BAB 04

Apa yang ditunjukkan **eksperimen** dan apa yang belum

Lima studi menjawab pertanyaan yang berbeda

Tabel berikut menyajikan temuan ringkas dan batas inferensinya. Besaran efek tidak boleh dijumlahkan atau dipakai sebagai peringkat produk.

Studi dan konteks	Hasil yang relevan	Batas penting
Bastani dkk. 2025 [20], matematika SMA di Türkiye	48% · 17% Latihan dengan GPT Base membaik 48%; tes tanpa AI menurun 17% relatif terhadap kontrol	Desain akses dan tahap penilaian menentukan arti hasil
Kestin dkk. 2025 [21], fisika universitas di US	Tutor khusus menghasilkan tes segera lebih baik daripada kelas pembelajaran aktif	Dua pelajaran; belum membuktikan retensi panjang
Wang dkk. 2025 [22], Tutor CoPilot di US	4 poin persentase Penguasaan topik naik 4 poin persentase; tes akhir tahun tidak meningkat signifikan	Outcome dekat dan outcome jauh berbeda
NFER dan EEF 2024 [23], guru sains England	56,2 vs 81,5 Persiapan 56,2 dibanding 81,5 menit per minggu	Ukuran produktivitas guru, bukan hasil belajar siswa
De Simone dkk. 2025 [24], Nigeria	0,31 Asesmen gabungan meningkat 0,31 simpangan baku	Program gabungan; kehilangan tindak lanjut membatasi inferensi

Identitas Dokumen

LAPORAN	**Kajian Komparatif AI dalam Pendidikan untuk Indonesia** Pendidikan dasar hingga perguruan tinggi. Laporan utama dan pilihan implementasi berbasis bukti
PENYUSUN	Asosiasi AI Indonesia, Riset 2026
BAB	Bab 4: Apa yang ditunjukkan eksperimen dan apa yang belum
HALAMAN	16–18 (nomor halaman laporan)
ISI BERKAS	Isi bab, Panduan Daftar Pustaka, dan Daftar Pustaka bab ini

PANDUAN PENGGUNAAN AI DALAM PROSES BELAJAR SISWA

AI Memberi Jawaban, Siswa Tetap Harus Berpikir Kritis

Eksperimen Türkiye [20] menunjukkan alasan untuk menguji kemampuan tanpa bantuan. Ketika siswa dapat meminta jawaban selama latihan, keberhasilan mengerjakan soal tidak selalu diikuti penguasaan mandiri. Versi tutor dengan petunjuk yang dirancang untuk menjaga proses belajar memitigasi kerugian tersebut. Ini mendukung pengujian guardrail pedagogis, bukan klaim bahwa seluruh AI merusak pembelajaran.

Interpretasi. Dalam implementasi Indonesia, pemberian petunjuk, permintaan alasan, dan pengurangan bantuan secara bertahap merupakan pilihan desain yang layak diuji. Masing-masing perlu dihubungkan dengan perilaku nyata: apakah siswa mencoba, menjelaskan kesalahan, serta menyelesaikan soal baru sendiri? Menambahkan kalimat "jangan langsung memberi jawaban" pada prompt saja belum membuktikan kontrol bekerja. Sebagai contoh desain, asisten LEA di Singapore Student Learning Space dirancang mengarahkan kembali murid yang keluar topik atau meminta jawaban langsung [25]. Sumber tersebut menjelaskan desain guardrail, bukan bukti bahwa kontrol itu meningkatkan hasil belajar.

Berhasil di satu tempat belum tentu berhasil di semua tempat

Studi Kestin [21] melibatkan 194 mahasiswa. Perbedaan tes segera mendukung potensi tutor yang disusun mengikuti pedagogi dan materi kuliah. Namun hasil ini tidak menguji penggantian dosen pada keseluruhan program studi, interaksi sosial kelas, atau kompetensi profesi jangka panjang.

Tutor CoPilot [22] memberi bantuan kepada tutor manusia. Versi EdWorkingPaper November 2025 yang diperiksa melaporkan manfaat pada penguasaan topik, sekaligus hasil nihil secara statistik pada tes matematika akhir tahun. Hasil nihil itu belum berarti tidak ada efek. Paparan siswa terhadap tutor yang memakai alat ini terbatas dan studi berlangsung sekitar dua bulan, sehingga desainnya tidak disiapkan untuk mengukur dampak pada tes akhir tahun. Menampilkan hanya hasil positifnya akan memberi gambaran yang tidak lengkap.

> **Implikasi.** Sekolah atau kampus dapat memilih satu outcome dekat untuk perbaikan harian dan satu outcome lebih jauh untuk keputusan perluasan. Sistem yang membantu hari ini tetapi belum mengubah hasil semester tetap mungkin bernilai, tetapi alasan pembelian dan klaim manfaatnya harus sesuai.

Penghematan kerja guru adalah manfaat tersendiri

Evaluasi EEF [23] melibatkan 259 guru pada 68 sekolah dan menilai persiapan sains Year 7–8. Selisih waktunya tidak boleh diubah menjadi klaim penghematan 31% dari seluruh jam kerja guru. Tidak ditemukan perbedaan kualitas bahan pada penilaian ahli yang dilakukan, tetapi studi ini tidak mengukur learning gain siswa.

Usulan. Guru dapat mencatat waktu membuat draf, memeriksa, merevisi, serta memperbaiki kesalahan setelah digunakan. Penghematan bersih menjadi ukuran yang lebih berguna daripada kecepatan menghasilkan teks. Waktu yang tersedia kembali kemudian dapat dialokasikan sesuai kebutuhan, misalnya umpan balik atau pendampingan.

Konteks terbatas sumber daya tetap membutuhkan rancangan dukungan

Program Nigeria [24] berlangsung enam minggu pada sembilan sekolah, dengan kegiatan sepulang sekolah dan bimbingan guru. Dari 1.328 siswa yang diacak, 759 mengikuti asesmen akhir (hitungan penulis dari 657 siswa kelompok perlakuan dan 671 siswa kelompok kontrol, serta 422 dan 337 siswa yang menyelesaikan asesmen akhir). Kehilangan tindak lanjut ini memperkuat kebutuhan kehati-hatian membaca hasil. Efek gabungan tidak dapat dipisahkan begitu saja menjadi efek perangkat lunak semata. Konversi hasil ke "tahun sekolah" juga bukan bukti bahwa siswa benar-benar melewati kurikulum beberapa tahun.

Implikasi untuk Indonesia. Perbandingan pilot sebaiknya menyamakan waktu belajar dan intensitas pendampingan. Jika kelompok AI mendapat jam tambahan sedangkan kelompok pembanding tidak, hasilnya menunjukkan manfaat paket program, bukan kontribusi AI yang terisolasi.

Bukti yang masih diperlukan

Untuk SD awal, penggunaan jangka panjang dan wellbeing, kajian ini tidak menemukan dasar yang cukup untuk menetapkan intensitas penggunaan universal. Untuk aksesibilitas, ukuran berhasil perlu disesuaikan: peningkatan akses pada materi dapat bernilai meskipun kecepatan bukan tujuan utama. Untuk learning analytics dan administrasi, akurasi prediksi atau kecepatan proses belum membuktikan keputusan akhirnya adil atau membantu siswa.

Sintesis OECD [1] mencakup tutoring, kolaborasi, kreativitas, dukungan guru, pengelolaan institusi dan penelitian. Cakupan itu menunjukkan banyak kemungkinan penggunaan; kekuatan bukti untuk masing-masing perlu diperiksa, dan tidak dapat diringkas sebagai satu angka manfaat AI pendidikan.

259
guru pada 68 sekolah (evaluasi EEF)

194
mahasiswa dalam studi Kestin

6 minggu
program Nigeria pada sembilan sekolah

Panduan Daftar Pustaka

Halaman-halaman sebelum ini adalah isi bab *Apa yang ditunjukkan eksperimen dan apa yang belum* (Bab 4) dari laporan *Kajian Komparatif AI dalam Pendidikan untuk Indonesia*, dengan satu halaman Identitas Dokumen yang disisipkan setelah halaman pembuka. Halaman-halaman setelah ini adalah Daftar Pustaka bab tersebut, memuat 7 sumber. Di laporan lengkap, daftar ini bernama Daftar Pustaka. Panduan ini menerangkan cara keduanya saling menunjuk.

Penanda di dalam teks

Sumber sebuah pernyataan ditandai dengan nomor dalam kurung siku, misalnya [12], atau beberapa nomor berurutan seperti [32][33]. Nomor itu adalah nomor catatan, bukan bagian dari kalimat dan bukan data.

Contoh bentuknya: “Kalimat yang dikutip dari sebuah sumber [12].” Nomor 12 menunjuk ke entri nomor 12 di Daftar Pustaka.

- Kalimat dengan nomor memuat pernyataan atau temuan yang diambil dari sumber lain. Sumbernya adalah entri bernomor sama di Daftar Pustaka.
- Kalimat tanpa nomor adalah analisis, pertimbangan, atau usulan penyusun laporan.
- Satu nomor menunjuk satu sumber di seluruh laporan, sehingga nomor di bab ini tidak harus berurutan dari 1.

Membaca Daftar Pustaka

Setiap entri di Daftar Pustaka dimulai dengan nomor catatan dalam kurung siku, diikuti nama penulis atau lembaga, judul, tahun terbit, dan alamat tautan bila ada. Untuk mengetahui asal sebuah pernyataan, cocokkan nomor di teks dengan nomor di awal entri.

Daftar Pustaka di berkas ini hanya memuat sumber yang dirujuk bab ini, diurutkan menurut nomor: 1, 20–25. Nomor yang tidak ditemukan di sana berarti sumbernya tidak dicantumkan di dokumen ini.

Menyebut tempat sebuah kutipan

Untuk menyebut asal sebuah pernyataan secara lengkap, ada dua keterangan yang saling melengkapi:

1. **Tempat di laporan ini:** judul laporan, nama bab, dan nomor halaman laporan.
2. **Sumber aslinya:** entri Daftar Pustaka dengan nomor catatan yang sama.

Nomor halaman laporan tercetak di bagian bawah setiap halaman isi. Nomor ini berbeda dari urutan halaman pada berkas PDF, karena berkas ini memuat halaman tambahan di luar isi. Identitas Dokumen, Panduan ini, dan Daftar Pustaka tidak bernomor halaman.

Daftar Pustaka

Bab 4: Apa yang ditunjukkan eksperimen dan apa yang belum. 7 sumber.

[1] OECD, OECD Digital Education Outlook 2026: Exploring Effective Uses of Generative AI in Education. Paris, Prancis: OECD Publishing, 2026, doi: 10.1787/062a7394-en. [Daring]. Tersedia: https://doi.org/10.1787/062a7394-en. Diakses: 30 Sep. 2026.

[20] H. Bastani, O. Bastani, A. Sungu, H. Ge, Ö. Kabakcı, dan R. Mariman, "Generative AI without guardrails can harm learning: Evidence from high school mathematics," Proc. Natl. Acad. Sci. U.S.A., vol. 122, no. 26, Art. no. e2422633122, 2025, doi: 10.1073/pnas.2422633122. [Daring]. Tersedia: https://doi.org/10.1073/pnas.2422633122. Diakses: 3 Okt. 2026.

[21] G. Kestin, K. Miller, A. Klales, T. Milbourne, dan G. Ponti, "AI tutoring outperforms in-class active learning: An RCT introducing a novel research-based design in an authentic educational setting," Sci. Rep., vol. 15, Art. no. 17458, 2025, doi: 10.1038/s41598-025-97652-6. [Daring]. Tersedia: http s://doi.org/10.1038/s41598-025-97652-6. Diakses: 30 Sep. 2026.

[22] R. E. Wang, A. T. Ribeiro, C. D. Robinson, S. Loeb, dan D. Demszky, "Tutor CoPilot: A human-AI approach for scaling real-time expertise," Annenberg Institute at Brown University, Providence, RI, AS, EdWorkingPaper No. 24- 1054, versi Nov. 2025, doi: 10.26300/81nh-8262. [Daring]. Tersedia: https://doi.org/10.26300/81nh-8262. Diakses: 3 Okt. 2026.

[23] P. Roy, H. Poet, R. Staunton, K. Aston, dan D. Thomas, "ChatGPT in lesson preparation: A teacher choices trial – Evaluation report," National Foundation for Educational Research untuk Education Endowment Foundation, London, Inggris, 2024. [Daring]. Tersedia: https://d2tic4wvo1iusb.clo udfront.net/production/documents/projects/chatgpt_in_less on_planning_-_evaluation_report.pdf. Diakses: 30 Sep. 2026.

[24] M. De Simone, F. Tiberti, M. Barron Rodriguez, F. Manolio, W. Mosuro, dan E. J. Dikoru, "From chalkboards to chatbots: Evaluating the impact of generative AI on learning outcomes in Nigeria," World Bank, Washington, DC, AS, Policy Research Working Paper 11125, 2025, doi: 10.1596/1813- 9450-11125. [Daring]. Tersedia: https://doi.org/10.1596/1813-9450-11125. Diakses: 3 Okt. 2026.

[25] Ministry of Education Singapore, "AI usage in schools," Parliamentary reply, 6 Mei 2026. [Daring]. Tersedia: https:// www.moe.gov.sg/news/parliamentary-replies/20260506-ai- usage-in-schools. Diakses: 3 Okt. 2026.