

**PENGEMBANGAN VIRTUAL LAB AI-STEM UNTUK MENINGKATKAN
LITERASI SAINS DAN KREATIVITAS SISWA PADA PEMBELAJARAN
BIOLOGI DI DAERAH KABUPATEN PULAU TALIBU**

Ikhlaashul Ihsan¹, Eva Vianti², Mujais Apling³

^{1,2,3} Pendidikan IPA, Fakultas Inovasi Pendidikan, Universitas Nahdlatul Ulama
Maluku Utara

¹ikhlaashulhsan96@gmail.com, ²evavianti08@gmail.com, ³mujaisnavis@gmail.com

ABSTRACT

Scientific literacy and creativity are essential competencies in biology learning, yet previous evidence from Taliabu Island indicates that both remain at moderate levels and are strongly shaped by external learning conditions. This study developed an AI-STEM Virtual Lab adapted to the characteristics of an island region and examined its feasibility and effectiveness. The study combined product development, expert validation, and a quasi-experimental nonequivalent control-group design involving 62 students (31 experimental; 31 control). Scientific literacy and creativity were measured using 15-item four-point Likert instruments, while media feasibility was assessed by two experts using 48 indicators. The corrected expert-validation score was 88.75%, indicating very high feasibility with minor revisions. Reliability coefficients ranged from 0.805 to 0.906. Wilcoxon tests showed significant pretest-posttest improvement in all groups ($p < .001$). Scientific literacy in the experimental group increased from 50.54 to 69.03 ($N\text{-gain} = 0.378$), whereas the control group increased from 41.67 to 61.56 ($N\text{-gain} = 0.321$); however, the between-group $N\text{-gain}$ difference was not significant ($p = .176$) and baseline literacy differed significantly. Creativity showed a clearer intervention pattern: baseline scores were equivalent ($p = .848$), while experimental $N\text{-gain}$ (0.313) exceeded control $N\text{-gain}$ (0.201; $p = .049$). The findings indicate that the AI-STEM Virtual Lab is highly feasible and promising, particularly for strengthening student creativity, while claims regarding superior scientific-literacy gains require further multi-school testing with better baseline equivalence.

Keywords: AI-STEM, virtual laboratory, scientific literacy, creativity, biology learning, Taliabu Island

ABSTRAK

Literasi sains dan kreativitas merupakan kompetensi penting dalam pembelajaran Biologi, sedangkan bukti sebelumnya di Pulau Taliabu menunjukkan bahwa kedua kemampuan tersebut masih berada pada kategori sedang dan sangat dipengaruhi kondisi eksternal pembelajaran. Penelitian ini mengembangkan Virtual Lab AI-STEM yang disesuaikan dengan karakteristik wilayah kepulauan serta menguji kelayakan dan efektivitasnya. Penelitian menggabungkan pengembangan produk, validasi ahli, dan kuasi eksperimen nonequivalent control group yang melibatkan 62 siswa (31 eksperimen dan 31 kontrol). Literasi sains dan kreativitas diukur dengan instrumen 15 butir skala Likert empat poin, sedangkan kelayakan media dinilai oleh dua ahli melalui 48 indikator. Hasil validasi terkoreksi sebesar 88,75% menunjukkan media sangat layak dengan revisi kecil. Reliabilitas instrumen berada pada rentang 0,805-0,906. Uji Wilcoxon menunjukkan peningkatan pretest-posttest signifikan pada seluruh kelompok ($p < 0,001$). Literasi sains eksperimen meningkat dari 50,54 menjadi 69,03 ($N\text{-gain} = 0,378$), sedangkan kontrol dari 41,67 menjadi 61,56 ($N\text{-gain} = 0,321$); perbedaan $N\text{-gain}$ antarkelompok belum

signifikan ($p = 0,176$) dan kemampuan awal literasi berbeda signifikan. Pada kreativitas, kemampuan awal setara ($p = 0,848$) dan N-gain eksperimen ($0,313$) lebih tinggi daripada kontrol ($0,201$; $p = 0,049$). Dengan demikian, Virtual Lab AI-STEM sangat layak dan menjanjikan terutama untuk penguatan kreativitas, sedangkan keunggulan terhadap peningkatan literasi sains masih memerlukan pengujian lebih luas dengan kesetaraan awal yang lebih baik.

Kata Kunci: AI-STEM, virtual lab, literasi sains, kreativitas, pembelajaran Biologi, Pulau Taliabu

A. Pendahuluan

Pembelajaran Biologi pada abad ke-21 tidak cukup diarahkan pada penguasaan fakta dan konsep, tetapi perlu membentuk kemampuan siswa untuk menjelaskan fenomena secara ilmiah, mengevaluasi bukti, merancang penyelidikan, menginterpretasikan data, dan mengambil keputusan yang dapat dipertanggungjawabkan (Priyani & Nawawi, 2021). Kompetensi tersebut merupakan bagian penting dari literasi sains (Kang, 2022). Literasi sains menempatkan pengetahuan ilmiah sebagai alat untuk memahami permasalahan nyata, sehingga siswa tidak hanya mengetahui konsep, tetapi mampu menggunakan konsep tersebut ketika berhadapan dengan persoalan lingkungan, kesehatan, teknologi, dan kehidupan social (Kelp et al., 2023). Serta (Turiman et al., 2012) menegaskan bahwa literasi sains berkaitan erat dengan proses ilmiah, penalaran, interpretasi

informasi, dan keterampilan abad ke-21.

Kondisi pendidikan sains di Indonesia masih menunjukkan kesenjangan antara tuntutan kompetensi tersebut dan capaian siswa (Amrianto et al., 2025). Artikel terdahulu yang secara khusus mengkaji siswa SMA di Pulau Taliabu melaporkan bahwa literasi sains berada pada kategori sedang dengan rerata 58,57, sedangkan kreativitas juga berada pada kategori sedang dengan rerata 60,77. Penelitian yang melibatkan 104 siswa dari lima SMA tersebut juga menemukan bahwa hubungan kemandirian belajar, kreativitas, dan literasi sains tidak signifikan serta hanya menjelaskan sebagian sangat kecil variasi literasi sains. Temuan tersebut mengarahkan perhatian pada faktor eksternal seperti kualitas proses pembelajaran, kesempatan melakukan praktikum, fasilitas laboratorium, dan akses terhadap sumber belajar (Ihsan et al., 2026).

Konteks	Pulau	Taliabu	
membuat persoalan tersebut semakin penting. (Ihsan et al., 2026)mencatat bahwa Kabupaten Pulau Taliabu terdiri atas delapan kecamatan dan 71 desa. Sebanyak 70 desa telah memiliki sinyal telepon seluler, tetapi 41 desa masih memiliki kekuatan sinyal yang lemah dan satu desa belum memiliki sinyal. Pada saat yang sama, karakter wilayah didominasi oleh aktivitas pertanian, kehutanan, dan perikanan serta memiliki berbagai fenomena lingkungan yang dapat dijadikan konteks autentik pembelajaran sains. Kondisi ini menunjukkan dua kebutuhan yang harus dipenuhi secara bersamaan, yaitu pembelajaran yang kaya konteks lokal dan media digital yang tetap ringan serta adaptif terhadap keterbatasan perangkat dan konektivitas.			dan bahan fisik. (Navarro et al., 2024) melaporkan bahwa virtual laboratory Biologi dinilai mudah digunakan dan mendukung motivasi serta kemandirian, tetapi tetap memerlukan dukungan teknis dan lebih tepat digunakan sebagai pelengkap pengalaman laboratorium nyata. (Yang et al., 2024)juga menunjukkan efek positif moderat teknologi virtual terhadap keterampilan praktis.
Keterbatasan laboratorium fisik dapat mengurangi frekuensi pengalaman eksperimen siswa. Virtual lab memungkinkan siswa memanipulasi variabel, mengulang prosedur, mengamati hubungan sebab-akibat, dan membaca data tanpa selalu bergantung pada alat			Berdasarkan kebutuhan tersebut, dikembangkan Virtual Lab AI-STEM untuk pembelajaran Biologi di Kabupaten Pulau Taliabu. Produk dirancang untuk menggabungkan simulasi eksperimen, manipulasi variabel, pengolahan data, integrasi elemen STEM, dukungan AI, dan konteks lokal. Salah satu bentuk kontekstualisasi yang dinilai oleh ahli adalah pengaitan materi fotosintesis dengan tumbuhan lokal, ekosistem hutan, pesisir, dan mangrove Pulau Taliabu serta pemanfaatannya dalam kehidupan masyarakat. Desain juga mempertimbangkan variasi perangkat Android, ukuran layar, dan koneksi internet yang terbatas.
			Kebaruan penelitian ini terletak pada upaya mengubah temuan diagnostik lokal menjadi intervensi

pembelajaran yang spesifik terhadap wilayah kepulauan. Penelitian Ihsan et al. (2026) sebelumnya menjelaskan bahwa kreativitas dan literasi sains siswa Taliabu masih berada pada kategori sedang dan bahwa faktor internal yang diuji belum mampu menjelaskan variasi literasi sains secara memadai (Ihsan et al., 2026). Penelitian ini bergerak pada tahap berikutnya dengan mengembangkan media AI-STEM yang dapat memfasilitasi pengalaman eksperimen dan menguji perubahan literasi sains serta kreativitas melalui kelompok eksperimen dan kontrol.

Penelitian bertujuan untuk: (1) menentukan tingkat kelayakan Virtual Lab AI-STEM berdasarkan validasi ahli media; (2) menganalisis perubahan literasi sains dan kreativitas siswa sebelum dan setelah pembelajaran; dan (3) membandingkan peningkatan literasi sains dan kreativitas antara kelompok eksperimen dan kelompok kontrol. Hasil penelitian diharapkan menjadi dasar empiris bagi pengembangan pembelajaran Biologi berbasis digital yang kontekstual, adaptif terhadap wilayah kepulauan, dan tetap berhati-hati dalam menyimpulkan efektivitas

berdasarkan karakteristik data awal siswa.

B. Metode Penelitian

Penelitian menggunakan pendekatan penelitian pengembangan yang dilanjutkan dengan uji efektivitas kuasi eksperimen. Tahap pengembangan mencakup perumusan kebutuhan pembelajaran, perancangan fitur Virtual Lab AI-STEM, pengembangan prototipe, validasi ahli, revisi produk, implementasi pada kelompok eksperimen, dan evaluasi hasil belajar. Penelitian tidak menetapkan nama model pengembangan tertentu karena dokumen penelitian yang dianalisis merekam tahapan pengembangan dan validasi produk, tetapi tidak mencantumkan secara eksplisit model R&D tertentu. Uji efektivitas menggunakan desain nonequivalent pretest-posttest control group dengan dua kelompok kelas yang tersedia tanpa randomisasi individu.

Subjek uji efektivitas berjumlah 62 siswa SMA yang berasal dari SMAN 1 Pulau Taliabu. Sebanyak 31 siswa berada pada kelompok eksperimen yang menggunakan Virtual Lab AI-STEM dan 31 siswa

berada pada kelompok kontrol. Masing-masing siswa mengisi instrumen pada tahap pretest dan posttest. Penggunaan kelompok yang telah tersedia membuat penelitian ini lebih sesuai dikategorikan sebagai kuasi eksperimen daripada eksperimen acak penuh. Konsekuensinya, kesetaraan kemampuan awal perlu diperiksa sebelum menarik kesimpulan tentang perbedaan hasil akhir.

Instrumen literasi sains terdiri atas 15 pernyataan skala Likert empat poin yang mencakup pemahaman konsep, penjelasan fenomena, pencarian informasi ilmiah, interpretasi tabel/grafik, penarikan kesimpulan, hubungan sains-lingkungan, pemecahan masalah, evaluasi informasi ilmiah, penerapan sains dalam kehidupan, serta kebiasaan bertanya. Skor dikonversi menjadi nilai 0-100.

Instrumen kreativitas terdiri atas 15 pernyataan skala Likert empat poin yang merepresentasikan kelancaran ide, fleksibilitas penyelesaian, keberanian mencoba pendekatan berbeda, pengembangan dan modifikasi gagasan, eksplorasi sumber belajar, inovasi sederhana,

serta diskusi untuk memperoleh ide baru. Nilai dikonversi ke skala 0-100 dengan kategori sangat rendah, rendah, sedang, tinggi, dan sangat tinggi.

Kelayakan media dinilai oleh dua ahli menggunakan 48 butir skala 1-5 pada sepuluh aspek: tujuan dan karakteristik media, antarmuka, navigasi, kualitas teknis, simulasi, integrasi AI, integrasi STEM, dukungan literasi sains, dukungan kreativitas, serta kontekstualitas-aksesibilitas-etika-keselamatan.

Persentase dihitung dari skor perolehan terhadap skor maksimum; 81-100% dikategorikan sangat layak dan 61-80% layak.

Sebelum analisis inferensial, konsistensi internal instrumen dihitung menggunakan Cronbach's Alpha dari respons seluruh siswa pada masing-masing waktu pengukuran. Distribusi data diperiksa dengan Shapiro-Wilk. Karena beberapa distribusi pretest, posttest, dan gain tidak memenuhi asumsi normalitas, perubahan berpasangan dianalisis dengan Wilcoxon Signed-Rank Test dan perbandingan antarkelompok dianalisis dengan

Mann-Whitney U. Taraf signifikansi ditetapkan pada $\alpha = 0,05$.

Peningkatan pembelajaran juga dihitung menggunakan normalized gain (N-gain) dengan rumus $g = (\text{posttest} - \text{pretest}) / (100 - \text{pretest})$. N-gain $< 0,30$ ditafsirkan rendah, $0,30-0,69$ sedang, dan $\geq 0,70$ tinggi. Selain perbandingan N-gain, perbedaan kemampuan awal diperiksa untuk menghindari interpretasi berlebihan terhadap posttest. Karena literasi sains menunjukkan perbedaan baseline yang signifikan, dilakukan analisis sensitivitas ANCOVA dengan pretest sebagai kovariat dan standard error robust HC3. Analisis sensitivitas ini tidak menggantikan analisis nonparametrik utama, tetapi digunakan untuk menilai apakah keunggulan posttest tetap terlihat setelah mempertimbangkan kemampuan awal.

Aiken's V digunakan sebagai diagnostik item untuk mengidentifikasi prioritas revisi. Karena hanya terdapat dua validator, nilai V tidak dipakai sebagai satu-satunya bukti validitas isi. Persentase kelayakan juga diverifikasi ulang dari

skor mentah agar rekap bebas dari kesalahan aritmetika.

C. Hasil Penelitian dan Pembahasan

1. Kelayakan Virtual Lab AI-STEM

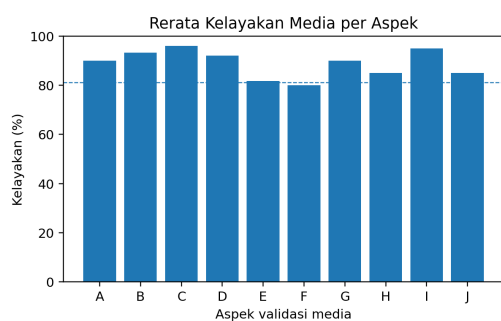
Hasil penilaian dua ahli menunjukkan bahwa Virtual Lab AI-STEM memperoleh skor 213 dari skor maksimum 240 pada masing-masing validator. Setelah persentase dihitung kembali langsung dari skor mentah, kedua penilaian menghasilkan 88,75%. Nilai tersebut termasuk kategori sangat valid/sangat layak. Kedua ahli tetap merekomendasikan media digunakan dengan revisi kecil, terutama pada kejelasan identitas versi produk, petunjuk penggunaan, batasan fungsi AI, optimasi perangkat Android berspesifikasi rendah, dan efisiensi akses pada koneksi terbatas.

Tabel 1. Rekapitulasi Kelayakan Virtual Lab AI-STEM (%)

Aspek	Ahli 1	Ahli 2	Rerata
A. Tujuan & karakteristik	90,00	90,00	90,00
B. Antarmuka visual	93,33	93,33	93,33
C. Navigasi	96,00	96,00	96,00
D. Teknis & kompatibilitas	92,00	92,00	92,00
E. Simulasi & interaktivitas	83,33	80,00	81,67
F. Integrasi AI & umpan balik	80,00	80,00	80,00
G. Integrasi STEM	88,00	92,00	90,00
H. Literasi sains	85,00	85,00	85,00
I. Kreativitas	95,00	95,00	95,00
J. Konteks, akses, etika	85,00	85,00	85,00
Total tertimbang	88,75	88,75	88,75

Aspek dengan nilai tertinggi adalah navigasi dan kemudahan penggunaan (96,00%) serta

dukungan terhadap kreativitas (95,00%). Aspek terendah adalah integrasi AI dan kualitas umpan balik (80,00%), diikuti simulasi dan interaktivitas virtual lab (81,67%). Pola ini konsisten dengan komentar validator bahwa AI perlu diberi batas fungsi yang lebih jelas dan simulasi perlu semakin menyerupai prosedur praktikum nyata. Rerata Aiken's V seluruh item sebesar 0,859. Item 22, 29, dan 48 memperoleh $V = 0,500$, sedangkan item 23 memperoleh $V = 0,625$. Item-item tersebut berkaitan dengan realisme fungsi alat/bahan virtual, manipulasi variabel, adaptivitas AI, serta perlindungan data dan etika AI. Karena hanya melibatkan dua ahli, nilai tersebut digunakan terutama sebagai dasar prioritas revisi.



Grafik 1. Rerata Persentase Kelayakan Media pada Sepuluh Aspek

Kelayakan yang tinggi mengindikasikan bahwa media telah memenuhi fungsi dasar sebagai lingkungan pembelajaran digital.

Temuan tersebut sejalan dengan Navarro et al. (2024) yang menunjukkan bahwa virtual laboratory pada Biologi dapat diterima pengguna ketika antarmuka, visual, dan alurnya mudah dipahami(Navarro et al., 2024). Namun, penilaian ahli dalam penelitian ini juga memperlihatkan bahwa kualitas media tidak cukup diukur dari tampilan dan navigasi. Interaktivitas eksperimen, adaptivitas AI, privasi data, dan etika harus diperlakukan sebagai komponen pedagogis utama. (Almasri, 2024)menekankan bahwa integrasi AI dalam pendidikan sains perlu diarahkan untuk mendukung proses belajar dan umpan balik, bukan sekadar menambahkan fitur teknologi.

2. Reliabilitas Instrumen

Tabel 2. Reliabilitas Instrumen

Variabel	Pengukuran	Alpha	Interpretasi
Literasi sains	Pretest gabungan	0,906	Sangat baik
Literasi sains	Posttest gabungan	0,846	Baik
Kreativitas	Pretest gabungan	0,805	Baik
Kreativitas	Posttest gabungan	0,856	Baik

Cronbach's Alpha seluruh instrumen berada di atas 0,80. Hasil tersebut menunjukkan konsistensi internal yang baik sampai sangat baik sehingga respons siswa cukup

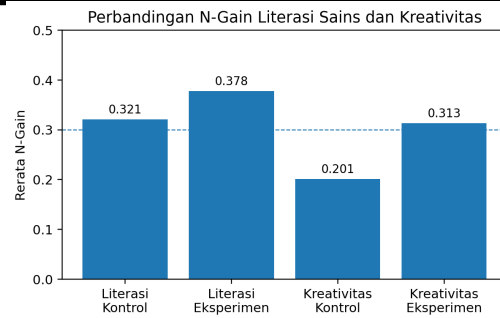
konsisten untuk digunakan dalam analisis perubahan pretest dan posttest. Hasil ini juga sejalan dengan penelitian lokal sebelumnya yang melaporkan reliabilitas konstruk kreativitas dan literasi sains yang baik dalam pengukuran siswa SMA di Pulau Taliabu (Ihsan et al., 2026)

3. Statistik Deskriptif dan N-Gain

Tabel 3. Pretest, Posttest, Gain, dan N-Gain

Variabel	Kelompok	Pretest	Posttest	Gain	N-Gain
Literasi	Kontrol	41,67 ± 11,43	61,56 ± 10,99	19,89	0,321
	Eksperimen	50,54 ± 12,59	69,03 ± 10,33	18,49	0,378
Kreativitas	Kontrol	50,48 ± 10,10	61,40 ± 9,45	10,91	0,201
	Eksperimen	51,29 ± 9,73	66,29 ± 12,02	15,00	0,313

Secara deskriptif, seluruh kelompok mengalami peningkatan. Pada literasi sains, rerata kelompok eksperimen meningkat 18,49 poin dan kelompok kontrol meningkat 19,89 poin. N-gain eksperimen (0,378) dan kontrol (0,321) sama-sama berada pada kategori sedang. Pada kreativitas, kelompok eksperimen meningkat 15,00 poin dengan N-gain 0,313 (sedang), sedangkan kontrol meningkat 10,91 poin dengan N-gain 0,201 (rendah). Dengan demikian, perbedaan pola peningkatan paling jelas secara deskriptif terlihat pada kreativitas.



Grafik 2. Perbandingan Rerata N-Gain Kelompok Kontrol dan Eksperimen

4. Uji Perubahan dan Perbandingan Antarkelompok

Tabel 4. Ringkasan Uji Inferensial

Variabel/Kel.	Uji	p-value	Keputusan
Lit. kontrol	Wilcoxon pre-post	<0,001	Signifikan
Lit. eksperimen	Wilcoxon pre-post	<0,001	Signifikan
Kreat. kontrol	Wilcoxon pre-post	<0,001	Signifikan
Kreat. eksperimen	Wilcoxon pre-post	<0,001	Signifikan
Literasi	MWU pretest	0,012	Berbeda
Literasi	MWU N-gain	0,176	Tidak berbeda
Kreativitas	MWU pretest	0,848	Setara
Kreativitas	MWU posttest	0,049	Berbeda
Kreativitas	MWU N-gain	0,049	Berbeda

Uji Wilcoxon menunjukkan seluruh kelompok mengalami peningkatan signifikan dari pretest ke posttest. Pada literasi sains, kelompok eksperimen dan kontrol sama-sama menunjukkan $p < 0,001$. Temuan ini membuktikan bahwa terjadi perubahan pembelajaran pada kedua kelompok, tetapi peningkatan dalam kelompok eksperimen saja belum dapat digunakan untuk menyatakan bahwa Virtual Lab AI-STEM lebih efektif daripada pembelajaran kontrol. Perbandingan antarkelompok merupakan tahap penting untuk menjawab klaim efektivitas relatif.

Pada literasi sains, kemampuan awal kelompok eksperimen telah lebih tinggi daripada kontrol (Mann-Whitney U, $p = 0,012$). Posttest eksperimen juga lebih tinggi ($p = 0,004$), tetapi perbedaan gain mentah tidak signifikan ($p = 0,949$) dan perbedaan N-gain tidak signifikan ($p = 0,176$). Analisis sensitivitas ANCOVA dengan pretest sebagai kovariat menghasilkan estimasi keunggulan posttest eksperimen sekitar 3,91 poin; dengan robust standard error HC3 pengaruh kelompok belum signifikan pada alpha 0,05 ($p = 0,105$). Oleh karena itu, data mendukung adanya peningkatan literasi sains pada kelompok eksperimen, tetapi belum mendukung klaim bahwa peningkatan tersebut secara statistik lebih besar daripada kelompok kontrol.

Interpretasi tersebut penting karena desain nonequivalent group rentan terhadap perbedaan awal. Jika hanya membandingkan posttest, keunggulan kelompok eksperimen dapat terlihat besar, padahal sebagian keunggulan sudah hadir sebelum perlakuan. Pendekatan yang hati-hati meningkatkan kredibilitas kesimpulan dan mencegah efek

media dibesar-besarkan. Secara substantif, N-gain literasi eksperimen tetap lebih tinggi daripada kontrol, tetapi selisih tersebut perlu diuji ulang pada sampel yang lebih besar, beberapa sekolah, dan prosedur pembentukan kelompok yang menghasilkan baseline lebih setara.

Pada kreativitas, kondisi awal berbeda. Pretest kelompok eksperimen dan kontrol tidak berbeda signifikan ($p = 0,848$), sehingga kedua kelompok relatif setara sebelum intervensi. Setelah pembelajaran, posttest eksperimen lebih tinggi berdasarkan Mann-Whitney U ($p = 0,049$), dan N-gain eksperimen juga lebih tinggi daripada kontrol ($p = 0,049$). Hedges' g untuk perbandingan N-gain sekitar 0,546, yang menunjukkan efek sedang. Analisis sensitivitas ANCOVA menunjukkan estimasi keunggulan posttest eksperimen sekitar 4,40 poin, meskipun dengan robust standard error hasilnya berada sedikit di atas batas signifikansi konvensional ($p = 0,069$). Secara keseluruhan, bukti terhadap kreativitas lebih konsisten daripada bukti terhadap literasi sains.

Hasil kreativitas memiliki dasar pedagogis yang kuat. Virtual lab memberi kesempatan siswa mencoba berbagai kombinasi variabel, mengulang eksperimen tanpa konsekuensi kerusakan alat, membandingkan alternatif, dan memperbaiki keputusan berdasarkan respons sistem. Ruang eksplorasi semacam ini relevan dengan indikator fluency, flexibility, elaboration, dan revisi gagasan. (Serna-Mendiburu & Guerra-Tamez, 2024) menemukan bahwa teknologi virtual memberikan efek positif moderat terhadap kreativitas dan bahwa dampaknya meningkat ketika kegiatan virtual dipadukan dengan inquiry-discovery, task-driven learning, serta umpan balik yang jelas. Hal ini sejalan dengan desain AI-STEM yang tidak hanya menyajikan informasi tetapi menuntut siswa berinteraksi dengan masalah dan respons sistem.

5. Pergeseran Kategori

Kemampuan Siswa

Tabel 5. Distribusi Kategori Pretest dan Posttest

Variabel/Kelompok	Pretest R/S/T/ST	Posttest R/S/T/ST
Literasi kontrol	20/11/0/0	2/19/7/3
Literasi eksperimen	10/16/5/0	0/5/19/7
Kreativitas kontrol	11/18/2/0	0/19/10/2
Kreativitas eksperimen	9/20/2/0	0/8/17/6

Keterangan R/S/T/ST berturut-turut menunjukkan kategori rendah, sedang, tinggi, dan sangat tinggi. Pada literasi sains eksperimen, sebelum perlakuan terdapat 10 siswa kategori rendah, 16 sedang, dan 5 tinggi. Setelah perlakuan, tidak terdapat lagi siswa kategori rendah; 19 siswa berada pada kategori tinggi dan 7 siswa pada kategori sangat tinggi. Pada kreativitas eksperimen, sembilan siswa yang semula berada pada kategori rendah tidak lagi muncul pada posttest; 17 siswa berada pada kategori tinggi dan enam pada kategori sangat tinggi. Pergeseran kategori ini memperlihatkan bahwa perubahan tidak hanya terjadi pada rerata, tetapi juga pada distribusi tingkat kemampuan siswa.

Perubahan distribusi perlu dibaca bersama hasil inferensial. Untuk kreativitas, pergeseran kategori yang lebih kuat pada kelompok eksperimen konsisten dengan perbedaan N-gain yang signifikan. Untuk literasi sains, perubahan kategori eksperimen terlihat sangat baik, tetapi kelompok eksperimen juga memulai penelitian dari rata-rata yang lebih tinggi. Hal ini

kembali menegaskan perlunya membedakan antara peningkatan internal kelompok dan bukti keunggulan kausal dibanding kelompok kontrol.

6. Pembahasan Integratif: AI-STEM, Literasi Sains, dan Konteks Kepulauan

Secara keseluruhan, hasil penelitian menunjukkan bahwa pengembangan Virtual Lab AI-STEM relevan dengan kebutuhan pembelajaran Biologi di Pulau Taliabu. Penelitian lokal sebelumnya menemukan bahwa kreativitas dan literasi sains siswa berada pada kategori sedang serta model yang hanya memasukkan faktor internal memiliki kemampuan prediksi yang sangat rendah (Ihsan et al., 2026). Temuan tersebut mengindikasikan pentingnya lingkungan belajar dan akses terhadap pengalaman ilmiah. Virtual lab dapat diposisikan sebagai salah satu intervensi eksternal yang memperluas kesempatan melakukan eksperimen tanpa menunggu ketersediaan laboratorium fisik lengkap.

Konteks geografis Taliabu juga menjelaskan mengapa aspek

kompatibilitas, efisiensi jaringan, dan aksesibilitas tidak dapat diperlakukan sebagai persoalan teknis sekunder. Media yang sangat interaktif tetapi membutuhkan koneksi stabil dan perangkat tinggi justru dapat memperlebar kesenjangan. Data wilayah yang dikutip Ihsan et al. (2026) menunjukkan banyak desa masih memiliki sinyal lemah. Oleh karena itu, optimasi ukuran aset, cache, pemuatan bertahap, dan kemungkinan mode luring sebagian merupakan bagian dari validitas ekologis media. Masukan validator agar aplikasi tetap ringan pada perangkat Android berspesifikasi rendah perlu menjadi prioritas pengembangan versi berikutnya.

Dukungan AI perlu tetap tunduk pada tujuan pedagogis. Nilai validasi aspek AI sebesar 80,00% merupakan nilai terendah dibanding aspek lain dan menunjukkan bahwa fitur AI masih menjadi area paling membutuhkan penyempurnaan. (Almasri, 2024) menjelaskan bahwa AI dapat mendukung penilaian, kuis, umpan balik, dan lingkungan belajar, tetapi implementasinya perlu memperhatikan peran guru, transparansi, dan tantangan etis.

Dalam konteks siswa SMA, AI sebaiknya memberikan pertanyaan penuntun dan umpan balik yang memancing alasan, bukan memberikan jawaban final sebelum siswa memproses data eksperimen.

Temuan kreativitas memperkuat asumsi bahwa pengalaman virtual paling bermanfaat ketika siswa diberi ruang tindakan. (Navarro et al., 2024) melaporkan virtual lab Biologi dapat meningkatkan motivasi, otonomi, minat, dan kepercayaan diri. (Serna-Mendiburu & Guerra-Tamez, 2024) menemukan pengaruh moderat teknologi virtual terhadap kreativitas, sedangkan (Yang et al., 2024) menemukan efek positif moderat terhadap keterampilan praktis. Hasil penelitian ini sejalan dengan kecenderungan tersebut, karena kreativitas kelompok eksperimen menunjukkan N-gain yang secara statistik lebih tinggi daripada kontrol. Namun, virtual lab tidak perlu diposisikan sebagai pengganti absolut laboratorium fisik. Penggunaan paling rasional adalah sebagai sarana persiapan, pengayaan, pengulangan eksperimen, dan alternatif ketika alat,

bahan, keselamatan, atau jarak membatasi praktikum langsung.

Pada literasi sains, hasil yang lebih kompleks tidak berarti media gagal. Kelompok eksperimen mengalami peningkatan yang besar secara internal dan berpindah menuju kategori tinggi dan sangat tinggi. Akan tetapi, peningkatan kontrol juga besar dan kelompok eksperimen sudah memiliki baseline lebih tinggi. Literasi sains merupakan konstruk luas yang mencakup pengetahuan, penalaran, evaluasi penyelidikan, interpretasi bukti, dan pengambilan keputusan. Dampak intervensi terhadap konstruk luas tersebut mungkin memerlukan durasi pembelajaran lebih lama, beberapa topik Biologi, serta asesmen berbasis performa dan skenario autentik, tidak hanya kuesioner persepsi.

Dari sisi pengembangan, hasil validasi dan efektivitas memberikan arah revisi yang jelas. Pertama, identitas produk perlu memuat versi dan tahun pengembangan. Kedua, pengguna pertama membutuhkan tutorial singkat yang menjelaskan alur eksperimen dan fungsi AI. Ketiga, representasi alat dan bahan virtual perlu dibuat lebih menyerupai

penggunaan praktikum nyata. Keempat, AI perlu meningkatkan adaptivitas tanpa mengambil alih proses berpikir. Kelima, privasi, etika AI, dan integritas akademik perlu ditampilkan sebagai bagian eksplisit dalam antarmuka. Keenam, aplikasi perlu terus dioptimalkan untuk perangkat dan jaringan yang sesuai dengan kondisi wilayah sasaran.

D. Keterbatasan Penelitian

Penelitian ini memiliki beberapa keterbatasan yang perlu diperhatikan. Pertama, uji efektivitas dilakukan pada satu sekolah dengan 62 siswa sehingga generalisasi ke seluruh Kabupaten Pulau Taliabu harus dilakukan secara hati-hati. Karakteristik akses internet, budaya belajar, fasilitas sekolah, dan kompetensi guru dapat berbeda antarsekolah maupun antarkecamatan.

Kedua, penggunaan kelompok kelas yang telah tersedia menyebabkan desain tidak sepenuhnya acak. Hal ini terlihat dari perbedaan pretest literasi sains antara kelompok eksperimen dan kontrol. Meskipun analisis N-gain dan ANCOVA sensitivitas telah digunakan, penelitian berikutnya

sebaiknya menerapkan matching, randomisasi pada tingkat kelas bila memungkinkan, atau desain multi-level pada beberapa sekolah.

Ketiga, instrumen literasi sains dan kreativitas berupa angket skala Likert sehingga lebih banyak menangkap persepsi atau penilaian diri siswa terhadap kompetensinya. Penelitian berikutnya perlu menambahkan asesmen kinerja berupa soal berbasis fenomena, analisis data, desain eksperimen, produk kreatif, dan log aktivitas virtual lab agar perubahan kompetensi dapat diamati secara langsung.

Keempat, validasi media melibatkan dua ahli. Hasil Aiken's V bermanfaat untuk mengidentifikasi butir yang perlu direvisi, tetapi jumlah validator yang kecil membatasi stabilitas estimasi validitas isi. Validasi lanjutan disarankan melibatkan ahli Biologi, ahli teknologi pembelajaran, ahli AI/etika digital, guru praktisi, dan pengguna siswa. Kelima, penelitian belum menguji retensi jangka panjang sehingga keberlanjutan dampak setelah pembelajaran berakhir belum diketahui.

D. Kesimpulan

Virtual Lab AI-STEM yang dikembangkan untuk pembelajaran Biologi di Kabupaten Pulau Taliabu memperoleh tingkat kelayakan terkoreksi sebesar 88,75% dan termasuk kategori sangat valid/sangat layak. Produk dinilai kuat pada aspek navigasi, antarmuka, dukungan kreativitas, teknis, dan integrasi STEM, sedangkan aspek integrasi AI, realisme simulasi, privasi, serta etika penggunaan AI menjadi prioritas revisi. Kedua validator merekomendasikan media layak digunakan dengan revisi kecil.

Pada uji efektivitas, seluruh kelompok mengalami peningkatan pretest-posttest yang signifikan. Literasi sains eksperimen meningkat dari 50,54 menjadi 69,03 dengan N-gain 0,378, sedangkan kontrol meningkat dari 41,67 menjadi 61,56 dengan N-gain 0,321. Karena pretest literasi kedua kelompok berbeda dan perbedaan N-gain tidak signifikan, penelitian ini tidak menyimpulkan bahwa peningkatan literasi sains eksperimen secara statistik lebih unggul dari kontrol. Sebaliknya, kreativitas memiliki baseline yang setara; N-gain eksperimen 0,313

lebih tinggi daripada kontrol 0,201 dan perbedaannya signifikan pada $p = 0,049$. Dengan demikian, bukti efektivitas paling kuat dalam dataset ini terdapat pada penguatan kreativitas.

Virtual Lab AI-STEM memiliki potensi sebagai media pembelajaran kontekstual untuk wilayah kepulauan, khususnya ketika dirancang ringan, adaptif terhadap keterbatasan konektivitas, memfasilitasi eksperimen berulang, dan memberikan dukungan AI yang transparan serta tidak menggantikan penalaran siswa. Penelitian lanjutan perlu melibatkan lebih banyak sekolah, durasi intervensi lebih panjang, kelompok yang lebih setara, asesmen performa, serta evaluasi retensi untuk memperoleh bukti yang lebih kuat mengenai dampak terhadap literasi sains.

DAFTAR PUSTAKA

- Almasri, F. (2024). Exploring the Impact of Artificial Intelligence in Teaching and Learning of Science: A Systematic Review of Empirical Research. *Research in Science Education*, 54(5), 977–997.
<https://doi.org/10.1007/s11165-024-10176-3>
- Amrianto, A., Rohman, F., Dharmawan, A., & Sari, M. S.

-
- (2025). PROBLEMS AND CHALLENGES IN 21ST CENTURY BIOLOGY LEARNING: A QUALITATIVE STUDY. *Research and Development Journal of Education*, 11(2), 927. <https://doi.org/10.30998/rdje.v11i2.26534>
- Ihsan, I., Vianti, E., Apling, M., Ayu Lestari, F., & Tuara, Z. I. (2026). Academy of Social Science and Global Citizenship Journal Analisis Hubungan Kemandirian Belajar, Kreativitas, dan Literasi Sains Siswa SMA di Pulau Taliabu Menggunakan Structural Equation Modeling-PLS. *AoSSaGCJ*, 6(1), 60–67. <https://doi.org/10.47200/aossagcj.v6i1.3746>
- Kang, J. (2022). Interrelationship Between Inquiry-Based Learning and Instructional Quality in Predicting Science Literacy. *Research in Science Education*, 52(1), 339–355. <https://doi.org/10.1007/s11165-020-09946-6>
- Kelp, N. C., McCartney, M., Sarvary, M. A., Shaffer, J. F., & Wolyniak, M. J. (2023). Developing Science Literacy in Students and Society: Theory, Research, and Practice. *Journal of Microbiology & Biology Education*, 24(2). <https://doi.org/10.1128/jmbe.00058-23>
- Navarro, C., Arias-Calderón, M., Henríquez, C. A., & Riquelme, P. (2024). Assessment of Student and Teacher Perceptions on the Use of Virtual Simulation in Cell Biology Laboratory Education. *Education Sciences*, 14(3). <https://doi.org/10.3390/educsci14030243>
- Priyani, N. E., & Nawawi, N. (2021). Analisis Pembelajaran STEM di Daerah Terluar Tertinggal Terdepan Indonesia Selama Masa Pandemi. *PSEJ (Pancasakti Science Education Journal)*, 6(1), 30–37. <https://doi.org/10.24905/psej.v6i1.30>
- Serna-Mendiburu, G. M., & Guerra-Tamez, C. R. (2024). Shaping the future of creative education: the transformative power of VR in art and design learning. *Frontiers in Education*, 9. <https://doi.org/10.3389/feduc.2024.1388483>
- Turiman, P., Omar, J., Daud, A. M., & Osman, K. (2012). Fostering the 21st Century Skills through Scientific Literacy and Science Process Skills. *Procedia - Social and Behavioral Sciences*, 59, 110–116. <https://doi.org/10.1016/j.sbspro.2012.09.253>
- Yang, C., Zhang, J., Hu, Y., Yang, X., Chen, M., Shan, M., & Li, L. (2024). The impact of virtual reality on practical skills for students in science and engineering education: a meta-analysis. In *International Journal of STEM Education* (Vol. 11, Number 1). Springer Science and Business Media Deutschland GmbH. <https://doi.org/10.1186/s40594-024-00487-2>
-