

PENGARUH INTENSITAS PENGGUNAAN GENERATIVE ARTIFICIAL INTELLIGENCE TERHADAP BERPIKIR KRITIS DAN HASIL BELAJAR MAHASISWA

THE EFFECT OF GENERATIVE ARTIFICIAL INTELLIGENCE USAGE INTENSITY ON STUDENTS' CRITICAL THINKING AND LEARNING OUTCOMES

PRASETYA YOGA SANTOSO, RESSA ULI PATRISSIA

UNIVERSITAS SAHID, UNIVERSITAS MUHAMMADIYAH PALANGKA RAYA

Received : Mar 31, 2026

Revised : Jun 5, 2026

Accepted : Jun 20, 2026

Abstract. This study analyses the effect of Generative Artificial Intelligence (Gen AI) usage intensity on the critical thinking and learning outcomes of higher-education students, positioning AI literacy as a moderating variable. A one-shot case study pre-experimental design was applied to 50 information-systems students at a private university in Jakarta over four weeks during the 2024/2025 odd semester. Participants were grouped by an AI Usage Intensity Index (AUII) into low ($n=16$), moderate ($n=17$), and high ($n=17$) categories using percentile-based classification. Data were collected through a validated questionnaire and tests of critical thinking and learning outcomes, then analysed using one-way ANOVA, the Tukey HSD post-hoc test, and two-way ANOVA for moderation. The results show significant differences in critical thinking ($F(2,47)=18.43$, $p<.001$, $\eta^2=.439$) and learning outcomes ($F(2,47)=22.17$, $p<.001$, $\eta^2=.485$) across groups; the high-intensity group attained the highest means for critical thinking ($M=84.71$) and learning outcomes ($M=87.12$). AI literacy significantly moderated the relationship between usage intensity and critical thinking. It is concluded that structured and literate Gen AI use is associated with higher-order cognition, so its curricular integration should be accompanied by strengthened AI literacy.

Abstrak. Penelitian ini menganalisis pengaruh intensitas penggunaan Generative Artificial Intelligence (Gen AI) terhadap kemampuan berpikir kritis dan hasil belajar mahasiswa, dengan literasi AI sebagai variabel moderator. Desain pra-eksperimen one-shot case study diterapkan pada 50 mahasiswa program studi sistem informasi di sebuah universitas swasta di Jakarta selama empat minggu pada semester ganjil 2024/2025. Partisipan dikelompokkan berdasarkan AI Usage Intensity Index (AUII) menjadi kategori rendah ($n=16$), sedang ($n=17$), dan tinggi ($n=17$) menggunakan klasifikasi berbasis persentil. Data dikumpulkan melalui kuesioner tervalidasi serta tes berpikir kritis dan hasil belajar, kemudian dianalisis dengan ANOVA satu arah, uji lanjut Tukey HSD, dan ANOVA dua arah untuk moderasi. Hasil menunjukkan perbedaan signifikan pada berpikir kritis ($F(2,47)=18,43$; $p<,001$; $\eta^2=,439$) dan hasil belajar ($F(2,47)=22,17$; $p<,001$; $\eta^2=,485$) antartiga kelompok; kelompok intensitas tinggi memperoleh rerata tertinggi untuk berpikir kritis ($M=84,71$) dan hasil belajar ($M=87,12$). Literasi AI terbukti memoderasi secara signifikan hubungan antara intensitas penggunaan dan berpikir kritis. Disimpulkan bahwa penggunaan Gen AI yang terstruktur dan berliterasi berkaitan dengan kognisi tingkat tinggi, sehingga integrasinya ke dalam kurikulum perlu disertai penguatan literasi AI.

Keywords:

Generative AI, critical thinking, learning outcomes, AI literacy, higher education

Kata kunci:

Generative AI, berpikir kritis, hasil belajar, literasi AI, pendidikan tinggi

(*) Corresponding Author: rpatriissia@gmail.com

How to Cite: Santoso, P., Y., Patrisia, R., U., (2026). Pengaruh Intensitas Penggunaan Generative Artificial Intelligence Terhadap Berpikir Kritis Dan Hasil Belajar Mahasiswa, 23 (1), 34-41. <https://doi.org/10.54124/jlmp.v23i1.192>

PENDAHULUAN

Kemunculan *Generative Artificial Intelligence* (Gen AI), khususnya model bahasa besar seperti ChatGPT, Google Gemini, dan Microsoft Copilot, telah mengubah cara mahasiswa belajar dan menyelesaikan tugas akademik. Secara ideal, kehadiran Gen AI seharusnya memperkuat kemampuan kognitif tingkat tinggi mahasiswa apabila digunakan secara terstruktur dan disertai literasi yang memadai. Namun kondisi faktual di lapangan memperlihatkan kesenjangan yang nyata antara potensi

tersebut dan praktik penggunaannya, sehingga diperlukan kajian empiris yang mengukur dampaknya secara terkendali.

Badan Pusat Statistik (2023) mencatat Indonesia memiliki 9,9 juta mahasiswa aktif, populasi terbesar keempat di dunia. Kemdikbudristek (2023) melaporkan bahwa 78,4% mahasiswa telah menggunakan minimal satu platform Gen AI, dan 62,1% di antaranya memanfaatkannya untuk mengerjakan tugas tanpa pemahaman mendalam terhadap *output* yang dihasilkan. Asosiasi Penyelenggara Jasa Internet Indonesia (2023) menunjukkan penetrasi internet mencapai 78,19% dengan kelompok usia 19–34 tahun sebagai pengguna terbesar (34,4%), sebuah katalis utama adopsi Gen AI di kalangan mahasiswa. Akan tetapi, *Indonesian Digital Literacy Survey* (2023) mengungkapkan bahwa hanya 31,2% mahasiswa memiliki literasi AI yang memadai, sedangkan 68,8% sisanya menggunakan Gen AI secara reaktif tanpa memahami keterbatasan, bias, dan etika penggunaannya. Lembaga Layanan Pendidikan Tinggi (2023) menambahkan bahwa 45,7% dosen belum memiliki panduan yang jelas mengenai penggunaan Gen AI di kelas.

Kesenjangan antara kondisi ideal dan kondisi faktual inilah yang menjadi dasar penelitian ini. Kondisi ideal mengharapkan Gen AI berfungsi sebagai penguat kognisi yang mendorong mahasiswa menjadi pembelajar mandiri, reflektif, dan kritis, sedangkan kondisi faktual justru memperlihatkan adopsi masif yang tidak diimbangi literasi, panduan kelembagaan, maupun kesiapan dosen. Ketimpangan ini berpotensi melahirkan ketergantungan dan *cognitive offloading* yang berlebihan. Gerlich (2025) bahkan menemukan korelasi negatif antara penggunaan AI yang tinggi dan kemampuan berpikir kritis ketika tidak disertai keterlibatan kognitif yang aktif. Temuan tersebut menegaskan bahwa hubungan antara penggunaan Gen AI dan capaian kognitif tidak bersifat linear, melainkan bergantung pada kualitas keterlibatan dan kapasitas kritis pengguna, sehingga menuntut pengujian empiris yang lebih cermat.

Sebagai solusi awal, penggunaan Gen AI perlu diarahkan melalui pendekatan pembelajaran terbimbing yang menempatkan literasi AI sebagai prasyarat, bukan sekadar pelengkap. *Generative AI* didefinisikan sebagai sistem kecerdasan buatan yang mampu menghasilkan konten orisinal berdasarkan pola yang dipelajari dari data pelatihan berskala besar (Feuerriegel et al., 2024). Dalam konteks pendidikan, Gen AI berpotensi berfungsi sebagai tutor adaptif dan alat *scaffolding* pembelajaran yang dapat disesuaikan dengan kebutuhan individual mahasiswa (Kasneci et al., 2023).

Penelitian ini berlandaskan dua teori utama. Pertama, *Cognitive Load Theory* (Sweller et al., 2019) yang membedakan beban kognitif intrinsik, *extraneous*, dan *germane*; Gen AI berpotensi mereduksi *extraneous load* sehingga sumber daya kognitif dapat dialihkan ke *germane load* yang berkontribusi pada konstruksi skema pengetahuan. Kedua, teori *Zone of Proximal Development* (Vygotsky, 1978) yang memandang Gen AI sebagai *More Knowledgeable Other* virtual yang memperluas zona perkembangan proksimal mahasiswa (Luckin, 2017). Kedua teori ini saling melengkapi: CLT menjelaskan mekanisme efisiensi pemrosesan kognitif, sedangkan ZPD menjelaskan dimensi sosial-mediasional dari interaksi mahasiswa dengan Gen AI sebagai mediator pembelajaran. Sintesis kedua kerangka inilah yang memandu perumusan hipotesis penelitian.

Konsep literasi AI menjadi penting karena memengaruhi cara mahasiswa memanfaatkan teknologi. Long & Magerko (2020) merumuskan literasi AI sebagai seperangkat kompetensi untuk memahami dan mengevaluasi AI secara kritis, sedangkan Ng et al. (2021) mengonseptualisasikannya sebagai kemampuan memahami, menggunakan, mengevaluasi, dan menyikapi AI secara etis. Dalam penelitian ini, literasi AI tidak diposisikan sebagai sekadar keterampilan teknis, melainkan sebagai kapasitas metakognitif yang memoderasi seberapa efektif intensitas penggunaan Gen AI ditransformasikan menjadi capaian kognitif yang bermakna.

Sejumlah penelitian internasional memperlihatkan kapasitas kognitif Gen AI yang tinggi sekaligus implikasinya bagi pembelajaran. Kasneci et al. (2023) memaparkan peluang dan tantangan model bahasa besar untuk pendidikan, sementara Kung et al. (2023) dan Gilson et al. (2023) menunjukkan bahwa ChatGPT mampu melampaui ambang kelulusan ujian lisensi medis, membuktikan kapasitas penalaran model yang relevan untuk konteks akademik. Temuan-temuan ini menegaskan bahwa Gen AI bukan sekadar mesin pencari, melainkan mitra kognitif yang berpotensi memengaruhi proses berpikir mahasiswa secara mendalam, baik ke arah positif maupun negatif.

Di Indonesia, kajian Gen AI dalam pendidikan masih berkembang. Nugroho et al. (2024)

mengidentifikasi pola dominan penggunaan Gen AI di kalangan mahasiswa, yaitu pencarian informasi, penyelesaian tugas, dan belajar mandiri. Namun studi-studi tersebut umumnya bersifat deskriptif-persepsional dan belum mengukur dampak terhadap kemampuan kognitif tingkat tinggi secara terkendali, apalagi mengintegrasikan literasi AI sebagai variabel yang memoderasi hubungan tersebut. Kesenjangan literatur ini bersifat substantif: belum ada penelitian di Indonesia yang secara simultan mengukur dampak intensitas penggunaan Gen AI terhadap berpikir kritis dan hasil belajar dengan menempatkan literasi AI sebagai moderator dalam satu model analitis yang utuh.

Kebaruan penelitian ini terletak pada dua aspek. Pertama, pengembangan *AI Usage Intensity Index* (AUII) yang mengukur intensitas penggunaan melalui tiga dimensi, yaitu frekuensi, kedalaman interaksi, dan diversitas, sehingga intensitas tidak direduksi menjadi sekadar frekuensi. Kedua, penempatan literasi AI sebagai variabel moderator dalam konteks sosiokultural pendidikan tinggi Indonesia yang dicirikan oleh orientasi belajar kolektif, tekanan akademik tinggi, dan keragaman akses infrastruktur digital. Sebagai jalan tengah antara melarang dan membiarkan penggunaan Gen AI secara bebas, penelitian ini mengusulkan pendekatan pembelajaran terbimbing berbasis literasi.

Berdasarkan latar belakang tersebut, penelitian ini bertujuan untuk menganalisis perbedaan kemampuan berpikir kritis mahasiswa berdasarkan intensitas penggunaan Gen AI; menganalisis perbedaan hasil belajar mahasiswa berdasarkan intensitas penggunaan Gen AI; dan menguji peran literasi AI sebagai moderator dalam hubungan antara intensitas penggunaan Gen AI dan kemampuan berpikir kritis. Sesuai tujuan tersebut, dirumuskan tiga hipotesis, yaitu H₁, terdapat perbedaan signifikan kemampuan berpikir kritis antarkelompok intensitas penggunaan Gen AI; H₂, terdapat perbedaan signifikan hasil belajar antarkelompok intensitas penggunaan Gen AI; dan H₃, literasi AI memoderasi secara signifikan hubungan antara intensitas penggunaan Gen AI dan kemampuan berpikir kritis.

METODE PENELITIAN

Penelitian ini menggunakan desain pra-eksperimen *one-shot case study* (Campbell & Stanley, 1963), yaitu satu kelompok yang diberi perlakuan kemudian diukur hasilnya tanpa kelompok kontrol dan tanpa *pretest*. Desain ini dipilih atas pertimbangan etika, karena membatasi akses teknologi bagi sebagian mahasiswa tidak memungkinkan, serta tujuan eksploratif untuk memperoleh gambaran penggunaan Gen AI dalam kondisi alamiah perkuliahan. Dengan demikian, penelitian ini bersifat *comparative explanatory* yang membandingkan capaian kognitif antarkelompok intensitas, dan secara eksplisit tidak mengklaim hubungan kausal yang definitif.

Penelitian dilaksanakan selama empat minggu pada semester ganjil tahun akademik 2024/2025 di sebuah universitas swasta di Jakarta, dengan subjek mahasiswa program studi sistem informasi. Populasi berjumlah 180 mahasiswa. Pengambilan sampel menggunakan teknik *purposive sampling* dengan kriteria telah menggunakan minimal satu platform Gen AI selama minimal tiga bulan, memiliki akses internet stabil, dan bersedia mengikuti seluruh prosedur penelitian. Berdasarkan kriteria tersebut diperoleh 50 mahasiswa (27,8% populasi) yang memenuhi syarat minimal sampel eksperimen (Cohen, 1988).

Meskipun desain *one-shot case study* memiliki keterbatasan dalam pengendalian validitas internal, desain ini relevan untuk tujuan eksploratif penelitian, yaitu memotret pengaruh intensitas penggunaan Gen AI dalam kondisi perkuliahan yang alamiah. Untuk memperkuat keabsahan, klasifikasi kelompok dilakukan secara empiris melalui distribusi persentil skor AUII, bukan berdasarkan ambang yang ditetapkan sebelumnya, sehingga setiap kelompok mencerminkan pola penggunaan yang benar-benar terjadi pada sampel.

Mahasiswa dikelompokkan berdasarkan *AI Usage Intensity Index* (AUII) yang diukur di akhir periode melalui kuesioner 30 butir skala Likert 1-5 mencakup tiga dimensi, yaitu frekuensi penggunaan, kedalaman interaksi, dan diversitas penggunaan. Skor dinormalisasi ke skala 0-100, lalu dikategorikan menggunakan *percentile-based grouping* menjadi rendah (MP33), sedang (P34-P66), dan tinggi (UP67), sehingga diperoleh Kelompok A (n=16), Kelompok B (n=17), dan Kelompok C (n=17). Pendekatan *data-driven classification* ini menghindari *cut-off* arbitrer dan meningkatkan validitas klasifikasi kelompok. Ketiga dimensi AUII beserta indikatornya disajikan pada Tabel 1.

Tabel 1. Dimensi AI Usage Intensity Index (AUII)

Dimensi	Definisi Operasional	Indikator Utama
Frekuensi penggunaan	Seberapa sering mahasiswa menggunakan Gen AI dalam aktivitas pembelajaran	Intensitas harian; durasi penggunaan; konsistensi penggunaan
Kedalaman interaksi	Tingkat kompleksitas dan kualitas interaksi kognitif dengan Gen AI	Penggunaan prompt kompleks; iterative prompting; evaluasi dan revisi keluaran; dialog reflektif
Diversitas penggunaan	Tingkat variasi konteks dan fungsi penggunaan Gen AI dalam aktivitas akademik	Penulisan akademik; problem solving; brainstorming; analisis konsep

Tiga instrumen utama digunakan dalam penelitian ini. Kuesioner AUII (Cronbach's Alpha = 0,891) mengukur intensitas penggunaan Gen AI. Tes kemampuan berpikir kritis diadaptasi dari *Watson-Glaser Critical Thinking Appraisal* (WGCTA) melalui prosedur adaptasi lintas budaya yang mencakup *forward-translation*, *back-translation*, validasi isi oleh tiga pakar pendidikan dan psikologi kognitif, serta uji coba terbatas (Beaton et al., 2000). Validitas konstruk diuji melalui analisis faktor yang mengonfirmasi terpeliharanya lima dimensi WGCTA, yaitu inferensi, pengenalan asumsi, deduksi, interpretasi, dan evaluasi argumen, dengan konsistensi internal Cronbach's Alpha U 0,80. Tes hasil belajar disusun mengacu pada taksonomi Bloom revisi (Anderson & Krathwohl, 2001) yang mencakup level C1-C6 dan divalidasi melalui *expert judgment* serta uji butir. Literasi AI sebagai variabel moderator diukur dengan skala Ng et al. (2021) yang mencakup dimensi pengetahuan, keterampilan penggunaan, dan kesadaran kritis, kemudian dikategorikan menjadi tinggi dan rendah melalui *median split*. Ringkasan konsistensi instrumen disajikan pada Tabel 2.

Tabel 2. Konsistensi dan Validitas Instrumen Penelitian

Instrumen	Variabel	Reliabilitas	Sumber
Kuesioner AUII	Intensitas penggunaan	0,891	Pengembangan peneliti
AI Literacy Scale	Moderator	U 0,85	Ng et al. (2021)
WGCTA modifikasi	Berpikir kritis	U 0,80	Watson & Glaser
Tes hasil belajar	Hasil belajar	U 0,78	Taksonomi Bloom

Prosedur penelitian terdiri atas pengukuran AUII awal, intervensi berupa *workshop* daring penggunaan Gen AI yang beretika dan efektif (mencakup *prompt engineering*, verifikasi keluaran, dan integrasi AI dalam belajar), periode observasi naturalistik, pengukuran AUII akhir beserta pengelompokan, serta pelaksanaan tes berpikir kritis dan hasil belajar sebagai observasi (O).

Data dianalisis menggunakan *Analysis of Variance* (ANOVA) satu arah untuk menguji perbedaan rerata antarkelompok pada kedua variabel dependen. Uji prasyarat normalitas Shapiro-Wilk dan homogenitas varians Levene dilakukan sebelum ANOVA. Bila hasil signifikan ($p < ,05$), dilanjutkan uji *post hoc* Tukey HSD untuk mengidentifikasi pasangan kelompok yang berbeda, dan ukuran efek dihitung dengan eta-squared (η^2) yang diinterpretasikan sebagai kecil ($\eta^2 < ,06$), sedang ($,06 - ,14$), dan besar ($,14$) mengikuti Cohen (1988). Peran moderasi literasi AI diuji dengan ANOVA dua arah. Seluruh analisis dilakukan menggunakan IBM SPSS Statistics 27.0 pada taraf signifikansi $\alpha = ,05$.

HASIL DAN PEMBAHASAN

Dari 50 mahasiswa, 58% perempuan dan 42% laki-laki dengan rentang usia 19-22 tahun ($M = 20,4$; $SD = 0,87$). Sebanyak 92% melaporkan akses internet di atas 10 Mbps. Platform yang paling banyak digunakan adalah ChatGPT (88%), diikuti Google Gemini (64%), Microsoft Copilot (42%), dan Perplexity AI (28%), dengan rerata lama penggunaan 14,3 bulan ($SD = 6,2$), yang menunjukkan mayoritas sampel merupakan pengguna cukup berpengalaman.

Uji normalitas Shapiro-Wilk menunjukkan distribusi skor berpikir kritis pada ketiga kelompok memenuhi asumsi normalitas, yaitu Kelompok A ($W = ,963$; $p = ,712$), Kelompok B ($W = ,951$; $p = ,541$), dan Kelompok C ($W = ,958$; $p = ,628$); demikian pula skor hasil belajar pada seluruh kelompok ($p > ,05$). Uji homogenitas Levene menunjukkan varians antarkelompok homogen baik untuk berpikir kritis ($F(2,47) = 1,23$; $p = ,301$) maupun hasil belajar ($F(2,47) = 0,98$; $p = ,383$). Dengan terpenuhinya kedua

asumsi tersebut, uji ANOVA dapat dilanjutkan.

Tabel 3. Statistik Deskriptif Berpikir Kritis dan Hasil Belajar per Kelompok

Kelompok	n	Berpikir Kritis M (SD)	Hasil Belajar M (SD)
A – Rendah	16	68,25 (7,12)	70,44 (6,88)
B – Sedang	17	77,41 (5,68)	79,06 (5,43)
C – Tinggi	17	84,71 (5,34)	87,12 (4,91)
Total	50	77,14 (8,94)	79,24 (9,11)

Tabel 3 menunjukkan peningkatan rerata yang monoton dari Kelompok A ke Kelompok C pada kedua variabel. Kelompok C memperoleh rerata tertinggi sekaligus simpangan baku terendah, yang mengindikasikan konsistensi performa lebih baik dibandingkan kelompok lain.

Tabel 4. Ringkasan Hasil Uji ANOVA Satu Arah

Variabel Dependen	F(2,47)	Sig. (p)	η^2	Interpretasi Efek
Berpikir Kritis	18,43	<,001	0,439	Besar
Hasil Belajar	22,17	<,001	0,485	Besar

Hasil ANOVA menolak H_0 dan menerima hipotesis penelitian pada kedua variabel. Nilai F hitung pada keduanya jauh melampaui F tabel ($F_{0,05}(2,47)=3,19$), sehingga H_1 dan H_2 diterima. Intensitas penggunaan Gen AI menjelaskan 43,9% varians berpikir kritis dan 48,5% varians hasil belajar, sebuah ukuran efek yang sangat substansial dalam penelitian pendidikan. Untuk mengetahui pasangan kelompok yang berbeda, dilakukan uji lanjut Tukey HSD yang hasilnya disajikan pada Tabel 5.

Tabel 5. Hasil Uji Post Hoc Tukey HSD

Pasangan Kelompok	Selisih Rerata BK	Sig. BK	Selisih Rerata HB	Sig. HB
A vs B	9,16	,002	8,62	,001
A vs C	16,46	<,001	16,68	<,001
B vs C	7,30	,014	8,06	,009

Keterangan: BK = Berpikir Kritis; HB = Hasil Belajar.

Uji Tukey HSD mengonfirmasi bahwa setiap pasangan kelompok berbeda signifikan ($p<,05$) pada kedua variabel, dengan perbedaan terbesar antara Kelompok A dan Kelompok C. Untuk menguji peran moderasi, ANOVA dua arah menemukan efek interaksi yang signifikan antara intensitas penggunaan dan literasi AI terhadap berpikir kritis ($F(2,44)=7,84$; $p=,001$; $\eta^2=,263$). Mahasiswa dengan literasi AI tinggi dan AUII tinggi memperoleh rerata berpikir kritis tertinggi ($M=91,20$), sedangkan mahasiswa dengan literasi rendah dan AUII tinggi memperoleh rerata lebih rendah ($M=78,14$), sehingga H_3 diterima.

INTENSITAS PENGGUNAAN GEN AI DAN BERPIKIR KRITIS

Perbedaan signifikan kemampuan berpikir kritis antarketiga kelompok ($F(2,47)=18,43$; $p<,001$) konsisten dengan prediksi *Cognitive Load Theory* (Sweller et al., 2019). Namun, interpretasi ini perlu dibatasi secara cermat. Temuan ini berkaitan dengan intensitas penggunaan yang diukur melalui AUII (frekuensi, kedalaman, dan diversitas), bukan semata kualitas interaksi. Dimensi kedalaman dalam AUII memang merepresentasikan aspek kualitatif penggunaan, tetapi analisis utama dilakukan terhadap skor intensitas komposit, sehingga klaim penelitian dibatasi pada hubungan intensitas-kognisi, sementara peran kualitas interaksi diposisikan sebagai penjelasan teoretis yang masih perlu diuji lebih lanjut.

Temuan ini selaras dengan Yan et al. (2024) dan Bewersdorff et al. (2025) yang menemukan bahwa manfaat kognitif Gen AI bergantung pada intensitas dan kualitas keterlibatan mahasiswa. Selisih rerata 16,46 poin antara kelompok AUII tinggi dan rendah menunjukkan bahwa intensitas keterlibatan, bukan sekadar ada-tidaknya penggunaan, berkaitan dengan perkembangan berpikir kritis. Berdasarkan pola interaksi pada kelompok intensitas tinggi.

Sebagai bentuk skeptisisme metodologis, perlu diakui bahwa hubungan yang teramati tidak serta-

merta menunjukkan bahwa Gen AI menyebabkan peningkatan berpikir kritis. Arah hubungan dapat pula bersifat sebaliknya, yaitu mahasiswa yang telah memiliki kemampuan berpikir kritis lebih tinggi cenderung menggunakan Gen AI secara lebih intensif dan beragam. Tanpa *pretest*, kelompok kontrol, dan randomisasi, kemungkinan *reverse causality* dan pengaruh variabel ketiga tidak dapat disingkirkan. Oleh karena itu, temuan ini sebaiknya dibaca sebagai pola asosiatif yang kuat, bukan bukti kausal.

Karena ketiga dimensi diintegrasikan menjadi satu indeks komposit, penelitian ini tidak memisahkan kontribusi masing-masing dimensi secara terpisah. Keterbatasan ini membuka peluang bagi penelitian lanjutan untuk menganalisis kontribusi relatif tiap dimensi melalui regresi berganda atau pemodelan persamaan struktural, sehingga hubungan antara intensitas, kualitas interaksi, dan capaian kognitif dapat dipetakan secara lebih presisi. Secara analitis, hasil ini memberikan indikasi empiris yang relevan untuk interpretasi ulang dalam kerangka CLT, dengan menempatkan Gen AI sebagai faktor eksternal yang dapat menggeser distribusi beban kognitif menuju *germane load* melalui interaksi yang reflektif.

Perlu ditegaskan bahwa hubungan ini bersifat kontingen, bukan deterministik. Pada kelompok intensitas rendah, penggunaan Gen AI cenderung berupa substitusi berpikir yang berisiko menimbulkan *automation bias* dan *cognitive offloading* berlebihan, sebagaimana diperingatkan Gerlich (2025). Sebaliknya, pada kelompok intensitas tinggi, Gen AI berperan sebagai *epistemic partner* yang memicu regulasi metakognitif. Dengan demikian, manfaat kognitif tidak ditentukan oleh keberadaan teknologi, melainkan oleh transformasi peran pengguna dari penerima jawaban menjadi pengonstruksi pengetahuan.

INTENSITAS PENGGUNAAN GEN AI DAN HASIL BELAJAR

Hasil ANOVA untuk hasil belajar ($F(2,47)=22,17$; $p<,001$; $\eta^2=,485$) menunjukkan intensitas penggunaan Gen AI berkaitan erat dengan capaian pembelajaran. Temuan ini memperkuat tinjauan Tlili et al. (2023) mengenai dampak positif AI terhadap hasil belajar kognitif. Dalam perspektif *Zone of Proximal Development* (Vygotsky, 1978), Gen AI berperan sebagai *scaffolding* dinamis yang memperluas zona perkembangan proksimal mahasiswa (Luckin, 2017), tercermin pada rerata hasil belajar Kelompok C ($M=87,12$) yang tertinggi.

Temuan ini relevan dengan konteks Indonesia. Karena 68,8% mahasiswa menggunakan Gen AI secara reaktif tanpa literasi memadai (*Indonesian Digital Literacy Survey*, 2023), peningkatan intensitas tanpa pendampingan literasi berisiko menimbulkan ketergantungan. Variabilitas Kelompok B yang lebih besar ($SD=5,43$) mengindikasikan pola penggunaan yang belum optimal. Variabel kontrol seperti lama penggunaan sebelumnya dan jenis platform telah dipertimbangkan dengan memastikan kesetaraan kriteria sampel antarkelompok, sehingga perbedaan capaian lebih dapat diatribusikan pada intensitas penggunaan, meskipun pengaruh faktor laten lain tetap tidak dapat sepenuhnya dikesampingkan.

Lebih jauh, pola peningkatan yang monoton dari Kelompok A ke Kelompok C mengindikasikan adanya hubungan bertingkat (*dose-response*) antara intensitas penggunaan dan hasil belajar. Namun, tanpa literasi AI yang memadai, intensitas tinggi justru berpotensi mendorong *surface learning* berbasis reproduksi jawaban, bukan konstruksi makna. Hal ini sejalan dengan Yan et al. (2024) yang menekankan bahwa hasil pembelajaran berbasis AI sangat bergantung pada orkestrasi pedagogis, bukan pada ketersediaan teknologi semata.

. Dalam lingkungan dengan kesiapan dosen yang belum merata (Lembaga Layanan Pendidikan Tinggi, 2023), peran literasi AI menjadi semakin menentukan, karena mahasiswa kerap mengembangkan pola penggunaan secara mandiri tanpa bimbingan terstruktur. Oleh karena itu, hasil penelitian ini menggarisbawahi pentingnya intervensi kelembagaan yang kontekstual, bukan sekadar mengadopsi model pemanfaatan AI dari konteks lain.

PERAN MODERASI LITERASI AI

Efek interaksi yang signifikan antara intensitas penggunaan dan literasi AI ($F(2,44)=7,84$; $p=,001$) menunjukkan bahwa literasi AI merupakan prasyarat untuk memaksimalkan manfaat Gen AI. Temuan ini memperluas konstruk literasi AI (Long & Magerko, 2020; Ng et al., 2021) ke domain

dampak pembelajaran yang lebih operasional. Perbedaan mencolok antara mahasiswa literasi tinggi ($M=91,20$) dan rendah ($M=78,14$) pada Kelompok C menegaskan bahwa investasi pada literasi AI memberikan manfaat lebih besar daripada sekadar memperluas akses, sejalan dengan rekomendasi UNESCO (2023). Literasi AI berfungsi sebagai *cognitive filter* yang menentukan bagaimana keluaran AI diproses, diverifikasi, dan diintegrasikan menjadi pengetahuan bermakna, sekaligus menjadi pelindung terhadap *surface learning* dan *automation bias*.

Implikasi penting dari temuan moderasi ini adalah bahwa kebijakan integrasi Gen AI tidak cukup berfokus pada penyediaan akses teknologi. Tanpa literasi yang memadai, peningkatan intensitas penggunaan berpotensi menghasilkan *illusory learning effect*, yaitu persepsi memahami materi padahal hanya mengandalkan reproduksi keluaran AI. Hal ini menjelaskan mengapa kelompok berliterasi rendah dengan frekuensi tinggi tetap menunjukkan performa lebih rendah dan variabilitas lebih besar. Karena itu, penguatan literasi AI yang mencakup evaluasi sumber, refleksi epistemik, dan etika penggunaan menjadi komponen yang tidak terpisahkan dari strategi pemanfaatan Gen AI di pendidikan tinggi.

Secara keseluruhan, ketiga temuan penelitian ini membentuk satu simpulan analitis yang koheren, yaitu intensitas penggunaan Gen AI berkaitan positif dengan capaian kognitif mahasiswa, tetapi kekuatan hubungan tersebut dimoderasi oleh literasi AI. Dengan demikian, kebijakan pendidikan tinggi tidak dapat memperlakukan akses teknologi dan kompetensi literasi sebagai dua hal yang terpisah. Model yang muncul dari penelitian ini menempatkan literasi AI sebagai jembatan yang menentukan apakah intensitas penggunaan Gen AI berujung pada penguatan kognisi atau justru pada pelemahannya, sebuah kerangka yang memperkaya diskursus pemanfaatan AI dalam pendidikan di konteks Indonesia.

KETERBATASAN DAN IMPLIKASI

Beberapa keterbatasan perlu diakui dalam menafsirkan hasil ini. Pertama, desain *one-shot case study* tanpa kelompok kontrol dan tanpa randomisasi membatasi inferensi kausal, sehingga seluruh temuan diposisikan sebagai pola asosiatif. Kedua, sampel terbatas pada satu program studi di satu universitas, sehingga generalisasi temuan perlu dibatasi. Ketiga, pengukuran intensitas melalui *self-report* berpotensi mengandung bias keinginan sosial. Keempat, durasi empat minggu belum cukup untuk menangkap dampak jangka panjang. Pengakuan keterbatasan ini penting agar pembaca tidak menafsirkan hasil secara berlebihan.

SIMPULAN DAN SARAN

Penelitian ini menyimpulkan, sesuai dengan tujuan dan rumusan masalah, bahwa terdapat perbedaan yang signifikan dalam kemampuan berpikir kritis dan hasil belajar mahasiswa berdasarkan intensitas penggunaan Gen AI, dengan kelompok intensitas tinggi memperoleh capaian tertinggi. Literasi AI terbukti memoderasi secara signifikan hubungan antara intensitas penggunaan Gen AI dan kemampuan berpikir kritis. Dengan demikian, intensitas penggunaan Gen AI berkaitan positif dengan capaian kognitif mahasiswa, dan kekuatan hubungan tersebut bergantung pada tingkat literasi AI.

Secara teoretis, hasil penelitian memberikan kontribusi pada pengembangan kerangka konseptual yang menempatkan literasi AI sebagai moderator hubungan antara intensitas penggunaan Gen AI dan capaian pembelajaran. Secara praktis, perguruan tinggi disarankan mengintegrasikan Gen AI secara terstruktur melalui pendekatan *guided AI learning* serta menjadikan pengembangan literasi AI sebagai prioritas kurikulum yang mencakup aspek teknis, epistemologis, dan etis.

Penelitian lanjutan direkomendasikan menggunakan desain yang lebih kuat, seperti eksperimen dengan kelompok kontrol atau analisis *moderated regression* dan *SEM-PLS*, cakupan sampel yang lebih luas, serta analisis kualitatif terhadap pola interaksi untuk memperdalam pemahaman tentang interaksi manusia-AI dalam pendidikan. Sebagai prospek pengembangan, hasil penelitian ini dapat menjadi dasar perumusan standar literasi AI bagi pendidikan tinggi serta pengembangan model pembelajaran berbantuan AI yang berorientasi pada penguatan berpikir kritis, yang pada kajian berikutnya diharapkan diuji pada beragam disiplin ilmu dan jenjang pendidikan.

PUSTAKA ACUAN

- Anderson, L. W., & Krathwohl, D. R. (2001). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Asosiasi Penyelenggara Jasa Internet Indonesia. (2023). *Laporan survei penetrasi internet Indonesia 2023*. APJII Press.
- Badan Pusat Statistik. (2023). *Statistik pendidikan tinggi Indonesia 2023*. BPS RI.
- Beaton, D. E., Bombardier, C., Guillemin, F., & Ferraz, M. B. (2000). Guidelines for the process of cross-cultural adaptation of self-report measures. *Spine*, 25(24), 3186–3191. <https://doi.org/10.1097/00007632-200012150-00014>
- Bewersdorff, A., Hartmann, C., Hornberger, M., Seßler, K., Bannert, M., Kasneci, E., Kasneci, G., Zhai, X., & Nerdel, C. (2025). Taking the next step with generative artificial intelligence: The transformative role of multimodal large language models in science education. *Learning and Individual Differences*, 118, 102601. <https://doi.org/10.1016/j.lindif.2025.102601>
- Campbell, D. T., & Stanley, J. C. (1963). *Experimental and quasi-experimental designs for research*. Rand McNally.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Feuerriegel, S., Hartmann, J., Janiesch, C., & Zschech, P. (2024). Generative AI. *Business & Information Systems Engineering*, 66(1), 111–126. <https://doi.org/10.1007/s12599-023-00834-7>
- Gerlich, M. (2025). AI tools in society: Impacts on cognitive offloading and the future of critical thinking. *Societies*, 15(1), 6. <https://doi.org/10.3390/soc15010006>
- Gilson, A., Safraneck, C. W., Huang, T., Socrates, V., Chi, L., Taylor, R. A., & Chartash, D. (2023). How does ChatGPT perform on the United States Medical Licensing Examination? The implications of large language models for medical education and knowledge assessment. *JMIR Medical Education*, 9(1), e45312. <https://doi.org/10.2196/45312>
- Indonesian Digital Literacy Survey. (2023). *AI literacy among Indonesian higher education students 2023*. IDLS Institute.
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., ... 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>
- Kemdikbudristek. (2023). *Laporan survei penggunaan kecerdasan buatan dalam pembelajaran mahasiswa Indonesia*. Kemendikbudristek Press.
- Kung, T. H., Cheatham, M., Medenilla, A., Sillos, C., De Leon, L., Elepaño, C., ... Tseng, V. (2023). Performance of ChatGPT on USMLE: Potential for AI-assisted medical education using large language models. *PLOS Digital Health*, 2(2), e0000198. <https://doi.org/10.1371/journal.pdig.0000198>
- Lembaga Layanan Pendidikan Tinggi. (2023). *Survei kesiapan dosen dalam menghadapi era AI di perguruan tinggi*. LL-Dikti Press.
- Long, D., & Magerko, B. (2020). What is AI literacy? Competencies and design considerations. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1–16. <https://doi.org/10.1145/3313831.3376727>
- Luckin, R. (2017). Towards artificial intelligence-based assessment systems. *Nature Human Behaviour*, 1(3), 0028. <https://doi.org/10.1038/s41562-016-0028>
- 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, 100041. <https://doi.org/10.1016/j.caeai.2021.100041>
- Nugroho, A., Syarif, M., & Prasetyo, D. (2024). Pola penggunaan kecerdasan buatan generatif di kalangan mahasiswa Indonesia: Sebuah studi nasional. *Jurnal Teknologi Pendidikan Indonesia*, 14(1), 45–62. <https://doi.org/10.26737/jtpi.v14i1.4401>
- Sweller, J., van Merriënboer, J. J. G., & Paas, F. (2019). Cognitive architecture and instructional design: 20 years later. *Educational Psychology Review*, 31(2), 261–292. <https://doi.org/10.1007/s10648-019-09465-5>
- Tlili, A., Shehata, B., Adarkwah, M. A., Bozkurt, A., Hickey, D. T., Huang, R., & Agyemang, B. (2023). What if the devil is my guardian angel: ChatGPT as a case study of using chatbots in education. *Smart Learning Environments*, 10(15), 1–24. <https://doi.org/10.1186/s40561-023-00237-x>
- UNESCO. (2023). *Guidance for generative AI in education and research*. UNESCO Publishing. <https://doi.org/10.54675/HMBB1350>
- Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Harvard University Press.
- Wardat, Y., Tashtoush, M. A., AIAli, R., & Jarrah, A. M. (2023). ChatGPT: A revolutionary tool for teaching and learning mathematics. *EURASIA Journal of Mathematics, Science and Technology Education*, 19(7), em2286. <https://doi.org/10.29333/ejmste/13272>
- Yan, L., Sha, L., Zhao, L., Li, Y., Martinez-Maldonado, R., Chen, G., Li, X., Jin, Y., & Gašević, D. (2024). Practical and ethical challenges of large language models in education: A systematic scoping review. *British Journal of Educational Technology*, 55(1), 90–112. <https://doi.org/10.1111/bjet.13370>