

**PENGARUH MODEL PROBLEM BASED LEARNING
BERPENDEKATAN PEMBELAJARAN MENDALAM
BERBANTUAN AI GENERATIF TERHADAP
KEMAMPUAN BERPIKIR KOMPUTASIONAL
DAN NUMERASI SISWA SMP**

Mustika Fitri Larasati Sibuea¹, Muhammad Ardiansyah Sembiring²,

Yudi Firdiansyah Putra Sitorus³

Pendidikan Matematika, Universitas Royal

e-mail: bukmus.inaction@gmail.com¹

Abstract: Indonesian students' mathematics achievement remains low: the 2026 national academic ability test (TKA) for junior high schools recorded an average mathematics score of 40.34, and Indonesia scored 364 in PISA 2025 mathematics, far below the OECD average of 463. Meanwhile, the government has introduced the Deep Learning approach and the Coding and Artificial Intelligence subject, which require computational thinking, while unrestricted generative AI has been shown to harm mathematics learning. This study examined the effect of Problem Based Learning with a Deep Learning approach assisted by guardrailed generative AI (PBL-DL + AI) on the computational thinking and numeracy of junior high school students. A quasi-experimental nonequivalent pretest-posttest control group design involved 95 eighth-grade students in three groups: PBL-DL + AI, PBL-DL without AI, and conventional instruction, on systems of linear equations in two variables. Data were analyzed using MANCOVA, ANCOVA, and Bonferroni post hoc tests. The results showed a significant multivariate effect (Wilks' $\lambda = 0.752$, $F = 6.81$, $p < .001$, $\eta^2 = 0.13$). For computational thinking, only the PBL-DL + AI group significantly outperformed the conventional group ($d = 1.02$), with the largest gains in abstraction and algorithm indicators. For numeracy, both PBL-DL groups outperformed the conventional group, with no additional benefit from AI. Guardrailed AI should therefore be positioned as a tutor that scaffolds thinking, not a replacement for meaningful contextual problems.

Keywords: Problem Based Learning, Deep Learning Approach, Generative AI, Computational Thinking, Numeracy

Abstrak: Capaian matematika siswa Indonesia masih rendah: rata-rata nilai matematika Tes Kemampuan Akademik (TKA) SMP 2026 hanya 40,34, dan skor matematika PISA 2025 Indonesia sebesar 364, jauh di bawah rata-rata OECD (463). Di sisi lain, pemerintah menerapkan pendekatan Pembelajaran Mendalam dan mata pelajaran Koding dan Kecerdasan Artifisial yang menuntut kemampuan berpikir komputasional, sementara penggunaan AI generatif tanpa pembatas terbukti dapat merugikan pembelajaran matematika. Penelitian ini bertujuan menganalisis pengaruh model *Problem Based Learning* berpendekatan Pembelajaran Mendalam berbantuan AI generatif ber-guardrail (PBL-PM + AI) terhadap kemampuan berpikir komputasional dan numerasi siswa SMP. Penelitian kuasi-eksperimen dengan desain *nonequivalent pretest-posttest control group* melibatkan 95 siswa kelas VIII dalam tiga kelompok, yaitu PBL-PM + AI, PBL-PM tanpa AI, dan pembelajaran konvensional, pada materi sistem persamaan linear dua variabel. Data dianalisis dengan MANCOVA, ANCOVA, dan uji lanjut Bonferroni. Hasil menunjukkan pengaruh multivariat yang signifikan (Wilks' $\lambda = 0,752$; $F = 6,81$; $p < 0,001$; $\eta^2 p = 0,13$). Pada berpikir komputasional, hanya kelompok PBL-PM + AI yang unggul signifikan atas kelompok konvensional ($d = 1,02$), dengan peningkatan terbesar pada indikator abstraksi dan algoritma. Pada numerasi, kedua kelompok PBL-PM unggul atas kelompok konvensional tanpa tambahan manfaat dari AI. AI ber-guardrail sebaiknya diposisikan sebagai tutor yang membimbing proses berpikir, bukan pengganti masalah kontekstual yang bermakna.

Kata Kunci: *Problem Based Learning*, Pembelajaran Mendalam, AI Generatif, Berpikir Komputasional, Numerasi.

PENDAHULUAN

Capaian matematika siswa Indonesia masih menjadi persoalan serius. Hasil Tes Kemampuan Akademik (TKA) jenjang SMP tahun 2026 menunjukkan rata-rata nasional matematika hanya 40,34, jauh di bawah Bahasa Indonesia sebesar 60,83 (Tribunnews.com, 2026). Hasil PISA 2025 juga menunjukkan skor matematika siswa Indonesia sebesar 364, jauh di bawah rata-rata OECD sebesar 463, dan tidak berubah secara signifikan dibandingkan PISA 2022 (OECD, 2026). Kondisi ini menunjukkan bahwa siswa belum terbiasa menggunakan matematika untuk bernalar dan menyelesaikan masalah kontekstual, yang merupakan inti kemampuan numerasi.

Pemerintah merespons kondisi tersebut melalui dua kebijakan terbaru. Kementerian Pendidikan Dasar dan Menengah menerapkan pendekatan Pembelajaran Mendalam (*deep learning*) pada semua mata pelajaran dengan penerapan yang disesuaikan dengan karakteristik tiap mata pelajaran (ANTARA News, 2026). Pendekatan ini menekankan pembelajaran yang berkesadaran (*mindful*), bermakna (*meaningful*), dan menggembirakan (*joyful*). Selain itu, Permendikdasmen Nomor 13 Tahun 2025 menetapkan mata pelajaran Koding dan Kecerdasan Artifisial yang diterapkan bertahap sejak tahun ajaran 2025/2026 untuk membekali siswa dengan kemampuan berpikir logis dan pemecahan masalah (ANTARA News, 2025).

Kedua kebijakan tersebut menempatkan berpikir komputasional sebagai kemampuan penting, yaitu kemampuan menguraikan masalah (dekomposisi), mengenali pola, melakukan abstraksi, dan menyusun algoritma (Wing, 2006; Weintrop et al., 2016). Kerangka PISA 2022 bahkan memasukkan berpikir komputasional ke dalam literasi matematika (OECD, 2023). Namun, penelitian yang dilakukan oleh Mustaqimah & Ni'mah (2024) menunjukkan capaian berpikir

komputasional siswa kelas VIII hanya 43%, dengan indikator abstraksi (23%) dan algoritma (36%) sebagai yang terendah.

Model *Problem Based Learning* (PBL) menghadirkan masalah autentik sebagai titik awal pembelajaran dan menempatkan guru sebagai fasilitator (Hmelo-Silver, 2004). Penelitian sebelumnya yang dilakukan oleh Malau et al. (2025) membuktikan bahwa PBL berpengaruh signifikan terhadap kemampuan berpikir komputasional siswa kelas VIII pada materi sistem persamaan linear dua variabel (SPLDV). Tinjauan sistematis Ye et al. (2023) juga menunjukkan bahwa integrasi berpikir komputasional melalui aktivitas pemecahan masalah umumnya meningkatkan hasil belajar matematika. Akan tetapi, penelitian tersebut belum mengintegrasikan prinsip Pembelajaran Mendalam dan belum memanfaatkan AI generatif.

AI generatif seperti ChatGPT dan Gemini dapat berperan sebagai tutor yang memberikan umpan balik personal, tetapi juga berisiko menimbulkan ketergantungan dan melemahkan kemampuan berpikir kritis (Kasneci et al., 2023). Eksperimen lapangan Bastani et al. (2025) pada hampir 1.000 siswa sekolah menengah menunjukkan bahwa akses AI tanpa pembatas menurunkan hasil ujian matematika, sedangkan tutor AI yang dirancang untuk memberi petunjuk bertahap mampu meredam dampak negatif tersebut. Temuan ini menunjukkan bahwa manfaat AI bergantung pada cara AI dirancang dan diintegrasikan ke dalam model pembelajaran.

Berdasarkan uraian tersebut, penelitian ini bertujuan menganalisis pengaruh model PBL berpendekatan Pembelajaran Mendalam berbantuan AI generatif ber-*guardrail* (PBL-PM + AI) terhadap kemampuan berpikir komputasional dan numerasi siswa SMP. Untuk memisahkan pengaruh AI dari pengaruh model, penelitian ini membandingkan tiga kelompok, yaitu PBL-

PM + AI, PBL-PM tanpa AI, dan pembelajaran konvensional.

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode kuasi-eksperimen dan desain *nonequivalent pretest-posttest control group design* (Shadish et al., 2002). Tahapan penelitian ditunjukkan pada Gambar 1.



Gambar 1 Tahapan Penelitian

Desain, Populasi, dan Sampel

Desain kelompok ditunjukkan pada Tabel 1, dengan O1 = pretest, O2 = posttest, X1 = PBL-PM + AI, dan X2 = PBL-PM tanpa AI. Populasi penelitian adalah siswa kelas VIII SMP Negeri [nama sekolah] di [kota]. Tiga kelas dipilih dengan teknik *cluster random sampling* lalu diundi untuk menentukan kelompok, sehingga diperoleh 95 siswa: 32 siswa PBL-PM + AI, 31 siswa PBL-PM, dan 32 siswa konvensional.

Tabel 1 Desain Penelitian

Kelompok	Pre	Perlakuan	Post
Eksperimen 1 (PBL-PM + AI)	O1	X1	O2
Eksperimen 2 (PBL-PM)	O1	X2	O2
Kontrol (Konvensional)	O1	—	O2

Prosedur Perlakuan

Perlakuan diberikan selama enam pertemuan (2×40 menit) pada materi SPLDV. Sintaks PBL (orientasi masalah, organisasi belajar, penyelidikan, penyajian hasil, serta analisis dan evaluasi) dipadukan dengan tiga prinsip Pembelajaran Mendalam: berkesadaran melalui penetapan tujuan dan refleksi diri; bermakna melalui masalah kontekstual seperti perbandingan tarif paket data dan harga bahan pokok; serta menggembirakan melalui kerja kelompok dan tantangan bertingkat. Kelas PBL-PM + AI dapat berkonsultasi dengan tutor AI yang diberi instruksi sistem (*prompt*) agar hanya memberikan pertanyaan pancingan, petunjuk bertahap, dan umpan balik, bukan jawaban akhir. Kelas PBL-PM menggunakan modul yang sama tanpa AI, sedangkan kelas kontrol belajar dengan penjelasan guru, contoh soal, dan latihan.

Instrumen

Instrumen terdiri atas tes berpikir komputasional berbentuk uraian (8 butir; indikator dekomposisi, pengenalan pola, abstraksi, dan algoritma), tes numerasi (20 butir), angket pengalaman Pembelajaran Mendalam (15 butir skala Likert 1–5), dan lembar observasi keterlaksanaan. Validitas isi dinilai oleh tiga ahli dan dua guru menggunakan indeks Aiken's V dengan kriteria $V \geq 0,80$ (Aiken, 1985), sedangkan reliabilitas diuji dengan Cronbach's alpha ($\geq 0,70$) pada kelas uji coba.

Teknik Analisis Data

Data dianalisis secara deskriptif dan inferensial. Peningkatan kemampuan dihitung dengan N-gain pada persamaan (1) dan dikategorikan tinggi ($g \geq 0,7$), sedang

($0,3 \leq g < 0,7$), atau rendah ($g < 0,3$) (Hake, 1998).

$$g = (\text{post} - \text{pre}) / (100 - \text{pre}) \quad (1)$$

Kesetaraan awal diuji dengan ANOVA satu jalur pada skor pretest. Uji prasyarat meliputi normalitas residual (Shapiro–Wilk), homogenitas varians (Levene), dan homogenitas kemiringan regresi. Hipotesis diuji dengan MANCOVA (Wilks' lambda) menggunakan skor pretest sebagai kovariat, dilanjutkan ANCOVA univariat dan uji lanjut berpasangan dengan koreksi Bonferroni. Besar efek dilaporkan dengan eta kuadrat parsial (η^2p) dan Cohen's d (Cohen, 1988). Taraf signifikansi yang digunakan adalah $\alpha = 0,05$.

HASIL DAN PEMBAHASAN

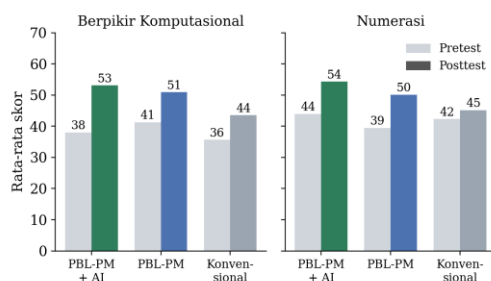
Deskripsi Data

Statistik deskriptif skor pretest, posttest, dan N-gain ditunjukkan pada Tabel 2 dan Gambar 2.

Tabel 2 Rata-rata Skor dan N-gain

Kelompok	Pre	Post	N-gain
PBL-PM + AI (BK)	38,03	53,12	0,24
PBL-PM (BK)	41,29	50,94	0,16
Konvensional (BK)	35,69	43,53	0,11
PBL-PM + AI (NUM)	43,94	54,34	0,20
PBL-PM (NUM)	39,48	50,13	0,18
Konvensional (NUM)	42,31	45,16	0,05

Keterangan: BK = berpikir komputasional; NUM = numerasi.



Gambar 2 Rata-rata Skor Pretest dan Posttest

Seluruh kelompok mengalami peningkatan. Peningkatan berpikir komputasional terbesar terjadi pada kelompok PBL-PM + AI, sedangkan peningkatan numerasi kelompok PBL-PM + AI dan PBL-PM relatif sama dan jauh di atas kelompok konvensional. Rata-rata N-gain seluruh kelompok masih berada pada kategori rendah. Uji ANOVA menunjukkan kemampuan awal berpikir komputasional antarkelompok tidak setara ($F = 3,18$; $p = 0,046$), sedangkan kemampuan awal numerasi setara ($F = 1,60$; $p = 0,208$), sehingga penggunaan skor pretest sebagai kovariat menjadi penting.

Uji Prasyarat dan Uji Hipotesis

Uji Shapiro–Wilk menunjukkan residual berdistribusi normal pada semua kelompok (p terkecil = 0,339 untuk berpikir komputasional dan = 0,510 untuk numerasi). Uji Levene menunjukkan varians homogen ($p = 0,060$ dan $p = 0,505$), dan kemiringan regresi juga homogen ($p = 0,118$ dan $p = 0,733$). Dengan demikian, analisis kovarians dapat dilanjutkan.

Hasil MANCOVA menunjukkan perbedaan simultan yang signifikan pada kedua kemampuan antarkelompok setelah mengontrol pretest (Wilks' $\lambda = 0,752$; $F(4; 178) = 6,81$; $p < 0,001$; $\eta^2p = 0,13$). Hasil ANCOVA univariat ditunjukkan pada Tabel 3.

Tabel 3 Hasil ANCOVA dan Rata-rata Terkoreksi

Var.	F	η^2p	E1	E2	K
BK	8,17	0,15	53,3	49,1	45,1
NUM	10,03	0,18	52,4	52,5	44,8

Keterangan: E1 = PBL-PM + AI; E2 = PBL-PM; K = konvensional; seluruh $p < 0,001$.

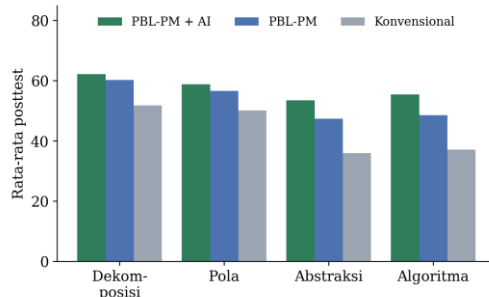
Model pembelajaran berpengaruh signifikan terhadap berpikir komputasional ($\eta^2p = 0,15$) dan numerasi ($\eta^2p = 0,18$), keduanya dengan efek besar. Hasil uji lanjut ditunjukkan pada Tabel 4.

Tabel 4 Uji Lanjut Bonferroni

Pasangan	Δ BK	p	Δ NUM	p
E1 vs K	8,15	< 0,001	7,62	< 0,001
E2 vs K	3,96	0,183	7,69	< 0,001
E1 vs E2	4,19	0,130	-0,07	1,000

Pembahasan

Pada berpikir komputasional, hanya kelompok PBL-PM + AI yang unggul signifikan atas kelompok konvensional (selisih 8,15 poin; $d = 1,02$). Kelompok PBL-PM tanpa AI unggul 3,96 poin atas kelompok konvensional, tetapi belum signifikan setelah koreksi Bonferroni. Gambar 3 menunjukkan bahwa selisih terbesar antara kelompok PBL-PM + AI dan konvensional terdapat pada indikator abstraksi (53,5 vs 36,0) dan algoritma (55,5 vs 37,2), yaitu dua indikator yang oleh Mustaqimah & Ni'mah (2024) dilaporkan paling rendah.



Gambar 3 Skor Posttest Per Indikator Berpikir Komputasional

Temuan ini diduga terjadi karena petunjuk bertahap dari tutor AI membantu siswa mengubah masalah cerita menjadi model persamaan (abstraksi) dan menyusun langkah penyelesaian (algoritma) tanpa memberikan jawaban akhir. Hasil ini sejalan dengan Bastani et al. (2025) bahwa AI yang dirancang dengan pembatas dapat mendukung proses berpikir siswa, serta memperkuat temuan Malau et al. (2025) dan Ye et al. (2023) tentang efektivitas pembelajaran berbasis masalah terhadap berpikir komputasional.

Pada numerasi, kedua kelompok PBL-PM sama-sama unggul signifikan atas kelompok konvensional (selisih 7,62 dan 7,69 poin), sedangkan kelompok dengan dan tanpa AI tidak berbeda ($p = 1,000$). Artinya, peningkatan numerasi terutama berasal dari masalah kontekstual yang bermakna dalam model PBL-PM, bukan

dari AI. Hal ini sesuai dengan kerangka PISA yang menekankan kemampuan merumuskan, menerapkan, dan menafsirkan matematika dalam konteks nyata (OECD, 2023), serta dengan karakter PBL yang berangkat dari masalah autentik (Hmelo-Silver, 2004).

Hasil angket menunjukkan siswa pada kelas PBL-PM + AI dan PBL-PM menilai pembelajaran berkesadaran (4,10 dan 3,98), bermakna (3,98 dan 4,04), dan menggembirakan (4,17 dan 4,10) dalam kategori baik. Meskipun demikian, N-gain yang masih rendah menunjukkan bahwa enam pertemuan pada satu materi belum cukup untuk menghasilkan peningkatan besar.

SIMPULAN

Berdasarkan hasil penelitian dan pembahasan, dapat ditarik kesimpulan sebagai berikut.

1. Terdapat perbedaan kemampuan berpikir komputasional dan numerasi secara simultan antara kelompok PBL-PM + AI, PBL-PM, dan konvensional setelah mengontrol kemampuan awal ($\eta^2p = 0,13$).
2. Model PBL-PM + AI memberikan rata-rata terkoreksi berpikir komputasional tertinggi (53,29) dan unggul signifikan atas pembelajaran konvensional dengan efek besar, terutama pada indikator abstraksi dan algoritma.
3. Kedua model PBL-PM meningkatkan numerasi secara signifikan dibanding pembelajaran konvensional, tetapi penambahan AI tidak memberikan tambahan peningkatan numerasi.
4. Siswa menilai pengalaman belajar berkesadaran, bermakna, dan menggembirakan dalam kategori baik.

Guru disarankan menerapkan PBL berpendekatan Pembelajaran Mendalam dengan masalah kontekstual dan

menggunakan AI hanya sebagai tutor *ber-guardrail*. Penelitian lanjutan perlu memperpanjang durasi perlakuan, memperluas materi dan jumlah sekolah, serta mengkaji variabel lain seperti kecemasan matematika dan regulasi diri.

DAFTAR PUSTAKA

- Aiken, L. R. (1985). Three Coefficients for Analyzing the Reliability and Validity of Ratings. *Educational and Psychological Measurement*, 45(1), 131–142.
<https://doi.org/10.1177/0013164485451012>
- ANTARA News. (2025, Juli 22). Permendikdasmen 13/2025: Koding dan AI Resmi Masuk Kurikulum 2025/2026.
<https://www.antaranews.com/berita/4983541/permendikdasmen-13-2025-koding-dan-ai-resmi-masuk-kurikulum-2025-2026>
- ANTARA News. (2026, Juli 10). Kemendikdasmen Siap Terapkan Deep Learning di Semua Mata Pelajaran.
<https://www.antaranews.com/berita/5643915/kemendikdasmen-siap-terapkan-deep-learning-di-semua-mata-pelajaran>
- Bastani, H., Bastani, O., Sungu, A., Ge, H., Kabakcı, Ö., & Mariman, R. (2025). Generative AI without Guardrails Can Harm Learning: Evidence from High School Mathematics. *Proceedings of the National Academy of Sciences*, 122(26), e2422633122.
<https://doi.org/10.1073/pnas.2422633122>
- Cohen, J. (1988). *Statistical Power Analysis for the Behavioral Sciences* (2nd ed.). Lawrence Erlbaum Associates.
- Hake, R. R. (1998). Interactive-Engagement versus Traditional Methods: A Six-Thousand-Student Survey of Mechanics Test Data for Introductory Physics Courses. *American Journal of Physics*, 66(1), 64–74.
<https://doi.org/10.1119/1.18809>
- Hmelo-Silver, C. E. (2004). Problem-Based Learning: What and How Do Students Learn? *Educational Psychology Review*, 16(3), 235–266.
<https://doi.org/10.1023/B:EDPR.0000034022.16470.f3>
- Kasneci, E., Seßler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., et al. (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>
- Malau, O. L., Sidabutar, R., & Tambunan, L. O. (2025). Pengaruh Model Problem Based Learning terhadap Kemampuan Berpikir Komputasi Siswa Kelas VIII SMP Cinta Rakyat 3 Pematangsiantar. *MES: Journal of Mathematics Education and Science*, 11(1), 46–52.
<https://doi.org/10.30743/mes.v11i1.11926>
- Mustaqimah, U. P. S., & Ni'mah, K. (2024). Profil Kemampuan Berpikir Komputasi Siswa SMP pada Soal Tantangan Bebras. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, 7(2), 297–308.
<https://doi.org/10.22460/jpmi.v7i2.21590>
- OECD. (2023). *PISA 2022 Assessment and Analytical Framework*. OECD Publishing.
<https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2026). Education GPS – Indonesia: Student Performance (PISA 2025).
<https://gpseducation.oecd.org/CountryProfile?primaryCountry=IDN&topic=PI>
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.
- Tribunnews.com. (2026, Mei). Hasil TKA SMP 2026: Nilai Rata-Rata Matematika 40,34, Bahasa Indonesia 60,83.
<https://www.tribunnews.com/nasional/7834348/hasil-tka-smp-2026-nilai-rata-rata-matematika-4034-bahasa-indonesia-6083>
- Weintrop, D., Beheshti, E., Horn, M., Orton, K., Jona, K., Trouille, L., & Wilensky, U. (2016). Defining Computational Thinking for Mathematics and Science Classrooms. *Journal of Science Education and*

Technology, 25(1), 127–147.
<https://doi.org/10.1007/s10956-015-9581-5>

Wing, J. M. (2006). Computational Thinking. *Communications of the ACM*, 49(3), 33–35.
<https://doi.org/10.1145/1118178.1118215>

Ye, H., Liang, B., Ng, O.-L., & Chai, C. S. (2023). Integration of Computational Thinking in K-12 Mathematics Education: A Systematic Review on CT-based Mathematics Instruction and Student Learning. *International Journal of STEM Education*, 10, 3.
<https://doi.org/10.1186/s40594-023-00396-w>