Journal of AI, Humanities, and New Ethics,
1(2), 81-88, December 2025
e-ISSN: 3062-2867
jaihne.com

Genc Bilge Publishing Ltd.
(Young Wise) © 2025
gencbilgeyayincilik.com

Research Article

The role of cognitive laziness as a mediator between the intensity of AI use and critical thinking ability among university students

Muhammad Bahrul Nabil[1] and Nur Eva[2]*

Psychology Faculty, Malang State University, Malang, Indonesia

Article Info

Received: 2 October 2025
Accepted: 21 December 2025
Online: 30 December 2025

Keywords
Cognitive laziness
Critical thinking
Generative AI
PLS-SEM
University students

Abstract

Departing from the increasing use of generative artificial intelligence (GenAI) in higher education, which on the one hand helps academic efficiency, but on the other hand has the potential to reduce the quality of thinking processes when used excessively. This study aims to analyze the effect of AI usage intensity on students' critical thinking ability and to examine the role of cognitive laziness as a mediator. The method used was a correlational quantitative study with a survey approach involving 200 students of Universitas Negeri Malang from several study programs (53% female; 46.9% male). Data were collected using a questionnaire containing three scales: AI usage intensity (10 items), cognitive laziness (adapted from LAS; 9 items), and critical thinking (12 items). The analysis was conducted using PLS-SEM. The results showed that the intensity of AI use negatively affected critical thinking ($\beta = -0.275$), while the effect of AI usage intensity on cognitive laziness was very weak ($\beta = 0.020$) with $R^2 = 0.000$. Cognitive laziness affected critical thinking ($\beta = 0.438$), but the mediation effect was not strongly supported because the indirect effect was very small (≈0.0088). The model explained 26.3% of the variation in critical thinking ($R^2 = 0.263$). These findings indicate that AI use is associated with a decline in critical thinking mainly through direct effects, while the mediation mechanism needs to be re-examined using stronger instruments.

To cite this article

Nabil, M.B., and Eva, N. (2025 The role of cognitive laziness as a mediator between the intensity of AI use and critical thinking ability among university students. *Journal of AI, Humanities, and New Ethics, 1*(2), 73-80. DOI: https://doi.org/10.5281/zenodo.

Introduction

The use of generative artificial intelligence (GenAI), such as large language models (LLMs), in higher education has increased rapidly due to its ability to assist information retrieval, summarization, idea generation, and academic writing support. Recent literature confirms that LLMs have the potential to accelerate cognitive work and support learning when used appropriately (Kasneci et al., 2023; Dwivedi et al., 2023). At the same time, international institutions also emphasize the need for governance and user capacity to ensure that GenAI does not create ethical risks and reduce the quality of learning (Miao & Holmes, 2023). In the context of Indonesian higher education, the need for guidelines on GenAI use has also begun to be formalized through guideline documents for the academic community (Kusumawardani et al., 2024).

In line with this, the present study positions the intensity of AI use as an academic reality that is "unavoidable," while simultaneously highlighting the dual aspects of AI use: improving efficiency and the quality of the learning process, yet

[1] Student, Malang State University, Malang, Indonesia.
[2] Corresponding Author: Assoc.Prof, , Faculty of Pscychology, Malang State University, Malang, Indonesia.. E-mail: nur.eva.fppsi@um.ac.id ORCID: 0000-0003-3584-5049

also potentially giving rise to "cognitive laziness" and a decline in critical thinking skills when students become overly dependent on AI-generated outputs. Pedagogically, recent findings indicate that the benefits of GenAI tend to emerge when its use is directed toward supporting self-regulation and reflection, for example through metacognitive support that encourages planning, monitoring, and learning evaluation (Xu et al., 2025). This indicates that the main issue is not merely whether to use AI or not, but rather how GenAI is used and how students' thinking habits change when AI becomes a source of instant answers.

However, concerns regarding the decline in cognitive abilities are not merely assumptions. Synthesized evidence shows that over-reliance on AI dialogue systems may be associated with the weakening of important cognitive abilities, including critical thinking and analytical reasoning, because users tend to accept AI recommendations without verification (Zhai et al., 2024). From a cognitive psychology perspective, this phenomenon is consistent with the concept of cognitive offloading, namely the tendency to transfer mental burdens to external tools so that internal thinking processes are reduced (Risko & Gilbert, 2016). When offloading becomes habitual in academic tasks, such as constructing arguments, evaluating evidence, or checking logical consistency, critical thinking skills are at risk of not being adequately trained.

At this point, cognitive laziness becomes a mechanism worthy of scientific examination. The literature on GenAI-based education has begun to emphasize that the ease of obtaining answers can trigger "metacognitive laziness" and weaken the motivation to conduct in-depth evaluation, which in turn affects critical thinking (Fan et al., 2024). Other studies also demonstrate relevant psychological pathways: dependence on AI may reduce critical thinking through mechanisms of fatigue and information overload, while information literacy and AI literacy function as socio-cognitive buffers (Tian & Zhang, 2025). Nevertheless, the current state of the art reveals several gaps: (1) many studies examine the impact of AI on performance, learning outcomes, or ethical risks in general; (2) specific psychological mechanisms that mediate the relationship between the intensity of AI use and critical thinking, particularly cognitive laziness, have not consistently become the primary focus in student populations; and (3) contextual evidence in higher education, with varying learning cultures and institutional policies, remains limited.

In addition to mediation mechanisms, protective factors are also important for explaining when AI use becomes problematic. The AI literacy framework emphasizes competencies to understand AI limitations, critically evaluate outputs, and collaborate effectively with AI (Long & Magerko, 2020). Empirically, AI literacy and information literacy have been shown to weaken the negative impact of AI dependence on critical thinking (Tian & Zhang, 2025). On the other hand, the issue of self-efficacy is also relevant: dependence on GenAI may encourage false self-efficacy, namely efficacy that is "perceived as high" but inconsistent with actual competence, and is correlated with decreased academic achievement (Jia et al., 2025). These findings strengthen the argument that the main risk is not merely technical dependence, but also the distortion of metacognitive processes, including self-assessment, information verification, and willingness to allocate cognitive effort.

Based on this mapping, several alternative solutions are commonly adopted by institutions: restricting or prohibiting GenAI to prevent academic dishonesty, integrating GenAI in a guided manner through task designs that require verification and reflection, and implementing AI literacy interventions so that students are able to calibrate their trust in AI. Global guidelines also encourage strengthening human capacity and regulating the use of GenAI to align with educational goals (Miao & Holmes, 2023). Based on this logic, the most rational solution to be tested and developed is not merely "prohibition," but rather ensuring that the use of GenAI continues to stimulate higher-order thinking processes, for example through metacognitive support and improved AI literacy (Xu et al., 2025; Long & Magerko, 2020). However, for such solutions to be appropriately targeted, evidence is needed regarding the psychological mechanisms that explain why the intensity of AI use may reduce critical thinking among some students.

Thus, the scientific novelty of this study lies in examining cognitive laziness as a mediator explaining the relationship between the intensity of AI use and critical thinking ability, as designed in the conceptual framework in the document (X: intensity of AI use → M: cognitive laziness → Y: critical thinking). This study also opens opportunities for strengthening the model through protective factors that have been considered in the document, such as AI literacy and self-efficacy, to explain variations in impact across individuals. Methodologically, the correlational quantitative analysis design based on surveys and path modeling, as proposed in the document, provides an adequate framework for examining direct effects, indirect effects (mediation), and potential moderation effects.

The purpose of this study is to analyze the effect of the intensity of AI use on students' critical thinking, examine the mediating role of cognitive laziness, and provide an empirical basis for designing GenAI usage strategies that encourage critical thinking through AI literacy and metacognitive support, rather than merely relying on restrictions.

Method

Research Model

The research model in this study examines the direct effect of AI usage intensity on students' critical thinking and the indirect effect through cognitive laziness as a mediating variable. The model consists of AI Usage Intensity (X) → Cognitive Laziness (M) → Critical Thinking (Y). This model was developed based on a quantitative correlational approach using path analysis to examine direct and indirect relationships among variables (Creswell & Creswell, 2018).

Participant

The respondents consisted of 200 UM students from several departments, namely Psychology, Guidance and Counseling, and Educational Management (53% female, 46.9% male), who participated in this study.

Data Collection Tool

The data collection instrument in this study was a structured questionnaire consisting of three main measurement tools to assess the intensity of AI use, cognitive laziness, and students' critical thinking ability. The intensity of AI use was measured using the AI Usage Intensity Scale, which consisted of 10 statement items regarding habits and the level of utilization of generative AI (such as ChatGPT, Gemini, Copilot, and similar tools) in academic activities, such as assisting with assignments, generating ideas/writing outlines, and explaining difficult concepts. Responses were provided using a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree). One item was reverse-coded, so score reversal was conducted before total score calculation. The total score was obtained by summing all items after reversal, where a higher score indicated a higher intensity of AI use.

Cognitive Laziness Scale for Students

The mediator variable, cognitive laziness, was measured using the Cognitive Laziness Scale for Students, which was a contextual adaptation of the Laziness Assessment Scale (LAS). This scale consisted of 9 items and was designed to capture the tendency to avoid cognitive effort, fail to complete readings/tasks, and choose to complete work minimally in the context of higher education, including when AI assistance was available. Responses used a 5-point frequency scale (1 = never to 5 = almost always). All items were scored in the same direction, and the total score was obtained by summing all items; higher scores indicated higher levels of cognitive laziness.

Critical Thinking Ability Scale

Critical thinking ability was measured using the Critical Thinking Ability Scale, which consisted of 12 statement items assessing aspects of argument evaluation, information verification, multi-perspective consideration, logical conclusion drawing, and the habit of checking the reasonableness of answers when using AI. Responses were provided using a 5-point Likert scale (1 = strongly disagree to 5 = strongly agree). Two items were reverse-coded, so score reversal was conducted before calculating the total score. The total score was calculated by summing all items after reversal; higher scores indicated better critical thinking ability.

Data Analysis

Questionnaire data were first coded and scored according to the guidelines of each scale. Reverse-coded items on the AI Usage Intensity Scale and the Critical Thinking Ability Scale were reversed before calculating the total score. The score for each construct was calculated by summing the item scores (or by using the average indicator score if analyzed as a latent construct). Subsequently, data feasibility examinations were conducted, including missing data, inappropriate responses (for example, identical response patterns across all items), and descriptive statistics (mean, standard deviation, minimum-maximum) to describe respondent profiles and trends in the research variables.

Hypothesis testing was conducted using PLS-SEM because the model contained direct and indirect (mediation) relationships among latent constructs. The evaluation of the measurement model was conducted by assessing (1) indicator reliability through outer loading values (referring to the general criterion of ≥ 0.70; indicators between 0.40–0.70 could be considered for retention if construct reliability/validity remained adequate), (2) internal consistency reliability through Cronbach's Alpha, rho_A, and Composite Reliability (general reference ≥ 0.70; in exploratory studies values approaching 0.60 could be considered), and (3) convergent validity through AVE (≥ 0.50). Discriminant validity evaluation could be complemented using HTMT criteria (for example < 0.85 or < 0.90) and/or Fornell-Larcker as well as cross-loading examination to ensure that indicators loaded more strongly on their own constructs than on other constructs.

Structural model evaluation was conducted by examining potential multicollinearity (for example VIF), assessing the magnitude and direction of path coefficients, and evaluating the explanatory power of the model through R Square (R^2) and Adjusted R^2. The significance of path coefficients and indirect effects was tested through a bootstrapping procedure (for example 5,000 resampling) to obtain t-statistics, p-values, and confidence intervals. The mediation test of Cognitive Laziness was conducted by assessing the significance of the indirect effect (X→M→Y) and comparing it with the direct effect X→Y to determine whether the mediation was absent/weak/partial/full.

Results and Discussion

The PLS-SEM model in the figure shows three latent constructs: AI Usage Intention, Cognitive Laziness, and Critical Thinking, along with path coefficient values (inner model), R-square values, and indicator outer loadings (measurement model).

In the structural model, three main path coefficients were obtained: (1) AI Usage Intention → Cognitive Laziness of 0.020, indicating a very weak/nearly zero effect; (2) Cognitive Laziness → Critical Thinking of 0.438, indicating a fairly strong effect in magnitude; and (3) AI Usage Intention → Critical Thinking of −0.275, meaning that the higher the intention to use AI, the lower the critical thinking score tended to be. The explanatory power of the model can be seen from the R^2 values: the Cognitive Laziness construct had an $R^2 = 0.000$ (almost not explained by AI Usage Intention), while Critical Thinking had an $R^2 = 0.263$, meaning that approximately 26.3% of the variation in critical thinking was explained by AI Usage Intention and Cognitive Laziness.

Conceptually, the mediation effect can be estimated from the product of the paths $a \times b = 0.020 \times 0.438 = 0.0088$. This indirect effect value is very small, so the mediating contribution of Cognitive Laziness to the relationship between AI Usage Intention and Critical Thinking appears practically minimal (confirmation of mediation significance still requires bootstrapping output).

In the measurement model, many indicators showed low outer loadings. For the AI Usage Intention construct, only one indicator appeared adequate (for example loading around 0.720), while many other indicators were low (around 0.20–0.42) and some were even negative (for example around −0.193; −0.267; −0.056). In the Cognitive Laziness construct, relatively stronger indicators were in the range of 0.472–0.633, but several other indicators were very low (for example 0.028; 0.162; 0.176; 0.317; 0.359). In the Critical Thinking construct, only a few indicators were relatively stronger (for example around 0.602; 0.453; 0.426), whereas most indicators were in the low range (around 0.171–0.366) and one indicator was negative (around −0.15).

The most substantively consistent finding from the structural model is the negative relationship between AI Usage Intention and Critical Thinking ($\beta = -0.275$). Interpretatively, this result points to the possibility that increased intention/use of AI for academic activities is associated with reduced student engagement in evaluative thinking processes, such as the habit of examining evidence, comparing arguments, and independently constructing reasoning, thereby resulting in lower critical thinking. In the context of generative AI use, this is consistent with concerns that AI may facilitate cognitive "shortcuts" (quick answers, instant summaries) that reduce higher-order thinking practice when not accompanied by verification.

However, the AI Usage Intention → Cognitive Laziness path, which is nearly zero (β = 0.020; R^2 Cognitive Laziness = 0.000), indicates that in these data, AI use does not sufficiently explain variations in cognitive laziness. There are two most plausible explanations. First, students' AI use may be heterogeneous: some use AI as an initial aid but continue to think and recheck independently, so AI intention does not automatically lead to cognitive laziness. Second, and most importantly, this result is very likely influenced by weak measurement construct quality, as indicated by the large number of low outer loadings and the presence of negative loadings, causing the relationships among latent variables to be reduced or unstable.

The most "unusual" finding is the positive Cognitive Laziness → Critical Thinking path (β = 0.438). Theoretically, cognitive laziness is usually predicted to reduce critical thinking. Therefore, this positive direction needs to be interpreted cautiously and most likely indicates issues in operationalization/scoring or construct validity. Two strong indications are: (1) the presence of indicators with negative loadings on the construct, which often occurs when reverse items have not been reverse-scored or when indicators are not aligned with the construct definition; and (2) the large number of indicators with very low loadings, making the latent construct less "clean" in representing the intended concept. If the "Cognitive Laziness" construct empirically captures something different (for example, certain reflective attitudes, or artifacts from non-aligned items), then the resulting path coefficients may contradict theoretical expectations.

When reviewed as a mediation test, the main requirement for strong mediation is that both the X→M and M→Y paths are adequate. In this model, the X→M path is very weak, so even though M→Y is large, the indirect effect a×b remains very small (≈0.0088). This indicates that in the current model, the explanation for the decline in critical thinking is more dominantly through the direct effect of AI use on critical thinking, rather than through the cognitive laziness mechanism.

Table 1. R-Square and adjusted R-Square values for endogenous constructs

Endogenous Variable	R Square (R^2)	Adjusted R Square (Adj. R^2)
Cognitive Laziness	0.000	-0.005
Critical Thinking	0.263	0.255

Table 1 (R Square) presents the R Square (R^2) and Adjusted R Square (Adj. R^2) values for each endogenous variable, namely variables explained by predictor constructs in the model.

For the Cognitive Laziness construct, the values obtained were R^2 = 0.000 and Adj. R^2 = −0.005. These findings indicate that the variation in cognitive laziness can hardly be explained by the predictor directed toward it (in this model: intention/intensity of AI use). The negative Adj. R^2 value indicates that after considering the number of predictors and sample size, the model does not provide improved explanatory power for cognitive laziness; statistically, its performance is even lower than a model without predictors. Practically, this indicates that the structural effect of the predictor on cognitive laziness is very weak.

In contrast, for the Critical Thinking construct, the values obtained were R^2 = 0.263 and Adj. R^2 = 0.255, meaning that approximately 26.3% of the variation in critical thinking can be explained by the predictors in the model (namely AI usage intention/intensity and cognitive laziness). The small difference between R^2 and Adj. R^2 indicates that correction due to model complexity does not meaningfully alter the results, so the explanatory power of the model on critical thinking is relatively stable. In social and behavioral research, an R^2 value around 0.25 is generally categorized as weak to moderate, yet still reasonable considering the many other factors outside the model that also influence critical thinking ability.

Overall, Table 1 shows that this structural model is more capable of explaining variation in Critical Thinking than in Cognitive Laziness. The implication is that the mechanism positioning cognitive laziness as a variable influenced by AI use has not yet obtained strong empirical support in these data, whereas variation in critical thinking can still be partially explained by the constructs in the model.

Table 2. Reliability and Convergent Validity

Construct	Cronbach's Alpha	rho_A	Composite Reliability	AVE
Cognitive Laziness	0.607	0.697	0.580	0.666
Critical Thinking	0.799	0.570	0.521	0.621
AI Usage Intention	0.525	0.571	0.703	0.610

Table 2 presents the evaluation of reliability and convergent validity for each construct, as indicated by Cronbach's Alpha, rho_A, Composite Reliability (CR), and Average Variance Extracted (AVE). Cronbach's Alpha and rho_A describe the internal consistency of items within a construct, whereas CR assesses construct reliability in the PLS-SEM measurement model. Generally, good reliability values are in the range of ≥ 0.70, although in exploratory studies values of 0.60 may still be considered acceptable. Meanwhile, an AVE ≥ 0.50 indicates that the construct is capable of explaining more than 50% of the variance of its indicators, thus meeting the criteria for convergent validity.

The results in the table show that the Cognitive Laziness construct has a Cronbach's Alpha of 0.607 and rho_A of 0.697, indicating that internal consistency falls within a moderate category. However, the CR value of 0.580 is still relatively low, suggesting that the reliability of this construct is not yet optimal, even though its AVE is high (0.666). This indicates that on average the construct is able to capture indicator variance well, but there are still indicators with weak contributions that reduce overall consistency.

For the Critical Thinking construct, Cronbach's Alpha reached 0.799, which generally indicates good reliability. However, the rho_A value (0.570) and CR (0.521) are relatively low, indicating inconsistency among the reliability indices. This condition usually suggests the presence of unstable indicators (for example low/negative loadings or reverse items that have not been reverse-scored), so although the alpha appears high, the construct reliability in the context of PLS-SEM is not yet strong. Nevertheless, the AVE value of 0.621 still meets the criteria for convergent validity.

Meanwhile, the AI Usage Intention construct shows a Cronbach's Alpha of 0.525 and rho_A of 0.571, both of which are relatively low, indicating weak internal consistency among the items. However, the CR value of 0.703 has reached the commonly accepted minimum threshold, and the AVE value of 0.610 also satisfies the convergent validity criterion. This pattern indicates that convergently the construct is able to explain indicator variance fairly well, but the homogeneity/accuracy among items still needs to be strengthened.

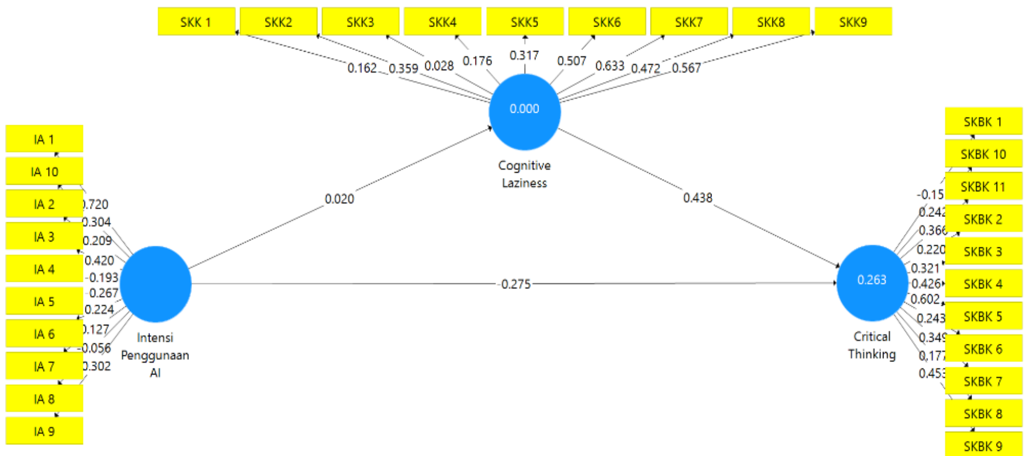

Figure 1. Structural equation model of the relationships among AI usage intensity, cognitive laziness, and critical thinking

Overall, Table 2 indicates that all three constructs have met convergent validity criteria (because all AVE values are > 0.50), but aspects of reliability, particularly those reflected in CR and several rho_A/Cronbach's Alpha values, still require attention. Therefore, future studies are recommended to conduct indicator evaluation (for example by examining reverse items, removing items with very low/negative loadings, and re-testing reliability) so that construct measurement becomes more robust.

Conclusion

This study shows that the intention/intensity of AI use has a negative effect on students' critical thinking ability (β = −0.275). This means that the higher students' use of AI in academic activities, the more it tends to be followed by a decrease in critical thinking scores. Meanwhile, the intention/intensity of AI use has almost no effect on cognitive laziness (β = 0.020) with R^2 = 0.000, so variation in cognitive laziness cannot be explained by AI use in this model. Overall, the model is able to explain 26.3% of the variation in critical thinking (R^2 = 0.263; Adj. R^2 = 0.255), indicating weak-to-moderate explanatory power.

The role of cognitive laziness as a mediator is not strongly supported, because the path from AI use to cognitive laziness is very weak and the indirect effect (a×b) is very small. Thus, in these data, the effect of AI use on critical thinking occurs more dominantly through a direct effect, rather than through the mediating mechanism of cognitive laziness. However, this conclusion should be interpreted cautiously because the evaluation of the measurement model shows that several indicators remain weak (the reliability of some constructs is not yet optimal). Therefore, future studies are recommended to improve the instrument, particularly items with low/negative loadings and the checking of reverse items, and to retest the model so that the results are more robust.

References

Creswell, J. W., & Creswell, J. D. (2018). *Research design: Qualitative, quantitative, and mixed methods approaches* (5th ed.). SAGE Publications.

Dwivedi, Y. K., Kshetri, N., Hughes, L., Slade, E. L., Jeyaraj, A., Kar, A. K., Baabdullah, A. M., Koohang, A., Raghavan, V., Ahuja, M. K., Albanna, H., Al-Busaidi, K. A., Balakrishnan, J., Barlette, Y., Bertin, E., ... Wright, R. (2023). So what if ChatGPT wrote it? Multidisciplinary perspectives on opportunities, challenges and implications of generative conversational AI for research, practice and policy. *International Journal of Information Management, 71*, 102642. https://doi.org/10.1016/j.ijinfomgt.2023.102642

Fan, Y., Tang, Y., Le, S., Shen, H., Tan, Y., Zhao, L., Shen, X., Li, H., & Gašević, D. (2024). Beware of metacognitive laziness: Effects of generative artificial intelligence use on learning motivation, self-regulated learning, and critical thinking. *British Journal of Educational Technology, 55*(3), 1049–1071. https://doi.org/10.1111/bjet.13544

Jia, W., Pan, L., & Neary, S. (2025). Effect of GenAI dependency on university students' academic achievement: The mediating role of self-efficacy and moderating role of perceived teacher caring. *Behavioral Sciences, 15*(10), 1348. https://doi.org/10.3390/bs15101348

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*, 102274. https://doi.org/10.1016/j.lindif.2023.102274

Kusumawardani, S. S., Wulandari, D., Pannen, P., Ekadiyanto, F. A., Wiryana, I. M., Purwarianti, A., & Alfarozi, S. A. I. (2024). *Panduan penggunaan generative artificial intelligence pada pembelajaran di perguruan tinggi* (*Guidelines for the use of generative artificial intelligence in higher education learning*) [Manual]. Direktorat Pembelajaran dan Kemahasiswaan, Jakarta. http://repositori.kemendikdasmen.go.id/id/eprint/32289

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

Miao, F., & Holmes, W. (2023). *Guidance for generative AI in education and research*. UNESCO.

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

Tian, Y., & Zhang, J. (2025). The dark side of generative AI: Unraveling the influence of AI dependence on college students' critical thinking. *Acta Psychologica, 252*, 105725. https://doi.org/10.1016/j.actpsy.2025.105725

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*(4), 1842–1863. https://doi.org/10.1111/bjet.13599

Zhai, X., Wibowo, S., & Li, Y. (2024). The effects of over-reliance on AI dialogue systems on students' cognitive abilities: A systematic review. *Smart Learning Environments*. https://doi.org/10.1186/s40561-024-00300-7