren-based and non-pesantren madrasah through permutation-based multigroup analysis after assessing measurement invariance with the MICOM procedure (Cheah et al., 2023; Hair et al., 2024).

## RESULTS AND DISCUSSION

### *4.1 Preliminary Analyses*

The 25 substantive items had no missing values. Item means ranged from 2.71 to 3.91, skewness from −0.79 to 0.33, and excess kurtosis from −0.96 to −0.04, indicating no severe departure from normality. Full collinearity VIFs ranged from 1.063 to 1.831, well below 3.3, suggesting that common method bias was unlikely to distort the results (N. Kock & Dow, 2025). Class-level intraclass correlations ranged from −0.030 to 0.042 with design effects not exceeding 1.54, and school-level intraclass correlations did not exceed 0.016 (design effects ≤ 1.65). Given these modest values, a single-level model was retained (McNeish, 2026).

### *4.2 Measurement Model*

**Table 2.** Measurement model: loadings, reliability, and convergent validity

| Construct | Item | Loading | α | $\rho_A$ | $\rho_c$ | AVE |
|---|---|---|---|---|---|---|
| Teacher metacognitive scaffolding (TMS) | TMS1 | 0.847 | 0.865 | 0.870 | 0.908 | 0.711 |
| | TMS2 | 0.862 | | | | |
| | TMS3 | 0.797 | | | | |
| | TMS4 | 0.866 | | | | |
| AI literacy (AIL) | AIL1 | 0.835 | 0.857 | 0.865 | 0.903 | 0.701 |
| | AIL2 | 0.842 | | | | |
| | AIL3 | 0.870 | | | | |
| | AIL4 | 0.800 | | | | |
| Peer collaborative AI use (PCAU) | PCAU1 | 0.763 | 0.792 | 0.880 | 0.871 | 0.694 |
| | PCAU2 | 0.888 | | | | |
| | PCAU3 | 0.843 | | | | |
| Productive cognitive offloading (PCO) | PCO1 | 0.877 | 0.887 | 0.887 | 0.922 | 0.746 |
| | PCO2 | 0.842 | | | | |
| | PCO3 | 0.871 | | | | |
| | PCO4 | 0.866 | | | | |
| Substitutive cognitive offloading (SCO) | SCO1 | 0.842 | 0.868 | 0.877 | 0.910 | 0.717 |
| | SCO2 | 0.824 | | | | |
| | SCO3 | 0.839 | | | | |
| | SCO4 | 0.881 | | | | |
| Self-regulated learning (SRL) | SRL1 | 0.853 | 0.904 | 0.905 | 0.926 | 0.677 |
| | SRL2 | 0.834 | | | | |
| | SRL3 | 0.813 | | | | |
| | SRL4 | 0.823 | | | | |
| | SRL5 | 0.815 | | | | |
| | SRL6 | 0.798 | | | | |

*Note.* All loadings were significant at $p < .001$ (10,000 bootstrap subsamples). Item wording is given in Appendix A.

Table 2 shows that all loadings exceeded 0.708 (0.763–0.888). Cronbach's alpha ranged from 0.792 to 0.904, $\rho_A$ from 0.865 to 0.905, and $\rho_c$ from 0.871 to 0.926, and AVE values ranged from 0.677 to 0.746, confirming convergent validity. Outer VIFs ranged from 1.562 to 2.547. The reliability of the six SRL items (α = 0.904) is noteworthy because de Araujo et al. (2023) found that the full metacognitive self-regulation subscale did not reach acceptable reliability in a Brazilian university sample. The newly developed scales also showed high reliability. High reliability alone does not establish construct validity, so the results for the new scales are treated as preliminary evidence, supported by the discriminant validity and criterion-related results reported below.

**Table 3.** Discriminant validity based on the heterotrait–monotrait ratio (HTMT)

| Construct | TMS | AIL | PCAU | PCO | SCO |
|---|---|---|---|---|---|
| AIL | 0.262 [0.385] | | | | |
| PCAU | 0.154 [0.277] | 0.264 [0.395] | | | |
| PCO | 0.482 [0.580] | 0.282 [0.406] | 0.131 [0.262] | | |
| SCO | 0.470 [0.569] | 0.369 [0.477] | 0.136 [0.268] | 0.251 [0.364] | |
| SRL | 0.568 [0.658] | 0.428 [0.528] | 0.095 [0.223] | 0.547 [0.627] | 0.537 [0.628] |

*Note.* Values in brackets are the upper bounds of the 95% percentile bootstrap confidence intervals (10,000 subsamples).

Table 3 shows that all HTMT values were below 0.85, the highest being between TMS and SRL (0.568), and no upper bound of the 95% bootstrap confidence intervals exceeded 0.658, which supports discriminant validity (Hair & Alamer, 2022). The HTMT between PCO and SCO was only 0.251, indicating that productive and substitutive offloading are empirically distinct constructs rather than opposite ends of one continuum. This result is consistent with Zhu et al. (2026), who treated dependent and autonomous offloading as separate constructs with different relationships to perceived cognitive outcomes.

***4.3 Structural Model***

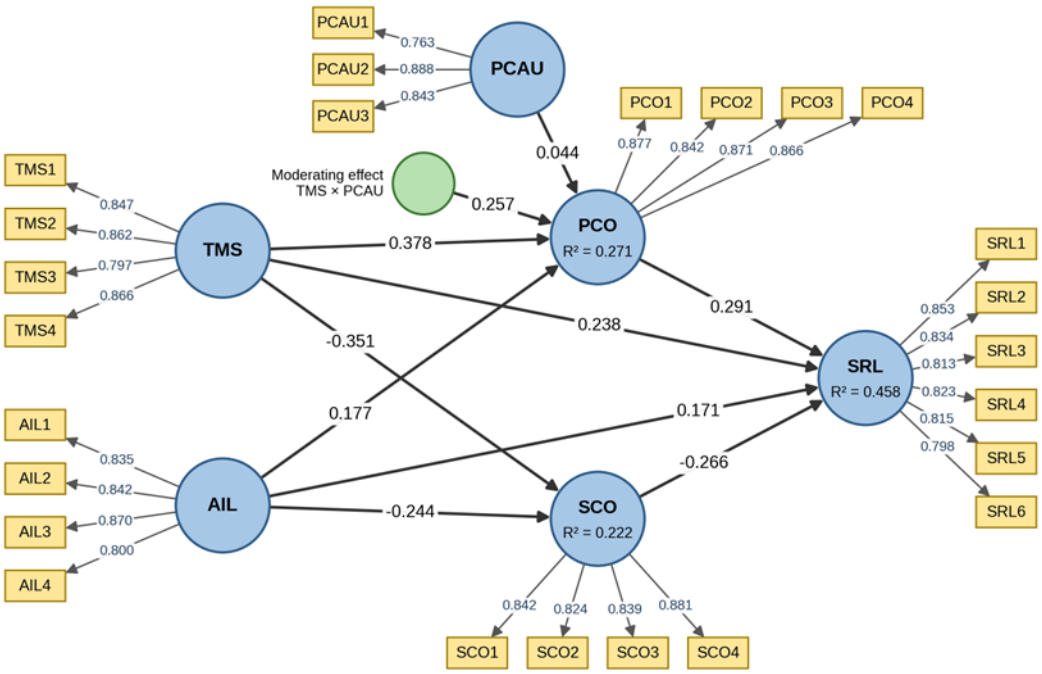



**Figure 2.** Results of the PLS algorithm: outer loadings, path coefficients, and $R^2$

Figure 2 displays the outer loadings, path coefficients, and $R^2$ values. Inner VIFs ranged from 1.009 to 1.411, so collinearity was not a concern. The model explained 27.1% of the variance in PCO, 22.2% in SCO, and 45.8% in SRL (adjusted $R^2$ = 0.262, 0.218, and 0.446). By common PLS-SEM benchmarks, the explained variance in SRL approaches the moderate level, while that of the two mediators lies around the weak level (Hair et al., 2022). Table 4 presents the tests of the direct effects.

**Table 4.** Results of hypothesis testing for direct effects

| Hypothesis | Path | β | *SD* | *t* | *p* | 95% CI | $f^2$ | Decision |
|---|---|---|---|---|---|---|---|---|
| H1 | TMS → PCO | 0.378 | 0.050 | 7.574 | < .001 | [0.277, 0.473] | 0.184 | Supported |
| H2 | TMS → SCO | −0.351 | 0.047 | 7.408 | < .001 | [−0.441, −0.257] | 0.150 | Supported |
| H3 | AIL → PCO | 0.177 | 0.051 | 3.432 | < .001 | [0.072, 0.274] | 0.039 | Supported |
| H4 | AIL → SCO | −0.244 | 0.046 | 5.265 | < .001 | [−0.335, −0.153] | 0.072 | Supported |
| H5 | PCO → SRL | 0.291 | 0.041 | 7.054 | < .001 | [0.209, 0.372] | 0.124 | Supported |
| H6 | SCO → SRL | −0.266 | 0.048 | 5.551 | < .001 | [−0.357, −0.171] | 0.101 | Supported |
| H7 | TMS → SRL | 0.238 | 0.049 | 4.849 | < .001 | [0.141, 0.335] | 0.074 | Supported |
| H11 | TMS × PCAU → PCO | 0.257 | 0.058 | 4.430 | < .001 | [0.114, 0.344] | 0.090 | Supported |
| – | PCAU → PCO | 0.044 | 0.053 | 0.820 | .412 | [−0.062, 0.152] | 0.002 | – |
| – | AIL → SRL | 0.171 | 0.048 | 3.553 | < .001 | [0.073, 0.263] | 0.040 | – |
| Control | Gender → SRL | 0.003 | 0.043 | 0.072 | .943 | [−0.080, 0.088] | 0.000 | – |
| Control | Grade → SRL | 0.079 | 0.040 | 1.965 | .049 | [−0.0003, 0.158] | 0.011 | – |
| Control | GenAI use frequency → SRL | −0.011 | 0.047 | 0.230 | .818 | [−0.102, 0.081] | 0.000 | – |

*Note.* $R^2$ (adjusted $R^2$): PCO = 0.271 (0.262), SCO = 0.222 (0.218), SRL = 0.458 (0.446). t and two-tailed p values are based on bootstrap standard errors; CI = percentile bootstrap confidence interval (10,000 subsamples).

As shown in Table 4, all seven hypothesised direct effects (H1–H7) were statistically significant in the expected directions, and the hypothesised interaction effect (H11) was also significant. TMS was positively related to PCO (H1: β = 0.378, $t$ = 7.574) and negatively related to SCO (H2: β = −0.351, $t$ = 7.408), with medium effect sizes. AIL was positively related to PCO (H3: β = 0.177, $t$ = 3.432) and negatively related to SCO (H4: β = −0.244, $t$ = 5.265), with small effect sizes. PCO was positively related to SRL (H5: β = 0.291, $t$ = 7.054), SCO was negatively related to SRL (H6: β = −0.266, $t$ = 5.551), and TMS retained a positive direct relationship with SRL (H7: β = 0.238, $t$ = 4.849), all at $p$ < .001. The non-hypothesised path from AIL to SRL was significant (β = 0.171), whereas the conditional effect of PCAU on PCO at average TMS was not (β = 0.044, $p$ = .412). Gender and GenAI use frequency were not related to SRL. The grade coefficient was very small (β = 0.079, $f^2$ = 0.011) and borderline, because its bootstrap $t$-test gave $p$ = .049, whereas its 95% percentile CI [−0.0003, 0.158] narrowly included zero. Given this disagreement at the margin and the negligible effect size, grade is not interpreted as a substantive predictor of SRL. The opposite signs of the two offloading paths mirror the dual pattern described by G. Wang et al. (2026), in which augmentation-oriented GenAI use supports learner agency whereas replacement-oriented use erodes it.

The PLSpredict results showed $Q^2_{\text{predict}}$ values above zero for all 14 indicators of the endogenous constructs (0.113–0.220) and construct-level $Q^2_{\text{predict}}$ values of 0.186 for PCO, 0.211 for SCO, and 0.310 for SRL. The PLS-SEM model produced a lower root mean squared error than the linear regression benchmark for all 14 indicators, indicating high predictive power (Hair, 2021).

### *4.4 Mediation and Moderation Analyses*

**Table 5.** Specific indirect effects and conditional effects of teacher metacognitive scaffolding

| Hypothesis | Effect | β | *SD* | *t* | *p* | 95% CI | Decision |
|---|---|---|---|---|---|---|---|
| H8 | TMS → PCO → SRL | 0.110 | 0.022 | 4.943 | < .001 | [0.070, 0.157] | Supported |
| H9 | TMS → SCO → SRL | 0.093 | 0.022 | 4.306 | < .001 | [0.054, 0.138] | Supported |
| H10 | AIL → PCO → SRL | 0.051 | 0.017 | 3.054 | .002 | [0.020, 0.087] | Supported |
| – | AIL → SCO → SRL | 0.065 | 0.018 | 3.656 | < .001 | [0.034, 0.102] | – |
| – | Total indirect effect of TMS on SRL | 0.203 | 0.031 | 6.659 | < .001 | [0.147, 0.267] | – |
| – | Total indirect effect of AIL on SRL | 0.116 | 0.024 | 4.743 | < .001 | [0.071, 0.166] | – |
| H11 (slope) | TMS → PCO at low PCAU (−1 SD) | 0.120 | 0.084 | 1.443 | .149 | [−0.017, 0.313] | – |
| H11 (slope) | TMS → PCO at high PCAU (+1 SD) | 0.635 | 0.069 | 9.227 | < .001 | [0.480, 0.749] | – |
| – | Index of moderated mediation (TMS → PCO → SRL by PCAU) | 0.075 | 0.019 | 3.844 | < .001 | [0.031, 0.109] | – |

*Note.* The direct effects of TMS and AIL on SRL were significant and had the same sign as the indirect effects, indicating complementary partial mediation. The positive indirect effects through SCO are products of two negative paths and reflect the benefit of lower substitutive offloading, not a benefit of SCO itself.

Table 5 shows that all hypothesised indirect effects were significant. The indirect effect of TMS on SRL was 0.110 through PCO (H8) and 0.093 through SCO (H9), and the indirect effect of AIL on SRL through PCO was 0.051 (H10). The non-hypothesised indirect effect of AIL through SCO was also significant (0.065). Because the direct effects of TMS and AIL on SRL were significant and had the same sign as the indirect effects, both relationships represent complementary partial mediation (Hair et al., 2022). The two offloading pathways together transmitted 46.1% of the total effect of TMS on SRL and 40.4% of the total effect of AIL.

The interaction term TMS × PCAU was positively related to PCO (H11: β = 0.257, $t$ = 4.430, $p$ < .001, $f^2$ = 0.090). The simple slope analysis showed that TMS was not significantly related to PCO when PCAU was one standard deviation below the mean (β = 0.120), whereas the relationship was significant at the mean (β = 0.378) and considerably stronger one standard deviation above the mean (β = 0.635). The significant index of moderated mediation (0.075) further indicates that the indirect effect of TMS on SRL through PCO increases with peer collaborative AI use. This pattern is consistent with the observation that student teams prompting ChatGPT together revised about half of its responses (Perifanou & Economides, 2025).

### *4.5 Multigroup Analysis and Robustness Checks*

The MICOM procedure established compositional invariance for all constructs, and the differences in composite means and variances between pesantren-based ($n$ = 160) and non-pesantren ($n$ = 167) madrasah fell within the 95% permutation intervals. Full measurement invariance therefore allowed the path coefficients to be compared (Cheah et al., 2023).

**Table 6.** Multigroup analysis of pesantren-based ($n$ = 160) and non-pesantren ($n$ = 167) madrasah

| Path | Pesantren-based | Non-pesantren | Difference | p (permutation) |
|---|---|---|---|---|
| TMS → PCO | 0.460* | 0.303* | 0.156 | .124 |
| AIL → PCO | 0.056 | 0.245* | −0.189 | .086 |
| PCAU → PCO | 0.080 | 0.023 | 0.057 | .533 |
| TMS → SCO | −0.366* | −0.338* | −0.028 | .771 |
| AIL → SCO | −0.234* | −0.251* | 0.017 | .851 |
| PCO → SRL | 0.229* | 0.286* | −0.057 | .486 |
| SCO → SRL | −0.135 | −0.348* | 0.213 | .029 |
| TMS → SRL | 0.440* | 0.119 | 0.321 | < .001 |
| AIL → SRL | 0.065 | 0.243* | −0.178 | .039 |

*Note.* Values are standardised path coefficients estimated without the interaction term and control variables. * $p$ < .05 within the group (5,000 bootstrap subsamples). p values for differences are based on 5,000 permutations.

Table 6 indicates that three paths differed significantly between the groups. The direct relationship between TMS and SRL was much stronger in pesantren-based madrasah (β = 0.440 versus 0.119), the negative relationship between SCO and SRL was weaker there (β = −0.135 versus −0.348), and the relationship between AIL and SRL was stronger in non-pesantren madrasah (β = 0.243 versus 0.065). These differences are in line with qualitative evidence that teacher competence and Islamic ethical values are central concerns when AI is integrated into madrasah aliyah (Hermawan et al., 2025).

Robustness checks supported the findings. Adding the control variables as predictors of PCO and SCO changed no hypothesised path by 0.015 or more, and after the non-hypothesised AIL–SRL path was removed, all hypothesised paths remained significant with unchanged signs ($R^2$ for SRL = 0.437). SRL scores correlated with students' self-reported report-card averages ($r$ = .589, $p$ < .001, $n$ = 297), which also correlated with PCO ($r$ = .346) and SCO ($r$ = −.319). These correlations provide criterion-related evidence, in line with Indonesian findings that SRL is associated with academic achievement in digital learning environments (Sutarni et al., 2021).

## DISCUSSION

### *5.1 Offloading as the Mechanism Linking Support and Self-Regulation*

Regarding the second and third research questions, PCO and SCO were related to SRL in opposite directions, yet both transmitted positive indirect effects of teacher scaffolding and AI literacy, because these antecedents raised productive offloading and lowered substitutive offloading. The positive indirect effect through SCO should therefore be read as a protective pathway. Students who receive stronger scaffolding report less substitutive offloading, and less substitutive offloading is associated with stronger SRL, while substitutive offloading itself remains harmful, as its negative relationship with SRL shows. This pattern agrees with the dual-pattern account of GenAI use in higher education (G. Wang et al., 2026) and extends the distinction between autonomous and dependent offloading (Zhu et al., 2026) to Islamic senior high school students. It also fits the warning about metacognitive laziness (Fan et al., 2025). The negative SCO–SRL relationship is consistent with the negative association between SRL and AI dependence reported by Liu (2026), and the positive role of PCO corresponds with evidence that offloading to GenAI, alongside shared metacognition, was positively related to preservice teachers' achievement (Iqbal et al., 2025).

These results help explain why some GenAI interventions in secondary schools have not improved strategy use (Fütterer et al., 2026). Nearly half of the total effect of teacher scaffolding on SRL was transmitted through the offloading pathways, which suggests that an intervention that makes GenAI more engaging without changing whether students verify and reconstruct its outputs may leave self-regulation largely unchanged. This interpretation agrees with evidence that GenAI without learning safeguards can harm later performance, whereas a version that prompts students' own reasoning largely avoids this harm (Bastani et al., 2025). Najib and

Darnoto (2026) found that most Islamic education students used AI as a learning aid and verified information through scholarly sources, and Salam et al. (2026) showed that AI over-reliance was negatively related to student integrity, whereas intentional ChatGPT use was positively related to it.

### *5.2 The Role of Teacher Metacognitive Scaffolding*

Regarding the first research question, teacher scaffolding was the strongest antecedent of both forms of offloading. Its opposite relationships with PCO and SCO support the argument that metacognitive prompts shift cognitive work back towards the learner (Molenaar, 2022), in line with experimental evidence that metacognitive prompts and support improve SRL activities (Guo, 2022; Xu et al., 2025) and with the finding that teacher support shapes secondary students' motivation to learn with AI chatbots (Chiu et al., 2024). The direct TMS–SRL relationship further suggests that teachers' prompts act as models of regulation in their own right. The prompt to set a learning goal before using AI is particularly relevant, because awareness of a learning goal can reduce the memory costs of offloading under some conditions (Grinschgl et al., 2021).

### *5.3 The Protective but Limited Role of AI Literacy*

AI literacy was positively related to PCO and negatively related to SCO, indicating a protective rather than an exacerbating role in this sample. This finding differs from the mixed pattern reported for Indonesian university students (Hermita et al., 2026). One explanation is that the AIL items emphasised awareness of AI errors, detection of AI-generated text, and academic honesty, all of which support verification. Another is that the more structured guidance in madrasah may channel students' technical skills towards checking rather than delegating. The effects of AIL were nevertheless small, and the AIL–PCO relationship was not significant in pesantren-based madrasah, although it did not differ significantly from that in non-pesantren madrasah. Consistent with evidence that structured chatbot scaffolding can compensate for lower AI literacy (Q. Zhang & Liang, 2026), these results suggest that AI literacy alone is insufficient and that teacher guidance remains the stronger lever. Integrating Islamic principles such as tabayyun, which requires verifying information before accepting it, and amanah into AI literacy instruction may strengthen the evaluative component that supports productive offloading (Ahmad et al., 2025; Syamsudduha et al., 2026).

### *5.4 Peer Collaboration as a Condition for Productive Offloading*

Regarding the fourth research question, peer collaborative AI use was not significantly related to PCO at average levels of scaffolding, but it strengthened the relationship between teacher scaffolding and productive offloading, and scaffolding was not significantly related to PCO when peer collaboration was low. Teacher guidance thus appears to become a student habit mainly when students practise verification together. This finding supports the view that humans and AI can share the regulation of learning within groups (Järvelä et al., 2023) and that relatedness supports the internalisation of externally prompted behaviour (Ryan & Deci, 2017). The absence of a significant conditional effect of PCAU is also informative, because collaboration without guidance may simply multiply AI-centred interaction (Feng, 2025). Similarly, Indonesian senior high school students who completed writing tasks in small groups reported using AI mainly for planning and drafting, and AI did not improve their papers on all indicators (Utami et al., 2023), which suggests that collaboration needs to be combined with explicit verification tasks.

### *5.5 Differences between Pesantren-Based and Non-Pesantren Madrasah*

The multigroup analysis showed a much stronger direct relationship between teacher scaffolding and SRL and a weaker negative relationship between substitutive offloading and SRL in pesantren-based madrasah. A plausible explanation is that these madrasah maintain structured study routines and close teacher supervision, so that teacher guidance shapes students' regulation directly and partly buffers the consequences of occasional substitutive use. This explanation agrees with evidence that teacher support in GenAI-supported learning strengthens students' relatedness and competence (Fang et al., 2026) and that the moral-religious authority of pesantren leaders shapes value-based collaboration among teachers (Listyasari et al., 2026). In non-pesantren madrasah, students' own AI literacy played a larger

role in SRL. Because madrasah type is a school-level attribute with only four schools per group, these differences may also reflect characteristics of individual schools, and require confirmation with more schools.

### *5.6 Theoretical and Practical Implications*

Theoretically, this study refines the concept of cognitive offloading in education by showing that productive and substitutive offloading are distinct constructs with opposite relationships to SRL. It qualifies the view that GenAI simply helps or harms self-regulation, since the data support a conditional dual-pathway model in which teacher scaffolding and AI literacy relate to SRL through the balance between the two forms of offloading and the productive pathway depends on peer collaboration. The model thereby connects the scaffolding tradition in SRL research (Molenaar, 2022) with recent work on offloading to GenAI (Gao & Zhang, 2026; Zhu et al., 2026).

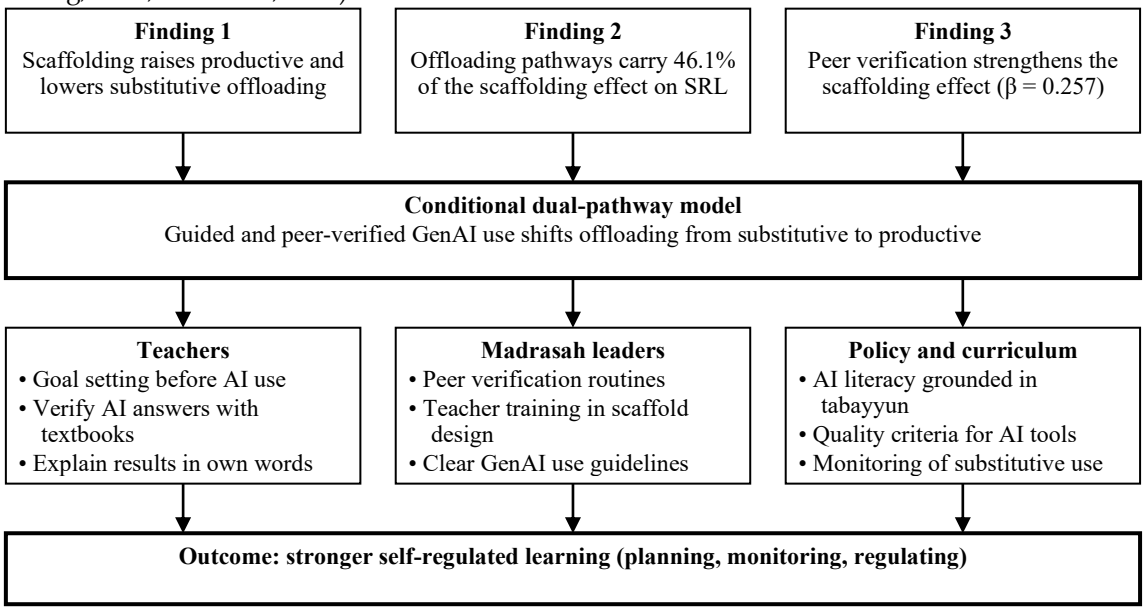



**Figure 3.** Implications of the study for GenAI-supported learning in madrasah aliyah

Figure 3 translates the empirical model into design levers for teachers and implications for madrasah leaders and policymakers. For teachers, GenAI assignments should include explicit prompts for goal setting, verification, and explanation, in line with recommendations to use GenAI as a scaffold rather than a source of direct answers (Tao et al., 2026). Verification should be organised as a peer activity, supported by tools that align ChatGPT activities with SRL phases (Chiu, 2024). Because perceived behavioural control was the strongest predictor of Islamic education teachers' intention to use AI (Febrianto et al., 2025) and teachers' AI literacy is closely linked to their acceptance of AI (Batubara et al., 2025), professional development should train teachers to design such scaffolds. Online AI use can also be followed by offline reflection grounded in local values (Andriana & Khalid, 2026), and madrasah should select AI tools against quality criteria that include data governance and equity (Haditia et al., 2026).

### *5.7 Limitations*

This study has several limitations. The cross-sectional design does not permit causal conclusions, and the reverse direction, in which students with stronger SRL choose more productive offloading, is plausible (Liu, 2026). The constructs were self-reported, and the SCO items may be affected by social desirability despite anonymity. The validity evidence for the newly developed scales, particularly PCO and SCO, is preliminary, because only two experts reviewed the items and the scales were developed and evaluated in the same sample. Confirmatory factor analysis in an independent sample and tests of measurement invariance in other groups and settings are needed. Although the intraclass correlations were small, students were nested in classes, and studies with more clusters should apply multilevel or cluster-robust estimation (McNeish, 2026). Finally, generalisation beyond the eight sampled madrasah requires caution.

## CONCLUSION

This study concludes that teacher metacognitive scaffolding is associated with the self-regulated learning of madrasah aliyah students both directly and through the way students offload cognitive work to GenAI. (1) Teacher metacognitive scaffolding and AI literacy were positively related to productive cognitive offloading and negatively related to substitutive cognitive offloading, with teacher scaffolding showing the stronger effects. (2) Productive offloading was positively related to SRL, whereas substitutive offloading was negatively related to it, and teacher scaffolding also retained a direct positive relationship with SRL. (3) The two forms of offloading partially mediated the relationships of teacher scaffolding and AI literacy with SRL, transmitting about 40% to 46% of their total effects. (4) Peer collaborative AI use strengthened the relationship between teacher scaffolding and productive offloading, and scaffolding was not significantly related to productive offloading when peer collaboration was low.

Future research should test the dual-pathway model with longitudinal or experimental designs, for example by comparing individual and peer-based verification scaffolds. Studies should also combine self-reports with behavioural data from students' interactions with AI and examine whether the model holds at other levels of education, in other regions of Indonesia, and in other collectivist learning cultures. Future studies should also cross-validate the productive and substitutive offloading scales in independent samples.

## REFERENCES


Abbas, M., Jam, F. A., & Khan, T. I. (2024). Is it harmful or helpful? Examining the causes and consequences of generative AI usage among university students. *International Journal of Educational Technology in Higher Education, 21*(1), Article 10. https://doi.org/10.1186/s41239-024-00444-7

Acosta-Enriquez, B. G., Arbulú Ballesteros, M. A., Guzman Valle, M. de los A., Morales Angaspilco, J. E., Aquino Lalupú, J. del R., Bravo Jaico, J. L., Germán Reyes, N. C., Alarcón García, R. E., & Janampa Castillo, W. E. (2025). The mediating role of academic stress, critical thinking and performance expectations in the influence of academic self-efficacy on AI dependence: Case study in college students. *Computers and Education: Artificial Intelligence, 8*, Article 100381. https://doi.org/10.1016/j.caeai.2025.100381

Ahmad, Y. S., Abdulrasaq, Y., Ali, A. M., & Ali, F. (2025). Integrating Islamic ethical principles into AI education and character building in digital society. *Tunjuk Ajar: Journal of Education and Culture, 1*(2), 130–146. https://doi.org/10.64929/ta.v1i2.38

Andriana, A. R., & Khalid, I. (2026). Integrating digital ethics and etiquette into digital literacy: A hybrid learning framework for elementary schools in Indonesia. *Tunjuk Ajar: Journal of Education and Culture, 2*(1), 60–75. https://doi.org/10.64929/ta.v2i1.48

Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakçı, Ö., & Mariman, R. (2025). Generative AI without guardrails can harm learning: Evidence from high school mathematics. *Proceedings of the National Academy of Sciences of the United States of America, 122*(26), Article e2422633122. https://doi.org/10.1073/pnas.2422633122

Batubara, H. H., Siddiq, M., Risfianti, A., & Saharani, A. (2025). Bridging artificial intelligence literacy and technology acceptance among elementary school teachers: Evidence from a structural equation model. *Al Ibtida: Jurnal Pendidikan Guru MI, 12*(1), 14–31. https://doi.org/10.24235/al.ibtida.snj.v12i1.20259

Becker, J.-M., Cheah, J.-H., Gholamzade, R., Ringle, C. M., & Sarstedt, M. (2023). PLS-SEM's most wanted guidance. *International Journal of Contemporary Hospitality Management, 35*(1), 321–346. https://doi.org/10.1108/IJCHM-04-2022-0474

Bong, M., Jeong, D., Park, S., Suh, J., Kim, S. G., Park, J. C., & Kim, J. (2026). Cultural insights into the motivation paradox of East Asian learners. *Educational Psychology Review, 38*(1), Article 41. https://doi.org/10.1007/s10648-026-10140-9

Carolus, A., Koch, M. J., Straka, S., Latoschik, M. E., & Wienrich, C. (2023). MAILS—Meta AI literacy scale: Development and testing of an AI literacy questionnaire based on well-founded competency models and psychological change- and meta-competencies.

*Computers in Human Behavior: Artificial Humans, 1*(2), Article 100014. https://doi.org/10.1016/j.chbah.2023.100014

Chang, D. H., Lin, M. P.-C., Hajian, S., & Wang, Q. Q. (2023). Educational design principles of using AI chatbot that supports self-regulated learning in education: Goal setting, feedback, and personalization. *Sustainability, 15*(17), Article 12921. https://doi.org/10.3390/su151712921

Cheah, J.-H., Amaro, S., & Roldán, J. L. (2023). Multigroup analysis of more than two groups in PLS-SEM: A review, illustration, and recommendations. *Journal of Business Research, 156*, Article 113539. https://doi.org/10.1016/j.jbusres.2022.113539

Cheah, J.-H., Nitzl, C., Roldán, J. L., Cepeda-Carrión, G., & Gudergan, S. P. (2021). A primer on the conditional mediation analysis in PLS-SEM. *ACM SIGMIS Database: The DATABASE for Advances in Information Systems, 52*(SI), 43–100. https://doi.org/10.1145/3505639.3505645

Chiu, T. K. F. (2024). A classification tool to foster self-regulated learning with generative artificial intelligence by applying self-determination theory: A case of ChatGPT. *Educational Technology Research and Development, 72*(4), 2401–2416. https://doi.org/10.1007/s11423-024-10366-w

Chiu, T. K. F., Moorhouse, B. L., Chai, C. S., & Ismailov, M. (2024). Teacher support and student motivation to learn with Artificial Intelligence (AI) based chatbot. *Interactive Learning Environments, 32*(7), 3240–3256. https://doi.org/10.1080/10494820.2023.2172044

Darvishi, A., Khosravi, H., Sadiq, S., Gašević, D., & Siemens, G. (2024). Impact of AI assistance on student agency. *Computers & Education, 210*, Article 104967. https://doi.org/10.1016/j.compedu.2023.104967

de Araujo, J., Gomes, C. M. A., & Jelihovschi, E. G. (2023). The factor structure of the Motivated Strategies for Learning Questionnaire (MSLQ): New methodological approaches and evidence. *Psicologia: Reflexão e Crítica, 36*, Article 38. https://doi.org/10.1186/s41155-023-00280-0

Faizin, N., Alfan, M., Basid, A., Ramadhan, M. R., Panatik, S. A., & Kawakip, A. N. (2025). Muslim students' acceptance of artificial intelligence in Islamic religious education: An extended TAM approach. *Discover Education, 4*, Article 304. https://doi.org/10.1007/s44217-025-00767-1

Fan, Y., Tang, L., Le, H., Shen, K., Tan, S., Zhao, Y., Shen, Y., Li, X., & Gašević, D. (2025). Beware of metacognitive laziness: Effects of generative artificial intelligence on learning motivation, processes, and performance. *British Journal of Educational Technology, 56*(2), 489–530. https://doi.org/10.1111/bjet.13544

Fang, X., Yang, M., Zhou, X., Li, Y., & Chiu, T. K. F. (2026). Using self-determination theory to explain how teacher support enhances student engagement and higher-order thinking in simulation-based learning with GenAI. *Journal of Educational Computing Research, 64*(4), 850–882. https://doi.org/10.1177/07356331261421079

Febrianto, A., Habibi, A., & Sofyan, S. (2025). AI readiness in 21st-century Islamic education: Teachers behavioral intentions and perceived control. *Tekno-Pedagogi: Jurnal Teknologi Pendidikan, 15*(2), 36–53. https://online-journal.unja.ac.id/pedagogi/article/view/42852

Feng, S. (2025). Group interaction patterns in generative AI-supported collaborative problem solving: Network analysis of the interactions among students and a GAI chatbot. *British Journal of Educational Technology, 56*(5), 2125–2145. https://doi.org/10.1111/bjet.13611

Fütterer, T., Bardach, L., Kuhn, J., Keller, S. D., & Gerjets, P. (2026). Enhancing school students' self-regulated learning through generative AI support: A randomized controlled trial. *Educational Psychology Review, 38*(1), Article 42. https://doi.org/10.1007/s10648-026-10133-8

Gao, L., & Zhang, F. (2026). Cognitive offloading and metacognitive calibration in generative AI-mediated doctoral learning: A grounded theory study of accountable reconstruction. *Frontiers in Psychology, 17*, Article 1920947. https://doi.org/10.3389/fpsyg.2026.1920947

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

Grinschgl, S., Papenmeier, F., & Meyerhoff, H. S. (2021). Consequences of cognitive offloading: Boosting performance but diminishing memory. *Quarterly Journal of Experimental Psychology, 74*(9), 1477–1496. https://doi.org/10.1177/17470218211008060

Guenther, P., Guenther, M., Ringle, C. M., Zaefarian, G., & Cartwright, S. (2023). Improving PLS-SEM use for business marketing research. *Industrial Marketing Management, 111*, 127–142. https://doi.org/10.1016/j.indmarman.2023.03.010

Guo, L. (2022). Using metacognitive prompts to enhance self-regulated learning and learning outcomes: A meta-analysis of experimental studies in computer-based learning environments. *Journal of Computer Assisted Learning, 38*(3), 811–832. https://doi.org/10.1111/jcal.12650

Haditia, M., Abubakar, I. Y., Rokimin, & Fauzi, A. (2026). Design and validation of a quality assurance framework for AI educational technology in low-resource settings. *Educational Technology in Developing Countries, 1*(1), 83–102. https://doi.org/10.64929/etdc.v1i1.68

Hair, J. F., Jr. (2021). Next-generation prediction metrics for composite-based PLS-SEM. *Industrial Management & Data Systems, 121*(1), 5–11. https://doi.org/10.1108/IMDS-08-2020-0505

Hair, J. F., Jr., & Alamer, A. (2022). Partial least squares structural equation modeling (PLS-SEM) in second language and education research: Guidelines using an applied example. *Research Methods in Applied Linguistics, 1*(3), Article 100027. https://doi.org/10.1016/j.rmal.2022.100027

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., & Sarstedt, M. (2022). *A primer on partial least squares structural equation modeling (PLS-SEM)* (3rd ed.). Sage.

Hair, J. F., Jr., Hult, G. T. M., Ringle, C. M., Sarstedt, M., Danks, N. P., & Ray, S. (2021). *Partial least squares structural equation modeling (PLS-SEM) using R: A workbook*. Springer. https://doi.org/10.1007/978-3-030-80519-7

Hair, J. F., Jr., Sarstedt, M., Ringle, C. M., & Gudergan, S. P. (2024). *Advanced issues in partial least squares structural equation modeling* (2nd ed.). Sage.

Hermawan, I., Nur, T., Mansyur, M. H., Bariah, O., Farida, N. A., Hermawan, R. A., Hermawan, A. S. W., & Billah, R. A. M. (2025). Integrating artificial intelligence in Islamic education: Insights from Madrasah Aliyah in Karawang, Indonesia. *Fikroh: Jurnal Pemikiran dan Pendidikan Islam, 18*(3), 505–515. https://doi.org/10.37812/fikroh.v18i3.2119

Hermita, M., Rohani, A., Syah, R. D., Suhatril, R. J., & Sartika, S. (2026). The Janus face of AI literacy: A dual-pathway model of its mitigating and exacerbating effects on academic procrastination. *Cogent Education, 13*(1), Article 2644679. https://doi.org/10.1080/2331186X.2026.2644679

Howard, J. L., Bureau, J., Guay, F., Chong, J. X. Y., & Ryan, R. M. (2021). Student motivation and associated outcomes: A meta-analysis from self-determination theory. *Perspectives on Psychological Science, 16*(6), 1300–1323. https://doi.org/10.1177/1745691620966789

Hox, J. J., Moerbeek, M., & van de Schoot, R. (2017). *Multilevel analysis: Techniques and applications* (3rd ed.). Routledge. https://doi.org/10.4324/9781315650982

Iqbal, J., Hashmi, Z. F., Asghar, M. Z., & Abid, M. N. (2025). Generative AI tool use enhances academic achievement in sustainable education through shared metacognition and cognitive offloading among preservice teachers. *Scientific Reports, 15*(1), Article 16610. https://doi.org/10.1038/s41598-025-01676-x

Järvelä, S., Nguyen, A., & Hadwin, A. (2023). Human and artificial intelligence collaboration for socially shared regulation in learning. *British Journal of Educational Technology, 54*(5), 1057–1076. https://doi.org/10.1111/bjet.13325

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

Kaya, D., & Yıldız Durak, H. (2026). A socially shared regulation of learning and self-determination theory-based classification of generative artificial intelligence-mediated mathematical modeling activities: The case of ChatGPT. *European Journal of Psychology of Education, 41*(1), Article 35. https://doi.org/10.1007/s10212-026-01085-2

Klar, M. (2025). Using ChatGPT is easy, using it effectively is tough? A mixed methods study on K-12 students' perceptions, interaction patterns, and support for learning with generative AI chatbots. *Smart Learning Environments, 12*, Article 32. https://doi.org/10.1186/s40561-025-00385-2

Kock, F., Berbekova, A., & Assaf, A. G. (2021). Understanding and managing the threat of common method bias: Detection, prevention and control. *Tourism Management, 86*, Article 104330. https://doi.org/10.1016/j.tourman.2021.104330

Kock, N., & Dow, K. E. (2025). Statistical significance and effect size tests in SEM: Common method bias and strong theorizing. In *Advances in management accounting* (Vol. 37, pp. 95–105). Emerald. https://doi.org/10.1108/S1474-787120250000037005

Listyasari, W. D., Kayen, H. S., & Haq, M. A. (2026). Charismatic kiai authority and professional learning communities in Indonesian pesantren. *Leadership in Muslim Societies, 1*(1), 55–74. https://doi.org/10.64929/lms.v1i1.59

Liu, B. (2026). How does self-regulated learning resist AI dependence? A mediating effect study based on college students who frequently use AIGC tools. *SAGE Open, 16*(1). https://doi.org/10.1177/21582440261420068

Lodge, J. M., Yang, S., Furze, L., & Dawson, P. (2023). It's not like a calculator, so what is the relationship between learners and generative artificial intelligence? *Learning: Research and Practice, 9*(2), 117–124. https://doi.org/10.1080/23735082.2023.2261106

McNeish, D. (2026). A practical guide to selecting and blending approaches for clustered data: Clustered errors, multilevel models, and fixed-effect models. *Psychological Methods, 31*(2), 225–251. https://doi.org/10.1037/met0000620

Molenaar, I. (2022). The concept of hybrid human-AI regulation: Exemplifying how to support young learners' self-regulated learning. *Computers and Education: Artificial Intelligence, 3*, Article 100070. https://doi.org/10.1016/j.caeai.2022.100070

Najib, A. C., & Darnoto, D. (2026). Intellectual degradation amidst the development of AI technology: A comparison between AI dependency students and AI assisted learning students. *Tafkir: Interdisciplinary Journal of Islamic Education, 7*(4), 1045–1060. https://doi.org/10.31538/tijie.v7i4.3207

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

Ng, D. T. K., Tan, C. W., & Leung, J. K. L. (2024). Empowering student self-regulated learning and science education through ChatGPT: A pioneering pilot study. *British Journal of Educational Technology, 55*(4), 1328–1353. https://doi.org/10.1111/bjet.13454

Perifanou, M., & Economides, A. A. (2025). Students collaboratively prompting ChatGPT. *Computers, 14*(5), Article 156. https://doi.org/10.3390/computers14050156

Pintrich, P. R., Smith, D. A. F., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (ED338122). National Center for Research to Improve Postsecondary Teaching and Learning, University of Michigan. https://eric.ed.gov/?id=ED338122

Podsakoff, P. M., Podsakoff, N. P., Williams, L. J., Huang, C., & Yang, J. (2024). Common method bias: It's bad, it's complex, it's widespread, and it's not easy to fix. *Annual Review of Organizational Psychology and Organizational Behavior, 11*(1), 17–61. https://doi.org/10.1146/annurev-orgpsych-110721-040030

Rahiem, M. D. H. (2026). Generative AI in higher education in Indonesia: Patterns of use and learning impact. *Social Sciences & Humanities Open, 13*, Article 102672. https://doi.org/10.1016/j.ssaho.2026.102672

Rinekso, A. B., Rodliyah, R. S., Purnawarman, P., & Lengkanawati, N. S. (2026). Indonesian EFL university students' ChatGPT acceptance: A cross-sectional study. *Journal of Education and Learning (EduLearn), 20*(3), 1959–1967. https://doi.org/10.11591/edulearn.v20i3.24855

Ringle, C. M., Wende, S., & Becker, J.-M. (2024). *SmartPLS 4* [Computer software]. SmartPLS GmbH. https://www.smartpls.com

Risko, E. F., & Gilbert, S. J. (2016). Cognitive offloading. *Trends in Cognitive Sciences, 20*(9), 676–688. https://doi.org/10.1016/j.tics.2016.07.002

Ryan, R. M., & Deci, E. L. (2017). *Self-determination theory: Basic psychological needs in motivation, development, and wellness*. Guilford Press. https://doi.org/10.1521/978.14625/28806

Salam, F. N. D., Rahman, E. S., Fakhri, M. M., & Soeharto, S. (2026). How ChatGPT usage and AI over-reliance influence student integrity: The moderating role of academic misconduct. *Tadris: Jurnal Keguruan dan Ilmu Tarbiyah, 11*(2), 627–638. https://ejournal.radenintan.ac.id/index.php/tadris/article/view/29449

Sardi, J., Darmansyah, Candra, O., Yuliana, D. F., Habibullah, Yanto, D. T. P., & Eliza, F. (2025). How generative AI influences students' self-regulated learning and critical thinking skills? A systematic review. *International Journal of Engineering Pedagogy, 15*(1), 94–108. https://doi.org/10.3991/ijep.v15i1.53379

Sarstedt, M., Hair, J. F., Pick, M., Liengaard, B. D., Radomir, L., & Ringle, C. M. (2022). Progress in partial least squares structural equation modeling use in marketing research in the last decade. *Psychology & Marketing, 39*(5), 1035–1064. https://doi.org/10.1002/mar.21640

Sarwanti, S., Sariasih, Y., Rahmatika, L., Islam, M. M., & Riantina, E. M. (2024). Are they literate on ChatGPT? University language students' perceptions, benefits and challenges in higher education learning. *Online Learning, 28*(3), 105–130. https://doi.org/10.24059/olj.v28i3.4599

Schunk, D. H., & Greene, J. A. (Eds.). (2018). *Handbook of self-regulation of learning and performance* (2nd ed.). Routledge. https://doi.org/10.4324/9781315697048

Stadler, M., Bannert, M., & Sailer, M. (2024). Cognitive ease at a cost: LLMs reduce mental effort but compromise depth in student scientific inquiry. *Computers in Human Behavior, 160*, Article 108386. https://doi.org/10.1016/j.chb.2024.108386

Sutarni, N., Ramdhany, M. A., Hufad, A., & Kurniawan, E. (2021). Self-regulated learning and digital learning environment: Its' effect on academic achievement during the pandemic. *Jurnal Cakrawala Pendidikan, 40*(2), 374–388. https://doi.org/10.21831/cp.v40i2.40718

Syamsudduha, S., Suwandi, S., & ZA, T. (2026). Islamic AI literacy as responsible knowledge practice: Development and validation of AILS-IIUS in Indonesian Islamic higher education. *Jurnal Pendidikan Islam, 12*(1), 187–201. https://doi.org/10.15575/jpi.v12i1.54254

Tao, S., Lan, M., Wang, M., & Li, H. (2026). Potential risks of generative artificial intelligence integration into K-12 education: A scoping review. *Computers and Education: Artificial Intelligence, 10*, Article 100561. https://doi.org/10.1016/j.caeai.2026.100561

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

Wang, B., Rau, P.-L. P., & Yuan, T. (2023). Measuring user competence in using artificial intelligence: Validity and reliability of artificial intelligence literacy scale. *Behaviour & Information Technology, 42*(9), 1324–1337. https://doi.org/10.1080/0144929X.2022.2072768

Wang, F., Jiang, C., King, R. B., & Leung, S. O. (2023). Motivated Strategies for Learning Questionnaire (MSLQ): Adaptation, validation, and development of a short form in the Chinese context for mathematics. *Psychology in the Schools, 60*(6), 2018–2040. https://doi.org/10.1002/pits.22845

Wang, G., Wang, W., Yang, D., & Ren, J. (2026). Generative AI, cognitive offloading, and learner agency in higher education: A scoping review. *Behavioral Sciences, 16*(7), Article 1150. https://doi.org/10.3390/bs16071150

Werdiningsih, I., Marzuki, Indrawati, I., Rusdin, D., Ivone, F. M., Basthomi, Y., & Zulfahreza. (2024). Revolutionizing EFL writing: Unveiling the strategic use of ChatGPT by Indonesian master's students. *Cogent Education, 11*(1), Article 2399431. https://doi.org/10.1080/2331186X.2024.2399431

Wood, D., Bruner, J. S., & Ross, G. (1976). The role of tutoring in problem solving. *Journal of Child Psychology and Psychiatry, 17*(2), 89–100. https://doi.org/10.1111/j.1469-7610.1976.tb00381.x

Xia, Q., Yang, Y., Wang, W., & Yin, H. (2026). Promoting interdisciplinary learning with generative AI through self-regulated scaffolding. *Journal of Computer Assisted Learning, 42*(2), Article e70214. https://doi.org/10.1002/jcal.70214

Xu, X., Qiao, L., Cheng, N., Liu, H., & Zhao, W. (2025). Enhancing self-regulated learning and learning experience in generative AI environments: The critical role of metacognitive support. *British Journal of Educational Technology, 56*(5), 1842–1863. https://doi.org/10.1111/bjet.13599

Zh, M. H. R., Sani, N. L., Kuswandi, D., & Fadhli, M. (2024). Needs analysis of development FBO media as a support for blended learning in Al-Qur'an Hadits lesson. *Jurnal Pendidikan Agama Islam Al-Thariqah, 9*(1), 17–32. https://doi.org/10.25299/al-thariqah.2024.vol9(1).15383

Zh, M. H. R., Sari, C. M., At Thaariq, Z. Z., Sumalee, I., & Kamilah, A. N. (2026). Enhancing conceptual understanding in Islamic education through a hypercontent digital flipbook in blended learning: Evidence from Indonesian madrasah. *Educational Technology in Developing Countries, 1*(1), 1–22. https://doi.org/10.64929/etdc.v1i1.60

Zhai, C., Wibowo, S., & Li, L. D. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments, 11*(1), Article 28. https://doi.org/10.1186/s40561-024-00316-7

Zhang, Q., & Liang, Y. (2026). Chatbot-assisted self-regulated learning scaffolding in EAP writing: The mediating role of SRL and the moderating role of AI literacy. *Education and Information Technologies, 31*(14), 6443–6482. https://doi.org/10.1007/s10639-026-14038-5

Zhang, S., Zhao, X., Zhou, T., & Kim, J. H. (2024). Do you have AI dependency? The roles of academic self-efficacy, academic stress, and performance expectations on problematic AI usage behavior. *International Journal of Educational Technology in Higher Education, 21*(1), Article 34. https://doi.org/10.1186/s41239-024-00467-0

Zhu, Q., Li, X., Dong, Y., Chang, P., & Fan, M. (2026). Not all cognitive offloading is equal: Distinguishing dependent and autonomous offloading to generative AI. *Frontiers in Psychology, 17*