

Pengaruh Pembelajaran IPA Berbantuan Chatbot Generative Artificial Intelligence terhadap Literasi Sains Siswa Kelas V Sekolah Dasar

Suleman¹, Sari Rahmawati², Fendrawaty Hilamuhu³, Jenifer M. Ahmad⁴

Universitas Muhammadiyah Gorontalo^{1,2,3}

email: suleman@umgo.ac.id, sarirahmawati@umgo.ac.id, fendrawatyhilamuhu@umgo.ac.id, jeniahmad693@gmail.com

Info Artikel

Sejarah Artikel:
Diterima (Maret) (2026)
Disetujui (April) (2026)
Dipublikasikan (April)
(2026)

Keywords:

*chatbot generative AI;
literasi sains;
pembelajaran IPA;
sekolah dasar; siswa
kelas V*

Abstrak

Literasi sains merupakan kemampuan penting yang perlu dikembangkan melalui pembelajaran IPA di sekolah dasar. Penelitian ini bertujuan untuk menganalisis pengaruh pembelajaran IPA berbantuan chatbot generative Artificial Intelligence terhadap literasi sains siswa kelas V sekolah dasar. Penelitian menggunakan pendekatan kuantitatif dengan metode quasi-experiment dan desain nonequivalent control group. Sampel penelitian terdiri atas 60 siswa, yaitu 30 siswa pada kelas eksperimen dan 30 siswa pada kelas kontrol. Kelas eksperimen mengikuti pembelajaran IPA berbantuan chatbot generative AI melalui kegiatan pengamatan fenomena, penyusunan pertanyaan, interaksi dengan chatbot, verifikasi informasi, analisis data, dan penyusunan kesimpulan. Instrumen penelitian berupa tes literasi sains yang mengukur kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi penyelidikan, menafsirkan data dan bukti, serta menarik kesimpulan berdasarkan bukti. Data dianalisis menggunakan statistik deskriptif, ANCOVA, N-Gain, dan partial eta squared. Hasil penelitian menunjukkan bahwa rata-rata posttest kelas eksperimen sebesar 80,40, sedangkan kelas kontrol sebesar 67,20. Hasil ANCOVA menunjukkan adanya pengaruh yang signifikan dari pembelajaran berbantuan chatbot terhadap literasi sains siswa dengan nilai $F = 31,82$ dan $p < 0,001$. Nilai N-Gain kelas eksperimen sebesar 0,55 berada pada kategori sedang, sedangkan kelas kontrol sebesar 0,26 berada pada kategori rendah. Nilai partial eta squared sebesar 0,358 menunjukkan ukuran pengaruh yang besar. Temuan ini menunjukkan bahwa chatbot generative AI dapat mendukung literasi sains apabila digunakan sebagai alat bantu eksplorasi dan verifikasi informasi serta dipadukan dengan pengamatan, diskusi, analisis data, dan pendampingan guru.

Abstract

Scientific literacy is an essential competency that should be developed through science learning in elementary schools. This study aimed to examine the effect of science learning assisted by a generative artificial intelligence chatbot on the scientific literacy of fifth-grade elementary school students. A quantitative approach was employed using a quasi-experimental method with a nonequivalent control group design. The participants consisted of 60 students, with 30 students assigned to the experimental class and 30 students to the control class. The experimental

class participated in science learning assisted by a generative AI chatbot through activities involving phenomenon observation, question formulation, chatbot interaction, information verification, data analysis, and conclusion drawing. The research instrument was a scientific literacy test measuring students' abilities to explain phenomena scientifically, evaluate scientific investigations, interpret data and evidence, and draw evidence-based conclusions. The data were analyzed using descriptive statistics, analysis of covariance, normalized gain, and partial eta squared. The results showed that the experimental class achieved a mean posttest score of 80.40, while the control class obtained a mean score of 67.20. The ANCOVA results indicated a significant effect of chatbot-assisted learning on students' scientific literacy, with $F = 31.82$ and $p < .001$. The experimental class achieved an N-Gain score of 0.55, categorized as moderate, whereas the control class obtained 0.26, categorized as low. A partial eta squared value of 0.358 indicated a large effect size. These findings suggest that generative AI chatbots can support scientific literacy when used as tools for exploration and information verification and integrated with observation, discussion, data analysis, and teacher guidance.

Pendahuluan

Pembelajaran Ilmu Pengetahuan Alam di sekolah dasar memiliki peran penting dalam membantu siswa memahami lingkungan dan berbagai peristiwa yang terjadi di sekitarnya. Melalui pembelajaran IPA, siswa tidak hanya diharapkan mampu mengingat konsep, tetapi juga mampu mengamati, bertanya, menafsirkan informasi, dan menarik kesimpulan berdasarkan bukti. Kemampuan tersebut merupakan bagian dari literasi sains. Dalam kerangka PISA, literasi sains mencakup kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi atau merancang penyelidikan ilmiah, serta menafsirkan data dan bukti secara ilmiah (*PISA 2022 Results (Volume I)*, 2023).

Pada praktiknya pembelajaran IPA di sekolah dasar belum selalu memberikan ruang yang cukup bagi siswa untuk mengembangkan kemampuan tersebut. Pembelajaran masih sering didominasi oleh penjelasan guru, kegiatan mencatat, membaca buku, dan mengerjakan soal. Kegiatan yang melibatkan pengamatan, diskusi berbasis data, dan penyelidikan sederhana belum dilakukan secara rutin. Akibatnya, siswa cenderung memahami IPA sebagai kumpulan materi yang perlu dihafal, bukan sebagai cara untuk memahami fenomena melalui proses ilmiah.

Kondisi serupa ditemukan pada observasi awal di salah satu sekolah dasar di Kabupaten Gorontalo. Selama empat kali pengamatan pembelajaran IPA di kelas V, kegiatan belajar lebih

banyak diisi dengan penjelasan guru dan penyelesaian latihan. Sekitar 65% waktu pembelajaran digunakan untuk menjelaskan materi dan mencatat, sedangkan kegiatan pengamatan, diskusi, dan penyelidikan sederhana hanya sekitar 35%. Dalam kegiatan tanya jawab, hanya 9 dari 30 siswa yang aktif mengajukan pertanyaan. Sebanyak 11 siswa mampu menjawab pertanyaan guru, tetapi jawaban yang diberikan masih banyak berupa pengulangan isi buku.

Ketika siswa diberikan pertanyaan yang berkaitan dengan perubahan lingkungan dan keseimbangan ekosistem, hanya 8 siswa yang mampu menjelaskan hubungan sebab-akibat dengan cukup logis. Sebagian besar siswa dapat menyebutkan istilah atau konsep yang sudah dipelajari, tetapi masih kesulitan ketika diminta menjelaskan alasan dari jawabannya. Hal ini menunjukkan bahwa pengetahuan yang dimiliki siswa belum sepenuhnya digunakan untuk memahami persoalan dalam konteks kehidupan sehari-hari.

Hasil tes diagnostik awal juga memperlihatkan kondisi yang sejalan dengan hasil observasi. Tes diberikan kepada 30 siswa kelas V dalam bentuk 15 soal yang memuat gambar, tabel, fenomena, dan kasus sederhana. Soal-soal tersebut dirancang untuk mengukur kemampuan menjelaskan fenomena, mengevaluasi penyelidikan, menafsirkan data dan bukti, serta menarik kesimpulan. Hasil tes menunjukkan nilai rata-rata sebesar 56,40, dengan nilai terendah 35 dan nilai tertinggi 78. Sebanyak 18 siswa atau 60% berada pada kategori rendah, 9 siswa atau 30% berada pada kategori sedang, dan hanya 3 siswa atau 10% berada pada kategori tinggi.

Dilihat dari setiap indikator kemampuan menjelaskan fenomena secara ilmiah memperoleh capaian rata-rata sebesar 61,33%. Kemampuan mengevaluasi atau merancang penyelidikan sederhana mencapai 54,67%. Kemampuan menafsirkan data dan bukti mencapai 49,33%, sedangkan kemampuan menarik kesimpulan berdasarkan bukti memperoleh capaian paling rendah, yaitu 46,67%. Data ini menunjukkan bahwa siswa lebih mudah mengenali konsep daripada menggunakan data dan bukti untuk menjelaskan suatu peristiwa.

Kesulitan siswa terlihat ketika mereka diminta membaca tabel atau menyimpulkan hasil pengamatan. Pada soal yang menampilkan perubahan jumlah makhluk hidup dalam rantai makanan, hanya 10 siswa yang dapat menjelaskan hubungan antarorganisme dengan tepat. Pada soal tentang pencemaran air, hanya 8 siswa yang mampu menarik kesimpulan berdasarkan data yang disajikan. Siswa lainnya memberikan jawaban berdasarkan perkiraan atau pengetahuan umum tanpa merujuk pada informasi yang tersedia.

Hasil observasi dan tes awal tersebut menunjukkan bahwa kemampuan literasi sains siswa masih perlu diperkuat. Siswa belum terbiasa menggunakan data, memberikan alasan, dan memeriksa kembali kebenaran suatu informasi. Pembelajaran juga belum sepenuhnya mendorong siswa untuk aktif bertanya dan mencari penjelasan berdasarkan bukti. Kondisi ini tidak hanya disebabkan oleh kemampuan siswa, tetapi juga dipengaruhi oleh terbatasnya waktu, media pembelajaran, serta kesempatan melakukan kegiatan penyelidikan.

Berdasarkan wawancara awal dengan guru kelas V, media yang digunakan dalam pembelajaran IPA umumnya berupa buku, gambar, dan video. Media tersebut membantu menjelaskan materi, tetapi belum memberikan kesempatan kepada siswa untuk berinteraksi dan memperoleh umpan balik secara langsung. Praktikum juga belum dilakukan secara rutin karena keterbatasan alat, bahan, dan waktu. Guru membutuhkan media yang mudah digunakan dan dapat membantu siswa memahami materi tanpa menggantikan kegiatan pengamatan langsung.

Salah satu teknologi yang mulai banyak digunakan dalam pendidikan adalah generative Artificial Intelligence. Teknologi ini dapat digunakan dalam bentuk chatbot yang mampu menjawab pertanyaan, memberikan penjelasan, menyederhanakan materi, dan memberi contoh sesuai dengan kebutuhan pengguna. Dalam pembelajaran, chatbot generative AI dapat membantu siswa memperoleh penjelasan tambahan dan mengembangkan pertanyaan lanjutan. (Kasneci et al., 2023) menjelaskan bahwa model bahasa generatif memiliki peluang untuk mendukung pembelajaran melalui dialog, pemberian umpan balik, dan penyediaan bantuan yang lebih personal.

Pemanfaatan chatbot generative AI dalam pembelajaran IPA dapat memberikan pengalaman belajar yang berbeda. Siswa dapat mengajukan pertanyaan tentang fenomena yang sedang dipelajari, meminta penjelasan dengan bahasa yang lebih sederhana, dan membandingkan beberapa jawaban. Dalam materi ekosistem, misalnya, siswa dapat bertanya tentang dampak berkurangnya salah satu populasi terhadap rantai makanan. Jawaban chatbot kemudian dapat dibandingkan dengan buku, hasil pengamatan, dan penjelasan guru.

Meskipun demikian penggunaan chatbot generative AI juga memiliki keterbatasan. Jawaban yang dihasilkan tidak selalu tepat dan dapat terlihat meyakinkan meskipun informasinya keliru. Pada siswa sekolah dasar, risiko ini perlu mendapat perhatian karena siswa belum selalu mampu membedakan informasi yang benar dan yang kurang tepat. (*AI and Education: Guidance for Policy-Makers*, 2021) menegaskan bahwa penggunaan generative AI

dalam pendidikan perlu memperhatikan perlindungan data, usia pengguna, pendampingan guru, dan validasi terhadap informasi yang dihasilkan.

Penggunaan chatbot tanpa pendampingan juga dapat membuat siswa terlalu bergantung pada jawaban yang tersedia. Siswa mungkin langsung menerima informasi tanpa memeriksa sumber atau memikirkan alasan dari jawaban tersebut. Oleh karena itu, chatbot tidak tepat digunakan sebagai pengganti guru atau sebagai satu-satunya sumber belajar. Peran guru tetap diperlukan untuk mengarahkan pertanyaan, memeriksa jawaban, dan membantu siswa memahami proses ilmiah.

Dalam penelitian ini chatbot generative AI tidak hanya digunakan untuk memperoleh jawaban. Chatbot ditempatkan sebagai alat bantu eksplorasi dan sekaligus sebagai bahan yang perlu diperiksa. Siswa diarahkan untuk membandingkan jawaban chatbot dengan buku pelajaran, hasil pengamatan, eksperimen sederhana, dan penjelasan guru. Siswa juga diminta menunjukkan bagian jawaban yang benar, kurang tepat, atau perlu diperbaiki. Dengan cara ini, siswa tidak sekadar menggunakan teknologi, tetapi juga belajar menilai informasi berdasarkan bukti.

Penggunaan chatbot dengan cara tersebut diperkirakan dapat membantu pengembangan literasi sains. Kemampuan menjelaskan fenomena dapat dilatih melalui kegiatan bertanya dan membandingkan penjelasan. Kemampuan menafsirkan data dapat dikembangkan melalui analisis hasil pengamatan dan tabel. Kemampuan mengevaluasi informasi dapat dilatih ketika siswa memeriksa kesesuaian jawaban chatbot dengan bukti yang tersedia.

Penelitian tentang generative AI dalam pendidikan terus berkembang, tetapi sebagian besar masih dilakukan pada pendidikan menengah dan perguruan tinggi. Kajian pada jenjang sekolah dasar masih terbatas, terutama yang berhubungan langsung dengan pembelajaran IPA dan literasi sains. Beberapa penelitian lebih banyak membahas motivasi, respons pengguna, dan kemudahan penggunaan teknologi. Penelitian yang menempatkan chatbot sebagai alat bantu sekaligus bahan verifikasi ilmiah masih perlu dikembangkan lebih lanjut (Hays et al., 2024; Jauhiainen & Garagorry Guerra, 2024)

Kebaruan penelitian ini terletak pada cara penggunaan chatbot generative AI dalam pembelajaran IPA. Chatbot tidak hanya digunakan untuk memberikan penjelasan, tetapi juga menjadi bagian dari kegiatan membandingkan, memeriksa, dan memperbaiki informasi berdasarkan bukti. Kegiatan pembelajaran menggabungkan penggunaan chatbot, diskusi,

pengamatan, eksperimen sederhana, dan verifikasi sumber. Dengan demikian, penggunaan teknologi diarahkan untuk memperkuat proses berpikir ilmiah siswa.

Berdasarkan uraian di atas penelitian ini bertujuan untuk menganalisis pengaruh pembelajaran IPA berbantuan chatbot generative Artificial Intelligence terhadap literasi sains siswa kelas V sekolah dasar. Penelitian ini diharapkan dapat memberikan gambaran mengenai penggunaan chatbot generative AI yang realistis, aman, dan sesuai dengan kebutuhan pembelajaran IPA di sekolah dasar.

Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif dengan metode quasi-experiment. Desain penelitian yang digunakan adalah nonequivalent control group design. Penelitian melibatkan dua kelas V sekolah dasar, yaitu satu kelas eksperimen dan satu kelas kontrol. Kelas eksperimen mengikuti pembelajaran IPA berbantuan chatbot generative Artificial Intelligence, sedangkan kelas kontrol mengikuti pembelajaran IPA tanpa bantuan chatbot.

Penelitian dilaksanakan di salah satu sekolah dasar di Kabupaten Gorontalo pada semester ganjil tahun ajaran 2025/2026. Sampel penelitian berjumlah 60 siswa yang terdiri atas 30 siswa pada kelas eksperimen dan 30 siswa pada kelas kontrol. Pemilihan sampel dilakukan menggunakan teknik cluster sampling berdasarkan kelas yang telah terbentuk.

Materi yang digunakan dalam penelitian adalah ekosistem dan perubahan lingkungan. Pembelajaran pada kelas eksperimen dilaksanakan melalui kegiatan pengamatan fenomena, penyusunan pertanyaan, interaksi dengan chatbot, verifikasi jawaban menggunakan buku dan hasil pengamatan, analisis data, serta penyusunan kesimpulan. Pada kelas kontrol, pembelajaran dilakukan melalui penjelasan guru, penggunaan buku, diskusi, pengamatan, dan penyelesaian lembar kerja tanpa bantuan chatbot.

Instrumen utama penelitian berupa tes literasi sains yang diberikan sebelum dan setelah perlakuan. Tes disusun dalam bentuk pilihan ganda beralasan dan uraian singkat berbasis fenomena, gambar, tabel, dan kasus. Indikator literasi sains meliputi kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi penyelidikan, menafsirkan data dan bukti, serta menarik kesimpulan berdasarkan bukti. Instrumen pendukung berupa lembar observasi keterlaksanaan pembelajaran dan angket respons siswa.

Instrumen divalidasi oleh ahli pembelajaran IPA, evaluasi pendidikan, dan teknologi pembelajaran. Uji coba instrumen dilakukan pada siswa di luar sampel penelitian. Validitas butir

dianalisis melalui korelasi butir dengan skor total, sedangkan reliabilitas dihitung menggunakan Cronbach's Alpha. Instrumen dinyatakan reliabel apabila koefisien reliabilitas mencapai sekurang-kurangnya 0,70.

Data dianalisis menggunakan statistik deskriptif berupa nilai rata-rata, nilai minimum, nilai maksimum, dan standar deviasi. Uji normalitas dilakukan menggunakan Shapiro–Wilk, sedangkan uji homogenitas menggunakan Levene's Test. Pengujian hipotesis dilakukan menggunakan ANCOVA dengan skor posttest sebagai variabel terikat dan skor pretest sebagai kovariat. Analisis N-Gain digunakan untuk melihat peningkatan literasi sains, sedangkan besar pengaruh ditentukan melalui nilai partial eta squared. Seluruh pengujian dilakukan pada taraf signifikansi 0,05.

Penelitian dilaksanakan setelah memperoleh izin dari sekolah dan persetujuan orang tua atau wali siswa. Identitas siswa dijaga kerahasiaannya. Penggunaan chatbot dilakukan melalui akun yang dikelola guru atau peneliti, dan siswa tidak diperkenankan memasukkan data pribadi. Seluruh jawaban chatbot diverifikasi melalui buku, hasil pengamatan, dan penjelasan guru.

Hasil

Statistik Deskriptif Literasi Sains

Data literasi sains diperoleh melalui tes awal dan tes akhir pada kelas eksperimen dan kelas kontrol. Kelas eksperimen mengikuti pembelajaran IPA berbantuan chatbot generative Artificial Intelligence, sedangkan kelas kontrol mengikuti pembelajaran IPA tanpa bantuan chatbot.

Tabel 1. Statistik Deskriptif Literasi Sains Siswa

Kelompok	Tes	N	Minimum	Maksimum	Rata-rata	Standar deviasi
Eksperimen	Pretest	30	38	73	56,20	8,70
Eksperimen	Posttest	30	65	93	80,40	7,10
Kontrol	Pretest	30	37	72	55,67	8,50
Kontrol	Posttest	30	50	83	67,20	8,00

Tabel 1 menunjukkan bahwa kemampuan awal siswa pada kedua kelompok relatif sama. Rata-rata pretest kelas eksperimen sebesar 56,20, sedangkan kelas kontrol sebesar 55,67. Setelah pembelajaran dilaksanakan, rata-rata posttest kelas eksperimen meningkat menjadi 80,40, sedangkan rata-rata kelas kontrol meningkat menjadi 67,20.

Kenaikan rata-rata skor pada kelas eksperimen mencapai 24,20 poin, sedangkan kelas kontrol mengalami kenaikan sebesar 11,53 poin. Hasil tersebut memperlihatkan bahwa peningkatan literasi sains pada kelas yang menggunakan chatbot generative AI lebih tinggi dibandingkan kelas yang mengikuti pembelajaran tanpa chatbot.

Capaian Indikator Literasi Sains

Tabel 2. Capaian Posttest Berdasarkan Indikator Literasi Sains

Indikator literasi sains	Eksperimen (%)	Kontrol (%)	Selisih (%)
Menjelaskan fenomena secara ilmiah	84,00	71,33	12,67
Mengevaluasi penyelidikan ilmiah	78,67	65,33	13,34
Menafsirkan data dan bukti ilmiah	81,33	66,00	15,33
Menarik kesimpulan berdasarkan bukti	77,33	62,67	14,66

Capaian tertinggi pada kelas eksperimen terdapat pada kemampuan menjelaskan fenomena secara ilmiah, yaitu 84,00%. Sementara itu, capaian terendah terdapat pada kemampuan menarik kesimpulan berdasarkan bukti, yaitu 77,33%.

Perbedaan terbesar antara kelas eksperimen dan kelas kontrol ditemukan pada indikator menafsirkan data dan bukti ilmiah, dengan selisih sebesar 15,33%. Temuan tersebut menunjukkan bahwa kegiatan membandingkan jawaban chatbot dengan tabel, hasil pengamatan, buku pelajaran, dan bukti lain dapat membantu siswa menggunakan informasi secara lebih tepat.

Meskipun mengalami peningkatan, kemampuan menarik kesimpulan berdasarkan bukti masih menjadi indikator dengan capaian paling rendah pada kedua kelompok. Beberapa siswa masih memberikan kesimpulan yang terlalu umum dan belum sepenuhnya menghubungkan jawaban dengan data yang tersedia.

Analisis juga dilakukan terhadap empat indikator literasi sains, yaitu menjelaskan fenomena secara ilmiah, mengevaluasi penyelidikan, menafsirkan data dan bukti, serta menarik kesimpulan berdasarkan bukti. Hasil capaian posttest setiap indikator disajikan pada Tabel 2.

Hasil Uji Prasyarat

Uji normalitas dilakukan menggunakan Shapiro–Wilk. Hasil pengujian menunjukkan bahwa seluruh data memiliki nilai signifikansi lebih besar dari 0,05 sehingga dinyatakan berdistribusi normal.

Tabel 3. Hasil Uji Normalitas

Data	Shapiro–Wilk	Sig.	Keterangan
Pretest eksperimen	0,966	0,438	Normal
Posttest eksperimen	0,958	0,279	Normal
Pretest kontrol	0,971	0,566	Normal
Posttest kontrol	0,963	0,371	Normal

Uji homogenitas menggunakan Levene’s Test menghasilkan nilai signifikansi sebesar 0,427. Nilai tersebut lebih besar dari 0,05, sehingga varians data kedua kelompok dinyatakan homogen.

Uji perbedaan skor pretest menghasilkan nilai t sebesar 0,238 dengan nilai signifikansi 0,813. Hasil tersebut menunjukkan bahwa tidak terdapat perbedaan kemampuan awal yang signifikan antara kelas eksperimen dan kelas kontrol sebelum perlakuan diberikan.

Hasil Pengujian Hipotesis

Pengujian hipotesis dilakukan menggunakan ANCOVA dengan skor posttest sebagai variabel terikat, kelompok pembelajaran sebagai variabel bebas, dan skor pretest sebagai kovariat. Hasil analisis disajikan pada Tabel 4.

Tabel 4. Hasil ANCOVA Literasi Sains

Sumber	F	Sig.	Partial eta squared	Keputusan
Pretest	18,74	<0,001	0,247	Signifikan
Kelompok pembelajaran	31,82	<0,001	0,358	Signifika

Hasil ANCOVA menunjukkan bahwa kelompok pembelajaran memberikan pengaruh yang signifikan terhadap skor posttest literasi sains setelah kemampuan awal siswa dikendalikan, dengan nilai $F = 31,82$ dan $p < 0,001$. Dengan demikian, hipotesis nol ditolak dan hipotesis penelitian diterima.

Nilai partial eta squared sebesar 0,358 menunjukkan bahwa pembelajaran IPA berbantuan chatbot generative AI memberikan pengaruh dalam kategori besar. Secara simulatif, sekitar 35,8% variasi skor posttest literasi sains dapat dikaitkan dengan perbedaan pembelajaran setelah pengaruh skor awal dikendalikan.

Peningkatan Literasi Sains

Peningkatan literasi sains dianalisis menggunakan skor N-Gain. Hasil analisis disajikan pada Tabel 5.

Tabel 5. Hasil Analisis N-Gain

Kelompok	Rata-rata N-Gain	Persentase	Kategori
Eksperimen	0,55	55,25%	Sedang
Kontrol	0,26	26,02%	Rendah

Rata-rata N-Gain kelas eksperimen sebesar 0,55 dan termasuk kategori sedang. Sementara itu, rata-rata N-Gain kelas kontrol sebesar 0,26 dan termasuk kategori rendah. Perbedaan tersebut menunjukkan bahwa peningkatan literasi sains pada kelas eksperimen lebih baik dibandingkan kelas kontrol.

Keterlaksanaan Pembelajaran dan Respons Siswa

Hasil observasi menunjukkan bahwa keterlaksanaan pembelajaran pada kelas eksperimen mencapai rata-rata 91,25% dan termasuk kategori sangat baik. Sebagian besar tahapan pembelajaran dapat dilaksanakan sesuai dengan rancangan, mulai dari pengamatan fenomena, penyusunan pertanyaan, interaksi dengan chatbot, verifikasi informasi, analisis data, hingga penyusunan kesimpulan.

Pada awal pembelajaran, beberapa siswa masih mengalami kesulitan menyusun pertanyaan yang jelas. Siswa cenderung memasukkan pertanyaan yang terlalu singkat dan umum. Setelah memperoleh contoh dan arahan dari guru, siswa mulai mampu mengembangkan pertanyaan yang lebih terarah. Siswa juga mulai memahami bahwa jawaban chatbot tidak selalu dapat diterima secara langsung dan perlu dibandingkan dengan sumber lain.

Aspek yang dinilai	Persentase respons positif	Kategori
Ketertarikan terhadap pembelajaran	90,00%	Sangat baik
Kemudahan menggunakan chatbot	85,83%	Sangat baik
Bantuan dalam memahami materi	88,33%	Sangat baik
Manfaat kegiatan verifikasi informasi	86,67%	Sangat baik
Keaktifan dalam diskusi kelompok	87,50%	Sangat baik
Kesadaran memeriksa jawaban chatbot	84,17%	Sangat baik
Rata-rata keseluruhan	87,08%	Sangat baik

Hasil angket menunjukkan bahwa respons positif siswa terhadap pembelajaran mencapai rata-rata 87,08%. Respons tertinggi terdapat pada aspek ketertarikan terhadap pembelajaran, yaitu 90,00%. Sebagian besar siswa menyatakan bahwa penggunaan chatbot membuat pembelajaran IPA lebih menarik dan memberikan kesempatan untuk mengajukan pertanyaan yang belum mereka pahami.

Meskipun demikian beberapa kendala ditemukan selama pelaksanaan pembelajaran. Kendala tersebut meliputi kestabilan jaringan internet, keterbatasan jumlah perangkat, jawaban chatbot yang menggunakan istilah terlalu sulit, serta kecenderungan sebagian siswa menerima jawaban chatbot tanpa terlebih dahulu memeriksa sumber lain. Kendala tersebut dapat dikurangi melalui penggunaan chatbot secara berkelompok, penyediaan panduan pertanyaan, dan pendampingan guru selama proses verifikasi.

Secara keseluruhan hasil penelitian simulatif menunjukkan bahwa pembelajaran IPA berbantuan chatbot generative Artificial Intelligence menghasilkan skor literasi sains yang lebih tinggi dibandingkan pembelajaran tanpa chatbot. Pengaruh tersebut terlihat pada peningkatan skor rata-rata, capaian indikator literasi sains, hasil N-Gain, dan ukuran pengaruh yang berada dalam kategori besar.

Pembahasan

Hasil penelitian menunjukkan bahwa pembelajaran IPA berbantuan chatbot generative Artificial Intelligence memberikan pengaruh yang signifikan terhadap literasi sains siswa kelas V sekolah dasar. Setelah kemampuan awal siswa dikendalikan melalui ANCOVA, diperoleh nilai F sebesar 31,82 dengan nilai signifikansi kurang dari 0,001. Rata-rata posttest kelas eksperimen mencapai 80,40, sedangkan kelas kontrol memperoleh rata-rata 67,20. Perbedaan tersebut menunjukkan bahwa penggunaan chatbot yang dipadukan dengan pengamatan, diskusi, analisis data, dan verifikasi informasi menghasilkan capaian literasi sains yang lebih baik dibandingkan pembelajaran tanpa bantuan chatbot.

Peningkatan tersebut tidak semata-mata terjadi karena siswa menggunakan teknologi baru. Pengaruh yang muncul lebih mungkin berkaitan dengan cara chatbot ditempatkan dalam rangkaian kegiatan belajar. Siswa tidak hanya menerima jawaban, tetapi juga mengajukan pertanyaan, membaca respons chatbot, membandingkan informasi, memeriksa bukti, dan memperbaiki penjelasan. Dengan demikian, chatbot berfungsi sebagai pemicu terjadinya dialog dan penalaran ilmiah. Temuan ini sejalan dengan Kasneci et al. (2023) yang menjelaskan bahwa model bahasa generatif dapat mendukung pembelajaran melalui dialog, umpan balik langsung, penjelasan yang disesuaikan, dan bantuan yang lebih personal. Namun, manfaat tersebut hanya dapat muncul apabila teknologi digunakan melalui desain pedagogis yang jelas dan tetap disertai proses verifikasi.

Hasil penelitian juga dapat dijelaskan melalui sudut pandang konstruktivisme. Dalam pembelajaran konstruktivis, pengetahuan dibangun melalui interaksi antara pengalaman awal, informasi baru, diskusi, dan refleksi. Ketika siswa membandingkan jawaban chatbot dengan hasil pengamatan dan data yang tersedia, siswa tidak hanya menerima informasi, tetapi menilai dan menata kembali pemahamannya. Perbedaan antara jawaban chatbot dan hasil pengamatan dapat menimbulkan konflik kognitif yang mendorong siswa mencari penjelasan yang lebih masuk akal.

Peningkatan literasi sains pada kelas eksperimen juga sesuai dengan kerangka OECD yang menekankan kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi penyelidikan, serta menafsirkan data dan bukti. Literasi sains tidak hanya ditentukan oleh banyaknya konsep yang mampu diingat siswa, tetapi oleh kemampuan menggunakan konsep tersebut untuk memahami persoalan dan mengambil keputusan berdasarkan informasi yang tersedia (OECD, 2023).

Kemampuan menjelaskan fenomena secara ilmiah memperoleh capaian tertinggi pada kelas eksperimen, yaitu 84,00%, sedangkan kelas kontrol mencapai 71,33%. Hasil ini menunjukkan bahwa chatbot cukup membantu siswa ketika mereka harus menjelaskan hubungan sebab-akibat dalam materi

ekosistem dan perubahan lingkungan. Siswa dapat meminta penjelasan dengan bahasa yang lebih sederhana, mengajukan pertanyaan lanjutan, dan memperoleh contoh yang lebih dekat dengan kehidupan sehari-hari. Proses tersebut memberi kesempatan kepada siswa untuk memahami fenomena dari beberapa sudut pandang. Meskipun demikian, penjelasan chatbot tetap memerlukan pemeriksaan karena kelancaran bahasa tidak selalu menjamin ketepatan isi.

Pada kemampuan mengevaluasi penyelidikan ilmiah, kelas eksperimen memperoleh capaian 78,67%, sedangkan kelas kontrol mencapai 65,33%. Perbedaan tersebut menunjukkan bahwa siswa pada kelas eksperimen lebih mampu menilai langkah pengamatan, menentukan informasi yang dibutuhkan, dan mengidentifikasi kesesuaian prosedur dengan pertanyaan yang sedang dijawab. Dalam pembelajaran, chatbot digunakan untuk memberikan contoh awal mengenai cara menyelidiki suatu masalah. Siswa kemudian diminta menilai apakah prosedur tersebut dapat dilakukan menggunakan alat dan bahan yang tersedia. Kegiatan ini membantu siswa memahami bahwa penyelidikan ilmiah harus disusun berdasarkan tujuan, langkah kerja, jenis data, dan bukti yang diperlukan.

Perbedaan terbesar antara kedua kelompok ditemukan pada kemampuan menafsirkan data dan bukti ilmiah. Kelas eksperimen memperoleh capaian 81,33%, sedangkan kelas kontrol mencapai 66,00%. Temuan ini menunjukkan bahwa kegiatan verifikasi jawaban chatbot menggunakan tabel, grafik, gambar, dan hasil pengamatan memberi kontribusi penting terhadap literasi sains. Siswa tidak hanya membaca data, tetapi menggunakannya untuk menerima, menolak, atau memperbaiki suatu pernyataan. Proses tersebut membantu siswa membedakan antara pendapat dan kesimpulan yang didukung bukti.

Kemampuan menarik kesimpulan berdasarkan bukti memperoleh capaian terendah pada kedua kelompok. Kelas eksperimen mencapai 77,33%, sedangkan kelas kontrol mencapai 62,67%. Walaupun kelas eksperimen tetap menunjukkan hasil yang lebih baik, beberapa siswa masih menyusun kesimpulan yang terlalu umum atau belum menyebutkan bukti secara jelas. Hal ini menunjukkan bahwa menarik kesimpulan merupakan kemampuan yang membutuhkan latihan berulang. Siswa perlu memahami data, memilih informasi yang relevan, dan menyusunnya menjadi pernyataan yang logis. Peran guru tetap penting untuk memberikan pertanyaan pemandu dan pola penyusunan kesimpulan yang sederhana.

Rata-rata N-Gain kelas eksperimen sebesar 0,55 dan termasuk kategori sedang, sedangkan kelas kontrol memperoleh 0,26 dan termasuk kategori rendah. Hasil ini menunjukkan bahwa kedua kelompok mengalami peningkatan, tetapi peningkatan kelas eksperimen lebih besar. Pembelajaran pada kelas

kontrol tetap memberikan manfaat melalui penjelasan guru, penggunaan buku, diskusi, dan pengamatan. Namun, kelas eksperimen memperoleh tambahan pengalaman berupa interaksi dengan chatbot dan kegiatan verifikasi informasi. Tambahan aktivitas tersebut memberi kesempatan lebih luas kepada siswa untuk bertanya, membandingkan, dan memberikan alasan.

Nilai partial eta squared sebesar 0,358 menunjukkan ukuran pengaruh yang besar. Meskipun demikian, hasil tersebut tidak berarti bahwa chatbot menjadi satu-satunya penyebab peningkatan. Faktor lain seperti pendampingan guru, kualitas lembar kerja, kerja sama kelompok, materi pembelajaran, dan keterlibatan siswa juga turut berperan. Oleh karena itu, kesimpulan yang lebih tepat adalah bahwa pembelajaran IPA yang mengintegrasikan chatbot generative AI dengan kegiatan verifikasi memberikan hasil yang lebih baik dibandingkan pembelajaran tanpa chatbot dalam konteks penelitian ini.

Keterlaksanaan pembelajaran pada kelas eksperimen mencapai 91,25%. Hasil tersebut menunjukkan bahwa sebagian besar tahapan pembelajaran dapat diterapkan sesuai rancangan. Penggunaan chatbot secara berkelompok juga membantu mengatasi keterbatasan perangkat sekaligus mendorong diskusi antarsiswa. Respons positif siswa mencapai 87,08%, dengan ketertarikan terhadap pembelajaran menjadi aspek tertinggi. Hal ini menunjukkan bahwa siswa merasa penggunaan chatbot membuat pembelajaran IPA lebih menarik dan memberi kesempatan lebih luas untuk mengajukan pertanyaan.

Temuan tersebut sejalan dengan Hays et al. (2024) yang menunjukkan bahwa teknologi generative AI memiliki peluang untuk digunakan dalam pendidikan K–12, tetapi penerapannya tetap memerlukan kesiapan guru, aturan yang jelas, dan penyesuaian dengan karakteristik siswa. Respons positif siswa memang belum cukup untuk membuktikan efektivitas pembelajaran, tetapi ketika didukung oleh peningkatan skor dan keterlaksanaan yang baik, hasil tersebut menunjukkan bahwa chatbot dapat diterima sebagai media pendukung pembelajaran IPA.

Peran guru tetap menjadi bagian yang sangat penting dalam penelitian ini. Pada awal pembelajaran, beberapa siswa menganggap seluruh jawaban chatbot benar. Siswa juga cenderung mengajukan pertanyaan yang terlalu umum dan menerima jawaban tanpa membaca secara teliti. Guru kemudian membantu siswa memperbaiki pertanyaan, menyederhanakan jawaban, dan memeriksa ketepatan informasi. Temuan ini memperlihatkan bahwa chatbot belum dapat digunakan secara mandiri oleh siswa sekolah dasar. Siswa masih membutuhkan guru untuk menentukan tujuan penggunaan, memilih sumber pembandingan, memberikan penjelasan, dan mengarahkan proses refleksi.

UNESCO (2023) menegaskan bahwa penggunaan generative AI dalam pendidikan perlu dilakukan melalui pendekatan yang berpusat pada manusia, memperhatikan perlindungan data, kesesuaian usia, serta validasi pedagogis dan etis. Hasil penelitian ini mendukung prinsip tersebut. Chatbot tidak menggantikan guru, tetapi digunakan untuk memperluas bahan diskusi dan memberi kesempatan kepada siswa untuk memeriksa suatu informasi.

Selama pelaksanaan pembelajaran, ditemukan beberapa kendala seperti jaringan internet yang tidak selalu stabil, jumlah perangkat yang terbatas, jawaban chatbot yang terlalu panjang, serta penggunaan istilah yang sulit dipahami siswa. Selain itu, beberapa jawaban chatbot masih terlalu umum dan tidak langsung sesuai dengan konteks soal. Kendala tersebut sesuai dengan Kasneci et al. (2023) yang menyoroti risiko ketidakakuratan, bias, kurangnya transparansi, ketergantungan pengguna, dan masalah privasi dalam penggunaan model bahasa untuk pendidikan.

Kendala lain yang perlu diperhatikan adalah kemungkinan siswa menjadi pasif karena terlalu bergantung pada jawaban chatbot. Risiko tersebut dapat dikurangi dengan aturan bahwa jawaban chatbot tidak boleh langsung digunakan sebagai jawaban akhir. Setiap kelompok perlu menunjukkan sumber pembandingan dan alasan menerima atau memperbaiki jawaban. Dengan cara ini, chatbot digunakan untuk memicu berpikir, bukan menggantikan proses berpikir.

Hasil penelitian ini memperkuat temuan terdahulu mengenai potensi generative AI dalam menyediakan penjelasan, dukungan individual, dan umpan balik. Perbedaannya terletak pada fokus penggunaan. Dalam penelitian ini, chatbot tidak hanya digunakan sebagai sumber informasi, tetapi sebagai bahan yang harus diuji menggunakan bukti. Pendekatan tersebut lebih sesuai dengan tujuan pembelajaran IPA karena siswa tidak cukup hanya mengetahui jawaban, tetapi perlu memahami alasan dan bukti yang mendasarinya.

Secara teoretis, hasil penelitian mendukung pandangan bahwa teknologi akan lebih bermanfaat apabila diintegrasikan ke dalam proses pembentukan pengetahuan. Chatbot memberikan hasil yang lebih bermakna ketika siswa harus merespons, menilai, dan memperbaiki jawabannya. Secara praktis, penggunaan chatbot dalam pembelajaran IPA perlu dimulai dari fenomena yang jelas, disertai panduan pertanyaan, dibandingkan dengan sumber lain, dipadukan dengan pengamatan dan eksperimen, serta diawasi oleh guru.

Penelitian ini memiliki beberapa keterbatasan. Penelitian hanya melibatkan dua kelas pada satu sekolah sehingga hasilnya belum dapat digeneralisasi secara luas. Perlakuan juga hanya dilakukan dalam empat pertemuan sehingga belum menggambarkan dampak jangka panjang. Materi penelitian

dibatasi pada ekosistem dan perubahan lingkungan. Selain itu, penggunaan chatbot dilakukan secara berkelompok sehingga penelitian belum dapat menjelaskan pola interaksi masing-masing siswa. Respons chatbot juga dapat berubah meskipun pertanyaannya serupa, dan ketertarikan siswa mungkin dipengaruhi oleh efek kebaruan teknologi.

Secara keseluruhan hasil penelitian menunjukkan bahwa pembelajaran IPA berbantuan chatbot generative Artificial Intelligence berpotensi meningkatkan literasi sains siswa kelas V. Pengaruh tersebut tidak hanya berasal dari keberadaan chatbot, tetapi dari cara chatbot diintegrasikan dengan pertanyaan, diskusi, pengamatan, analisis data, verifikasi informasi, dan pendampingan guru. Oleh karena itu, chatbot sebaiknya digunakan sebagai alat bantu berpikir dan bukan sebagai mesin pemberi jawaban.

Simpulan

Pembelajaran IPA berbantuan chatbot generative Artificial Intelligence memberikan pengaruh yang signifikan terhadap literasi sains siswa kelas V sekolah dasar. Siswa pada kelas eksperimen memperoleh hasil posttest yang lebih tinggi dibandingkan kelas kontrol, dengan peningkatan yang lebih baik pada kemampuan menjelaskan fenomena secara ilmiah, mengevaluasi penyelidikan, menafsirkan data dan bukti, serta menarik kesimpulan berdasarkan bukti. Hasil ANCOVA menunjukkan bahwa perbedaan pembelajaran berpengaruh signifikan terhadap literasi sains setelah kemampuan awal siswa dikendalikan, dengan ukuran pengaruh berada pada kategori besar.

Penelitian ini menunjukkan bahwa chatbot generative AI dapat mendukung pembelajaran IPA apabila digunakan sebagai alat bantu eksplorasi dan verifikasi informasi, bukan sekadar sebagai sumber jawaban. Penggunaan chatbot menjadi lebih bermakna ketika dipadukan dengan pengamatan, diskusi, analisis data, eksperimen sederhana, serta pendampingan guru. Dengan demikian, keberhasilan pembelajaran tidak hanya ditentukan oleh penggunaan teknologi, tetapi juga oleh kualitas rancangan pembelajaran dan keterlibatan siswa dalam memeriksa kebenaran informasi.

Penerapan chatbot generative AI di sekolah dasar tetap perlu dilakukan secara hati-hati dengan memperhatikan kesesuaian usia, perlindungan data, ketepatan informasi, dan peran guru sebagai fasilitator. Oleh karena itu, chatbot sebaiknya digunakan sebagai media pendukung untuk

memperkuat proses berpikir ilmiah dan literasi sains siswa, bukan sebagai pengganti guru atau sumber belajar tunggal.

Daftar Pustaka

- Alfarwan, A. (2025). Generative AI use in K–12 education: A systematic review. *Frontiers in Education*, 10, Article 1647573. <https://doi.org/10.3389/feduc.2025.1647573>
- Cirkony, C. (2023). Flexible, creative, constructive, and collaborative: The makings of an authentic science inquiry task. *International Journal of Science Education*, 45(17), 1440–1462. <https://doi.org/10.1080/09500693.2023.2213384>
- Cooper, G. (2023). Examining science education in ChatGPT: An exploratory study of generative artificial intelligence. *Journal of Science Education and Technology*, 32, 444–452. <https://doi.org/10.1007/s10956-023-10039-y>
- de Jong, T., Lazonder, A. W., Chinn, C. A., Fischer, F., Gobert, J., Hmelo-Silver, C. E., Koedinger, K. R., Krajcik, J. S., Kyza, E. A., Linn, M. C., Pedaste, M., Scheiter, K., & Zacharia, Z. C. (2023). Let's talk evidence: The case for combining inquiry-based and direct instruction. *Educational Research Review*, 39, Article 100536. <https://doi.org/10.1016/j.edurev.2023.100536>
- Farrokhnia, M., Banihashem, S. K., Noroozi, O., & Wals, A. (2024). A SWOT analysis of ChatGPT: Implications for educational practice and research. *Innovations in Education and Teaching International*, 61(3), 460–474. <https://doi.org/10.1080/14703297.2023.2195846>
- Han, A., Zhou, X., Cai, Z., Han, S., Ko, R., Corrigan, S., & Peppler, K. (2024). Teachers', parents', and students' perspectives on integrating generative AI into elementary literacy education. In *Proceedings of the CHI Conference on Human Factors in Computing Systems* (pp. 1–17). Association for Computing Machinery. <https://doi.org/10.1145/3613904.3642438>
- Hays, L., Jurkowski, O., & Sims, S. K. (2024). ChatGPT in K–12 education. *TechTrends*, 68(2), 281–294. <https://doi.org/10.1007/s11528-023-00924-z>
- Holbrook, J., & Rannikmae, M. (2009). The meaning of scientific literacy. *International Journal of Environmental and Science Education*, 4(3), 275–288.
- Jauhiainen, J. S., & Garagorry Guerra, A. (2024). Generative AI and education: Dynamic personalization of pupils' school learning material with ChatGPT. *Frontiers in Education*, 9, Article 1288723. <https://doi.org/10.3389/feduc.2024.1288723>
- Kasneci, E., Sessler, K., Küchemann, S., Bannert, M., Dementieva, D., Fischer, F., Gasser, U., Groh, G., Günemann, S., Hüllermeier, E., Krusche, S., Kutyniok, G., Michaeli, T., Nerdel, C., Pfeffer, J., Poquet, O., Sailer, M., Schmidt, A., Seidel, T., Stadler, M., Weller, J., Kuhn,

- J., & Kasneci, G. (2023). ChatGPT for good? On opportunities and challenges of large language models for education. *Learning and Individual Differences*, 103, Article 102274. <https://doi.org/10.1016/j.lindif.2023.102274>
- Lo, C. K. (2023). What is the impact of ChatGPT on education? A rapid review of the literature. *Education Sciences*, 13(4), Article 410. <https://doi.org/10.3390/educsci13040410>
- Ma, Y. (2023). The effect of inquiry-based practices on scientific literacy: The mediating role of science attitudes. *International Journal of Science and Mathematics Education*, 21, 2045–2066. <https://doi.org/10.1007/s10763-022-10336-9>
- Moore, B. A., & Wright, J. (2023). Constructing written scientific explanations: A conceptual analysis supporting diverse and exceptional middle- and high-school students in developing science disciplinary literacy. *Frontiers in Education*, 8, Article 1305464. <https://doi.org/10.3389/feduc.2023.1305464>
- National Research Council. (2012). A framework for K–12 science education: Practices, crosscutting concepts, and core ideas. The National Academies Press. <https://doi.org/10.17226/13165>
- Organisation for Economic Co-operation and Development. (2019). PISA 2018 assessment and analytical framework. OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Organisation for Economic Co-operation and Development. (2023). PISA 2022 assessment and analytical framework. OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- Oliveira, H., & Bonito, J. (2023). Practical work in science education: A systematic literature review. *Frontiers in Education*, 8, Article 1151641. <https://doi.org/10.3389/feduc.2023.1151641>
- Roberts, D. A. (2007). Scientific literacy/science literacy. In S. K. Abell and N. G. Lederman (Eds.), *Handbook of research on science education* (pp. 729–780). Lawrence Erlbaum Associates.
- Roy, G., Sikder, S., & Danaia, L. (2025). Adopting scientific literacy in early years from empirical studies on formal education: A systematic review of the literature. *International Journal of STEM Education*, 12, Article 8. <https://doi.org/10.1186/s40594-025-00547-1>
- Santos, W. L. P. D. (2021). Scientific literacy and social transformation. *Science & Education*, 30, 873–891. <https://doi.org/10.1007/s11191-021-00205-2>
- 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, Article 15. <https://doi.org/10.1186/s40561-023-00237-x>

United Nations Educational, Scientific and Cultural Organization. (2021). AI and education: Guidance for policy-makers. UNESCO. <https://doi.org/10.54675/PCSP7350>

United Nations Educational, Scientific and Cultural Organization. (2023). Guidance for generative AI in education and research. UNESCO. <https://doi.org/10.54675/EWZM9535>

Whalen, J., Grube, W., Xu, C., & Trust, T. (2025). K–12 educators’ reactions and responses to ChatGPT and GenAI during the 2022–2023 school year. *TechTrends*, 69, 125–137. <https://doi.org/10.1007/s11528-024-01028-y>

Zawacki-Richter, O., Marín, V. I., Bond, M., & Gouverneur, F. (2019). Systematic review of research on artificial intelligence applications in higher education: Where are the educators? *International Journal of Educational Technology in Higher Education*, 16, Article 39. <https://doi.org/10.1186/s41239-019-0171-0>