Pengembangan Game Berbasis Algoritma Logika pada Roblox untuk Melatih Problem Solving dengan Model ADDIE

Development of a Logical Algorithm-Based Game on Roblox for Training Problem Solving Using the ADDIE Model

Adel Aulia Dwi Murti[1*], Agung Kharisma Hidayah[2], Sastya Hendri Wibowo[3], Muhammad Imanullah[4]
Program Studi Teknik Informatika, Universitas Muhammadiyah Bengkulu
Jl. Bali, Kota Bengkulu, Indonesia
[1]auliaadeldm@gmail.com
[2]kharisma@umb.ac.id
[3]sastiahendriwibowo@gmail.com
[4]muhammad.imanullah@umb.ac.id

[*]Corresponding Author: auliaadeldm@gmail.com

Riwayat Artikel
Naskah diterima: 05/09/2026 | Naskah direvisi: 29/10/2026 | Naskah diterbitkan: 31/10/2026

Creative Commons License 4.0 (CC BY –NC)

Abstract*— Problem-solving ability is an essential skill for university students, yet interactive three-dimensional media for practising it remain limited. This study develops Nalar.exe, an educational puzzle game built with Roblox Studio and the Lua programming language, to train problem solving through three logical algorithm concepts: sequential logic, pattern recognition, and conditional logic. The game was developed using the Research and Development approach with the ADDIE model and consists of 10 levels arranged in three phases. It was evaluated through black-box testing, expert validation, the System Usability Scale (SUS), and a pre-test–post-test design. All 10 levels functioned as designed (100%). The logical algorithm expert and the game design expert gave feasibility scores of 76.67% and 75.00%, respectively. The average SUS score from 15 respondents was 72.17, above the benchmark of 68 and categorised as good. In the pre-test–post-test with 10 students, five participants already obtained the maximum pre-test score, a ceiling effect that prevented measuring their gain. Recalculated following Hake, the class-average N-Gain is 0.33 and the mean individual N-Gain of the five computable respondents is 0.40, both in the medium category. Because the estimate rests on few respondents with room to improve, it is only an initial indication and cannot be taken as proof of effectiveness. A more challenging instrument, a larger sample, and a control group are needed in further research.

Keywords*— Educational Game, Logical Algorithm, ADDIE Model, Roblox Studio, Problem Solving

Intisari— Kemampuan pemecahan masalah merupakan keterampilan penting bagi mahasiswa, tetapi media latihan interaktif berbasis tiga dimensi masih terbatas. Penelitian ini mengembangkan Nalar.exe, permainan edukatif berbasis Roblox Studio dan bahasa pemrograman Lua, untuk melatih pemecahan masalah melalui tiga konsep algoritma logika, yaitu logika berurutan, pengenalan pola, dan logika bersyarat. Pengembangan menggunakan pendekatan Research and Development dengan model ADDIE dan menghasilkan 10 level dalam tiga fase. Evaluasi dilakukan melalui pengujian black-box, validasi ahli, kuesioner System Usability Scale (SUS), serta pre-test dan post-test. Seluruh 10 level (100%) berfungsi sesuai rancangan. Ahli algoritma logika dan ahli game design memberikan persentase kelayakan masing-masing 76,67% dan 75,00%. Rata-rata skor SUS dari 15 responden adalah 72,17, di atas acuan 68 dan tergolong baik. Pada pre-test dan post-test terhadap 10 mahasiswa, lima responden telah memperoleh skor maksimal pada pre-test (efek langit-langit) sehingga peningkatannya tidak dapat diukur. Setelah dihitung ulang menurut Hake, N-Gain rata-rata kelas sebesar 0,33 dan rata-rata N-Gain individu lima responden yang dapat dihitung sebesar 0,40, keduanya berkategori sedang. Karena estimasi ini bertumpu pada sedikit responden yang masih memiliki ruang peningkatan, hasil ini baru berupa indikasi awal dan belum dapat dijadikan bukti efektivitas. Diperlukan instrumen yang lebih menantang, sampel yang lebih besar, dan kelompok kontrol pada penelitian lanjutan.

Kata Kunci— Game Edukasi, Algoritma Logika, Model ADDIE, Roblox Studio, Pemecahan Masalah

I. PENDAHULUAN

Kemampuan pemecahan masalah (*problem solving*) merupakan salah satu keterampilan abad ke-21 yang perlu dikuasai generasi muda, bersama kemampuan berpikir kritis, komunikasi, dan kolaborasi [1], [2]. Namun, pembelajaran masih sering berfokus pada penguasaan teori dan hafalan, sehingga mahasiswa kurang memperoleh ruang untuk berlatih menyelesaikan masalah secara aktif [1], [2]. Permainan edukatif digital menjadi salah satu kandidat media latihan mandiri karena menghadirkan tantangan berjenjang, umpan balik visual, dan ruang eksplorasi [3], [4]. Sebuah meta-analisis melaporkan bahwa permainan digital lebih efektif daripada pengajaran konvensional dalam pembelajaran STEM, dan jenis permainan teka-teki memberi manfaat yang menonjol bagi keterampilan berpikir dan pemecahan masalah [5].

Roblox Studio menarik sebagai platform pengembangan karena memungkinkan pembuatan ruang interaktif 3D yang dapat diakses lintas perangkat [6], dan telah ditinjau sebagai lingkungan belajar yang bermanfaat, termasuk bagi pembelajaran pemrograman pada bidang STEM [7]. Bahasa skrip yang digunakannya, Lua, bersifat ringan dan telah dimanfaatkan untuk membangun permainan maupun simulasi edukasi [8], [9].

TABEL I
PERBANDINGAN DENGAN PENELITIAN DAN PERMAINAN TERDAHULU

Peneliti	Media	Metode	Keterbatasan
Weng dkk. [10]	Aplikasi web teka-teki logika	Eksperimen, 121 peserta	Tanpa pengalaman 3D
Wardani dkk. [11]	Permainan papan fisik	ADDIE	Tidak digital
Azhar dkk. [12]	Game Roblox etnomatematika	ADDIE	Berfokus pada materi etnomatematika, bukan algoritma logika
Romadhan dkk. [13]	Modul pembuatan *obby* Roblox	PjBL	Tanpa evaluasi kelayakan produk yang menyeluruh
Tes IQ (Cranky Games)	Game Roblox, logika umum	Observasi	Soal acak, tidak berjenjang
Squid Game X (redifi)	Game Roblox, *if-else* implisit	Observasi	Satu konsep, berorientasi hiburan, tanpa evaluasi
Temukan dan Tandai (Blue Capybara Studios)	Game Roblox, pengenalan pola dan deduksi	Observasi	Kompetitif PvP, bukan latihan individual
Escape Room (seriif)	Game Roblox, logika berurutan	Observasi	Alur linier dan sekali pakai
Penelitian ini	Game 3D Roblox, 10 level	ADDIE	Tiga konsep logika dalam satu rangkaian; dievaluasi dengan validasi ahli, SUS, dan tes awal-akhir

Penelitian terdahulu telah menunjukkan manfaat permainan untuk melatih pemecahan masalah, tetapi dengan keterbatasan masing-masing. Weng dkk. [10] melaporkan bahwa permainan teka-teki logika meningkatkan kemampuan pemecahan masalah 121 peserta, namun produknya berupa aplikasi web sederhana tanpa pengalaman tiga dimensi. Wardani dkk. [11] menggunakan model ADDIE dan pendekatan berpikir komputasional melalui permainan papan fisik, sehingga belum bersifat digital. Pada platform Roblox, Azhar dkk. [12] mengembangkan permainan dengan model ADDIE untuk materi etnomatematika guna melatih berpikir kritis, sedangkan Romadhan dkk. [13] menyusun modul pembuatan permainan *obby* dengan pendekatan *project-based learning* tanpa evaluasi kelayakan produk yang menyeluruh. Observasi penulis terhadap empat permainan logika di Roblox (Tabel 1) juga menunjukkan bahwa permainan tersebut cenderung mengangkat satu konsep, bersifat linier atau kompetitif, dan tidak memiliki sistem evaluasi pemecahan masalah.

Perbandingan penelitian dan permainan tersebut dengan penelitian ini diringkas pada Tabel 1.

Sepengetahuan penulis, belum ada permainan logika di Roblox Studio yang memadukan logika berurutan, pengenalan pola, dan logika bersyarat dalam satu rangkaian level berjenjang, sekaligus dievaluasi melalui validasi ahli, uji *usability*, serta pengukuran *pre-test* dan *post-test*. Penelitian ini bertujuan (1) mengembangkan permainan Nalar.exe dengan model ADDIE; (2) menilai kelayakannya melalui pengujian fungsional dan validasi ahli algoritma logika serta ahli *game design*; (3) menilai *usability*-nya dengan *System Usability Scale* (SUS); dan (4) memperoleh indikasi awal mengenai peningkatan kemampuan pemecahan masalah. Kontribusi artikel ini adalah prototipe permainan beserta bukti awal kelayakan dan kebergunaannya, serta gambaran jujur mengenai keterbatasan pengukuran efektivitasnya.

Artikel ini disusun sebagai berikut. Bagian II memaparkan landasan teori, Bagian III menjelaskan metodologi dan prosedur pengujian, Bagian IV menyajikan hasil dan analisisnya, dan Bagian V memuat kesimpulan beserta saran.

II. LANDASAN TEORI

A. Algoritma Logika dan Berpikir Komputasional

Algoritma adalah rangkaian langkah yang jelas dan terurut untuk menyelesaikan suatu masalah, dan menjadi inti berpikir komputasional [14]. Berpikir komputasional mencakup dekomposisi, pengenalan pola, abstraksi, dan perancangan algoritma [15]. Tiga konsep digunakan sebagai dasar mekanik permainan. Logika berurutan (*sequential logic*) menuntut instruksi dijalankan secara beruntun sehingga urutan langkah menentukan hasil. Pengenalan pola (*pattern recognition*) adalah kemampuan mengidentifikasi keteraturan agar solusi ditemukan lebih efisien [16]. Logika bersyarat (*conditional logic*) menjalankan tindakan berdasarkan suatu kondisi, umumnya dengan struktur *if-else*.

B. Pemecahan Masalah

Polya [17] membagi pemecahan masalah menjadi empat langkah: memahami masalah, menyusun rencana,

JUTEI Edisi Volume. 10 No.2 Oktober 2026
ISSN 2579-3675, e-ISSN-5538
DOI 10.21460/jutei.2026.102.538

melaksanakan rencana, dan memeriksa hasil. Tahapan serupa digunakan pada kajian terkini mengenai mahasiswa [16]. Tantangan pada Nalar.exe dirancang mengikuti urutan itu: pemain menangkap kondisi masalah, memilih langkah, menerapkannya, lalu menilai hasilnya dari umpan balik permainan.

C. Roblox Studio dan Bahasa Lua

Roblox Studio adalah lingkungan pengembangan untuk membuat pengalaman 3D pada platform Roblox. Objek, antarmuka, dan perilaku permainan disusun di dalam editor dan dikendalikan melalui skrip Lua yang menangani peristiwa seperti sentuhan, penekanan tombol, dan perpindahan pemain [8]. Platform ini telah dikaji sebagai ruang belajar interaktif pada pendidikan tinggi dan pembelajaran pemrograman [6], [7], sehingga relevan untuk permainan latihan logika.

D. Game Edukatif dan Pengukuran Peningkatan

Game edukatif memadukan tujuan pembelajaran dengan unsur permainan seperti tantangan berjenjang, umpan balik langsung, dan kesempatan mencoba ulang, sehingga pemain belajar melalui tindakan [3], [4]. Manfaat tersebut baru dapat dinyatakan apabila diukur dengan instrumen yang mampu mendeteksi perubahan. Instrumen yang terlalu mudah menimbulkan *ceiling effect*, yaitu banyak peserta memperoleh skor tertinggi sebelum perlakuan sehingga peningkatan tidak dapat teramati. Dalam kondisi itu, selisih skor kecil dan N-Gain menjadi tidak stabil atau tidak terdefinisi [18].

E. Model ADDIE

ADDIE terdiri atas lima tahap, yaitu *Analysis*, *Design*, *Development*, *Implementation*, dan *Evaluation* [19]. Model ini telah digunakan pada pengembangan permainan edukatif [20], [11].

F. Pengukuran Berpikir Komputasional

Kemampuan berpikir komputasional dan pemecahan masalah dapat diukur melalui tes berbasis skenario yang menuntut peserta menerapkan dekomposisi, pengenalan pola, abstraksi, dan algoritma pada situasi konkret [15]. Agar tes mampu membedakan kemampuan, butir perlu memiliki tingkat kesukaran yang beragam dan opsi pengecoh yang masuk akal. Tes yang seluruh butirnya mudah dijawab hanya memberi informasi bahwa peserta menguasai konsep dasar, tetapi tidak mampu menunjukkan perubahan setelah suatu perlakuan.

G. Persentase Kelayakan Validasi Ahli

Kelayakan produk dinyatakan sebagai persentase skor angket Likert terhadap skor maksimal:

$$P = \frac{S_{\text{peroleh}}}{S_{\text{maks}}} \times 100\% \qquad (1)$$

dengan S_{peroleh} jumlah skor yang diperoleh dan S_{maks} jumlah butir dikalikan skor tertinggi (5).

H. System Usability Scale

SUS dikembangkan oleh Brooke [21] sebagai instrumen sederhana untuk menilai kebergunaan sistem. Instrumen terdiri atas 10 pernyataan berskala 1–5; pernyataan ganjil bersifat positif dan pernyataan genap bersifat negatif. Skor kontribusi butir ganjil adalah nilai jawaban dikurangi 1, sedangkan butir genap adalah 5 dikurangi nilai jawaban. Skor SUS responden ke-k dihitung dengan

$$SUS_k = 2{,}5 \times \sum_{j=1}^{10} X_{kj} \qquad (2)$$

dengan X_{kj} skor kontribusi butir ke-j. Skor akhir adalah rata-rata seluruh responden pada rentang 0–100. Skor 68 digunakan sebagai rata-rata acuan [22], dan skor sekitar 71 setara dengan penilaian adjektiva “baik” [23].

I. Normalized Gain (N-Gain)

Hake [18] mendefinisikan *gain* ternormalisasi sebagai perbandingan antara peningkatan skor aktual dan peningkatan maksimum yang mungkin dicapai. Untuk responden ke-i, N-Gain dihitung dengan

$$g_i = \frac{Post_i - Pre_i}{Maks - Pre_i} \qquad (3)$$

sedangkan untuk satu kelompok Hake menggunakan rata-rata skor kelas:

$$\langle g \rangle = \frac{\langle Post \rangle - \langle Pre \rangle}{Maks - \langle Pre \rangle} \qquad (4)$$

Kategori yang digunakan adalah tinggi ($g \geq 0{,}70$), sedang ($0{,}30 \leq g < 0{,}70$), dan rendah ($g < 0{,}30$) [18]. N-Gain individu tidak terdefinisi bila skor awal telah maksimal karena penyebutnya nol, dan kurang bermakna bagi kelompok yang rata-rata skor awalnya mendekati maksimal.

III. Metodologi Penelitian

A. Rancangan Penelitian

Penelitian ini merupakan penelitian dan pengembangan (*Research and Development*) dengan model ADDIE yang dipilih karena tahapannya terstruktur dan lazim dipakai untuk media pembelajaran [19]. Produk dievaluasi melalui validasi ahli, *black-box testing*, uji *usability* dengan SUS, dan desain satu kelompok *pre-test–post-test* tanpa kelompok kontrol. Dengan desain ini, perubahan skor tidak dapat dikaitkan secara kausal dengan permainan sehingga hasilnya diperlakukan sebagai indikasi awal.

B. Tahapan Pengembangan

1) Analysis: Kebutuhan dipetakan melalui observasi terhadap empat permainan logika di Roblox dan kajian literatur; wawancara tidak dilakukan karena kebutuhan dinilai sudah cukup terpetakan dari kedua sumber tersebut. Pengguna sasaran adalah mahasiswa berusia 18–25 tahun, konten berupa tiga konsep algoritma logika, dan perangkat berupa Roblox Studio dengan Lua.

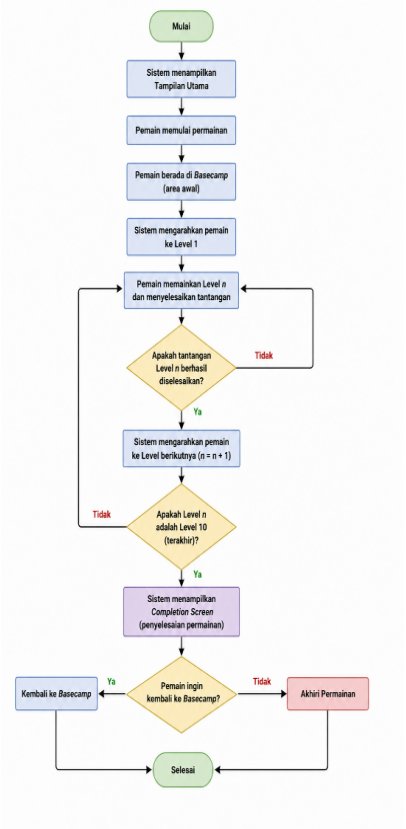
2) Design: Disusun *game design document*, diagram alur sistem, *storyboard* antarmuka, rancangan 10 level berjenjang, serta instrumen penelitian (angket validasi, kuesioner SUS, dan soal tes).

3) Development: Elemen 3D dan antarmuka dibuat di Roblox Studio, skrip Lua mengatur validasi solusi dan progresi level, kemudian dilakukan *alpha testing* dan perbaikan bertahap.

4) Implementation: Permainan diujicobakan kepada mahasiswa dengan prosedur pada Bagian III-D.

5) Evaluation: Data validasi ahli, SUS, serta *pre-test* dan *post-test* dianalisis dengan (1)–(4).

Alur sistem permainan dirancang linier namun dapat diulang (Gambar 1): pemain memulai dari tampilan utama dan *basecamp*, memainkan level ke-*n*, dan hanya berpindah ke level berikutnya jika tantangan terselesaikan. Setelah Level 10, sistem menampilkan layar penyelesaian dan pemain dapat kembali ke *basecamp* atau mengakhiri permainan.



Gambar. 1. Diagram alur sistem permainan Nalar.exe.

C. Rancangan Permainan

Nalar.exe dirancang sebagai permainan teka-teki berbasis algoritma logika yang menekankan penyelesaian tantangan secara bertahap, bukan alur cerita. Sistem progresi mewajibkan pemain menuntaskan setiap level sebelum berpindah ke level berikutnya, sehingga konsep dipelajari dari yang dasar hingga yang kompleks. Pembagian level pada tiap fase disesuaikan dengan porsi soal tes agar tantangan dalam permainan selaras dengan kemampuan yang diukur. Seluruh level memakai pola interaksi yang konsisten, yaitu pemain berinteraksi melalui objek yang dapat diakses dan sistem memverifikasi setiap langkah.

Mahasiswa dipilih sebagai pengguna sasaran karena tingkat kematangan kognitifnya, keakrabannya dengan platform digital dan budaya permainan, kemampuan pemecahan masalah yang telah terlatih, serta kapasitasnya untuk belajar secara implisit. Antarmuka disusun berdasarkan prinsip kejelasan, kemudahan akses, dan keterlibatan, dengan estetika futuristik bergaya siber yang dekat dengan selera visual kelompok usia tersebut.

D. Subjek, Instrumen, dan Prosedur Pengujian

Uji coba melibatkan mahasiswa berusia 18–25 tahun: 10 responden untuk *pre-test* dan *post-test*, serta 15 responden untuk SUS. Ukuran sampel mengacu pada uji lapangan awal penelitian pengembangan yang lazim melibatkan 10–30 peserta [24] dan pada penelitian SUS dengan 10 responden [25]. Validasi dilakukan oleh satu ahli algoritma logika (12 butir) dan satu ahli *game design* (8 butir) menggunakan angket Likert 1–5.

Instrumen tes disusun peneliti berdasarkan tiga konsep logika dan proporsi level, yaitu 10 soal pilihan ganda berbasis skenario dengan empat opsi (3 soal logika berurutan, 4 pengenalan pola, dan 3 logika bersyarat). Setiap jawaban benar bernilai satu poin dan skor dikonversi ke skala 0–100. Instrumen yang sama digunakan pada *pre-test* dan *post-test* dengan waktu pengerjaan 30 menit. Validitas dan reliabilitas butir belum diuji secara statistik.

Prosedur pengujian terdiri atas empat langkah: (1) responden mengerjakan *pre-test* secara individual; (2) responden memainkan Nalar.exe secara mandiri dari Level 1 hingga Level 10 sesuai skenario yang ditetapkan; (3) responden mengerjakan *post-test* dengan instrumen yang sama; dan (4) responden mengisi kuesioner SUS. Pengujian fungsional dilakukan peneliti pada seluruh level.

Instrumen tes memuat 10 soal dengan sebaran yang mengikuti jumlah level pada tiap fase (Tabel 2). Kuesioner SUS memuat 10 pernyataan pada Tabel 3 yang dijawab dengan skala 1 (sangat tidak setuju) sampai 5 (sangat setuju).

TABEL II
SEBARAN SOAL PRE-TEST DAN POST-TEST

Konsep Logika	Level Terkait	Jumlah Soal
Sequential logic	1–3	3
Pattern recognition	4–7	4
Conditional logic	8–10	3
Total	**1–10**	**10**

TABEL III
PERNYATAAN KUESIONER SUS

No.	Pernyataan
1	Saya akan sering menggunakan sistem ini.
2	Sistem ini terasa terlalu rumit tanpa perlu.
3	Sistem ini mudah digunakan.
4	Saya memerlukan bantuan orang teknis untuk menggunakan sistem ini.
5	Fungsi-fungsi pada sistem ini terintegrasi dengan baik.
6	Terdapat terlalu banyak inkonsistensi pada sistem ini.
7	Kebanyakan orang akan cepat belajar menggunakan sistem ini.
8	Sistem ini sangat merepotkan untuk digunakan.
9	Saya merasa sangat percaya diri menggunakan sistem ini.
10	Saya perlu mempelajari banyak hal sebelum dapat menggunakan sistem ini.

Validasi dilakukan sebelum permainan diujicobakan kepada mahasiswa. Setiap ahli memainkan permainan, mengisi angket Likert, dan memberikan catatan tertulis yang kemudian dicatat sebagai bahan perbaikan (Tabel 7).

Validator menilai aspek konten, penyampaian, bahasa, tampilan, dan fungsi media dengan skala Likert, dan persentase kelayakan yang diperoleh dibandingkan dengan standar kelayakan yang ditetapkan untuk menentukan apakah produk siap digunakan atau perlu diperbaiki. Kuesioner SUS

JUTEI Edisi Volume. 10 No.2 Oktober 2026
ISSN 2579-3675, e-ISSN-5538
DOI 10.21460/jutei.2026.102.538

diisi setelah responden menyelesaikan permainan. Identitas responden dicatat dengan kode (R1–R15) sehingga data dianalisis tanpa nama.

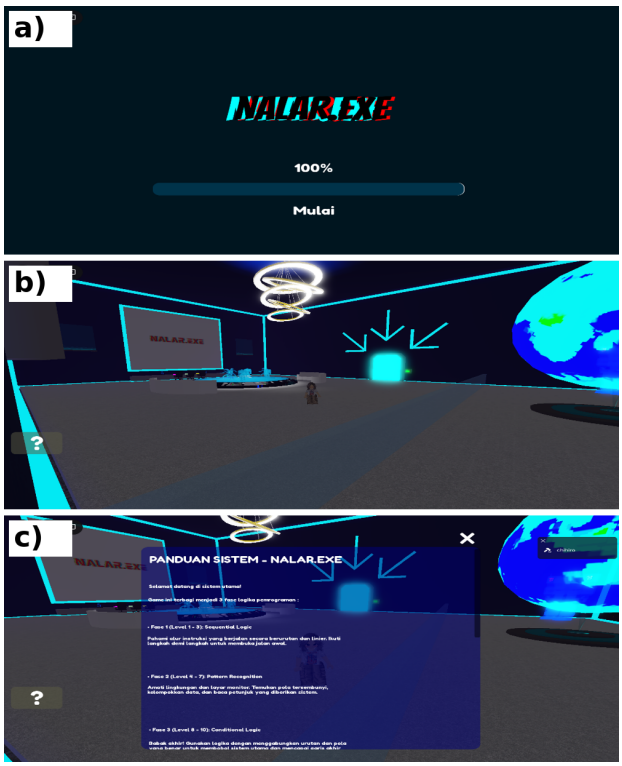
E. Teknik Analisis Data

Kelayakan dianalisis dengan (1), SUS dengan (2), dan peningkatan skor dengan N-Gain pada (3) dan (4). Karena N-Gain individu tidak terdefinisi bila skor awal telah maksimal, perhitungan individu hanya mencakup responden dengan skor awal di bawah 100, sedangkan N-Gain kelas dihitung dari rata-rata seluruh responden. Hasil dilaporkan secara deskriptif tanpa uji inferensial karena jumlah responden kecil.

IV. HASIL DAN ANALISIS

A. Produk yang Dihasilkan

Pengembangan menghasilkan Nalar.exe, permainan teka-teki 3D dengan 10 level yang diselesaikan berurutan, mulai dari *basecamp* hingga layar penyelesaian. Antarmukanya bertema digital dan siber (Gambar 2). Layar pemuatan menampilkan judul "NALAR.EXE" dengan efek *glitch* merah dan cyan, serta bilah kemajuan dan tombol Mulai yang muncul setelah pemuatan selesai. *Basecamp* berupa ruang bernuansa navy gelap dengan garis cyan bercahaya dan pintu menuju Level 1. Panel Panduan Sistem dapat dibuka kapan saja melalui tombol bantuan dan merangkum tiga fase logika sebagai rujukan bila pemain kebingungan.



Gambar. 2. Antarmuka umum: (a) layar pemuatan, (b) basecamp, dan (c) panel panduan sistem.

Setiap level dirancang agar tantangannya meningkat bertahap dalam satu fase, kemudian berganti konsep pada fase berikutnya. Pemain memperoleh umpan balik visual langsung setelah setiap percobaan dan tidak dikenai penalti; kesalahan hanya mengembalikan *puzzle* ke keadaan awal. Rancangan ini memberi latihan pada langkah memahami masalah, mencoba rencana, dan memeriksa hasil sebagaimana dalam kerangka Polya [17]. Tampilan seluruh level disajikan pada Gambar 3, dengan rincian tiap fase sebagai berikut.

1) Fase Sequential Logic (Level 1–3): Level 1 mengenalkan gagasan bahwa keberhasilan bergantung pada urutan tindakan, dengan dua tombol yang terlihat jelas. Level 2 menambah tahap eksplorasi karena pemain harus menemukan tombol tersembunyi sebelum menentukan urutan. Level 3, yang tersulit pada fase ini, mengangkat cerita sistem yang sedang diretas sehingga pemain menyusun tahapan penyelamatan data (*scan*, *extract*, *encrypt*, *backup*).

2) Fase Pattern Recognition (Level 4–7): Level 4 menampilkan sembilan lingkaran dengan satu posisi kosong yang diisi dengan objek dari meja di samping. Level 5 menambah variasi warna melalui meja penggabungan sehingga pemain harus mengenali pola dari kombinasi dua warna. Level 6 memuat dua tipe bentuk (segitiga dan lingkaran) dengan unsur yang hilang pada kelompok segitiga. Level 7 mengangkat narasi *data corruption*: pemain memulihkan *node* data yang hilang, dan *node* terakhir berbentuk berbeda sehingga harus dipulihkan lebih dahulu sebelum dapat diaktifkan. Level ini memadukan pengenalan pola dengan urutan tindakan.

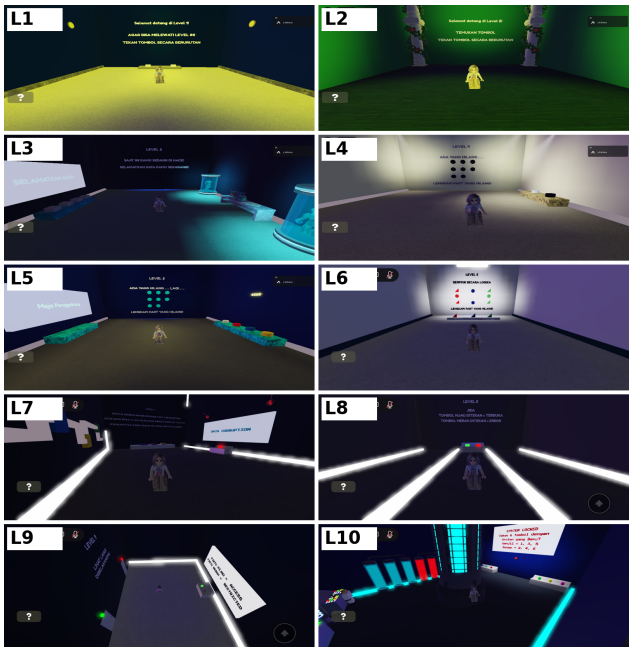
TABEL IV
RANCANGAN LEVEL NALAR.EXE

Level	Mekanik Utama	Tantangan Kognitif
1	Dua tombol harus ditekan berurutan	Urutan aksi
2	Mencari tombol tersembunyi, lalu menekan berurutan	Eksplorasi dan urutan
3	Menyusun empat tahap penyelamatan data	Urutan proses
4	Melengkapi pola sembilan lingkaran	Pola dasar
5	Menggabungkan dua warna untuk melengkapi pola	Pola dengan variasi warna
6	Melengkapi bentuk segitiga di antara lingkaran	Dua tipe bentuk
7	Memulihkan *node* data berbentuk berbeda	Pola dan urutan
8	Memilih tombol hijau atau merah sesuai instruksi	Percabangan sederhana
9	Memilih pintu berdasarkan lampu yang diacak	Banyak sumber informasi
10	Menekan enam tombol: ganjil lalu genap	Syarat dan kategori

3) Fase Conditional Logic (Level 8–10): Level 8 mengenalkan struktur *if-else* secara visual: tombol hijau membuka pintu, tombol merah menimbulkan kesalahan. Level 9 menambah sumber informasi, yaitu dua pintu berlabel akses dan terbatas serta lampu pada layar yang menunjukkan pintu yang benar. Pengacakan pada Level 9 membantu menghindari konten tetap (*hard-coded*) yang dapat mengurangi daya main ulang [26]. Level 10 menampilkan pesan "System Locked" yang mengharuskan pemain menekan enam tombol dengan urutan ganjil (1, 3, 5) lebih dulu, lalu genap (2, 4, 6). Setelah Level 10, layar "You

Win!" bergaya piksel muncul dan pemain dapat kembali ke *basecamp*.

Rancangan mekanik dan tantangan kognitif tiap level diringkas pada Tabel 4.



Gambar. 3. Tampilan Level 1 sampai Level 10 Nalar.exe (L1–L10), diurutkan dari kiri ke kanan dan atas ke bawah.

B. *Implementasi Logika Program*

Skrip Lua pada tiap fase menerapkan pemeriksaan kondisi terhadap tindakan pemain. Pada Level 1, sistem mencatat penekanan tombol pendukung; saat tombol utama ditekan, sistem memeriksa apakah tombol pendukung sudah aktif. Bila ya, penghalang dihancurkan dan jalan terbuka; bila tidak, hanya efek kesalahan sementara yang muncul dan pemain mengulang. Pada Level 4, sistem memeriksa jenis objek yang dipasang: objek benar menyelesaikan level, sedangkan objek pengecoh memunculkan visual berbeda dan memberi kesempatan mencoba objek lain. Pada Level 9, sistem mengacak pintu yang benar setiap level dimulai, menyalakan lampu penanda, lalu memeriksa pilihan pemain; pilihan salah mengembalikan pemain ke titik awal dan mengacak ulang kondisi. Logika Level 1 diringkas pada pseudokode berikut.

```
if tombolUtama ditekan then
    if tombolPendukung aktif then bukaJalan()
    else efekSalah() end
end
```

C. *Pengujian Fungsional*

Pengujian *black-box* dilakukan pada seluruh 10 level dengan membandingkan hasil yang diharapkan dan hasil aktual, seperti dirinci pada Tabel 5. Seluruh level (100%) berjalan sesuai rancangan, mulai dari deteksi masukan, validasi urutan aksi, hingga transisi antarlevel. Pengujian ini hanya memastikan fungsi program, bukan efektivitas pembelajaran.

TABEL V
HASIL PENGUJIAN FUNGSIONAL PER LEVEL

Level	Fase	Skenario Pengujian	Hasil
1	*Sequential*	Menekan dua tombol sesuai urutan	Berhasil
2	*Sequential*	Mencari tombol tersembunyi lalu menekan sesuai urutan	Berhasil
3	*Sequential*	Mengurutkan empat tahap penyelamatan data	Berhasil
4	*Pattern*	Melengkapi pola sembilan lingkaran	Berhasil
5	*Pattern*	Menggabungkan dua warna untuk melengkapi pola	Berhasil
6	*Pattern*	Melengkapi bentuk segitiga yang hilang	Berhasil
7	*Pattern*	Memulihkan *node* data berbentuk berbeda	Berhasil
8	*Conditional*	Menekan tombol sesuai instruksi *if-else*	Berhasil
9	*Conditional*	Memilih pintu berdasarkan petunjuk lampu (acak)	Berhasil
10	*Conditional*	Menekan enam tombol: ganjil lalu genap	Berhasil
Total berhasil			**10/10 (100%)**

D. *Hasil Validasi Ahli*

Hasil validasi disajikan pada Tabel 6. Ahli algoritma logika memberi persentase kelayakan 76,67% dan menyimpulkan bahwa permainan layak diproduksi. Ahli *game design* memberi 75,00% dan menyimpulkan bahwa permainan baik dengan revisi kecil, dengan catatan perlunya penghilangan *bug* yang masih ditemukan serta penambahan *sound effect*. Hasil ini bersumber dari satu ahli untuk setiap bidang sehingga belum dapat digeneralisasi.

Menurut aspeknya, ahli algoritma logika memberi persentase yang lebih tinggi pada aspek pembelajaran implisit (80,00%) daripada aspek kompetensi dan pembelajaran (73,33%). Hal ini menunjukkan bahwa penyisipan konsep logika ke dalam mekanik permainan dinilai cukup baik, sedangkan keterkaitan eksplisit dengan kompetensi masih dapat diperkuat. Pada ahli *game design*, aspek performa teknis memperoleh 80,00% dan aspek desain, interaksi, dan alur memperoleh 73,33%, sejalan dengan catatan bahwa pengalaman bermain masih dapat diperkaya melalui efek suara.

TABEL VI
HASIL VALIDASI AHLI

Validator	Aspek	Butir	Skor	%
Ahli algoritma logika	Kompetensi dan pembelajaran	6	22/30	73,33
	Pembelajaran implisit	6	24/30	80,00
Total		**12**	**46/60**	**76,67**
Ahli game design	Desain, interaksi, dan alur	6	22/30	73,33
	Performa teknis dan kualitas keseluruhan	2	8/10	80,00
Total		**8**	**30/40**	**75,00**

JUTEI Edisi Volume. 10 No.2 Oktober 2026
ISSN 2579-3675, e-ISSN-5538
DOI 10.21460/jutei.2026.102.538

E. Tindak Lanjut Masukan Ahli dan Temuan Uji Coba

Masukan ahli dan temuan uji coba awal dicatat pada Tabel 7. Perbaikan tersebut belum diterapkan pada versi yang diuji dan menjadi agenda pengembangan berikutnya.

TABEL VII
MASUKAN AHLI DAN TINDAK LANJUT

Sumber	Masukan atau Temuan	Tindak Lanjut
Ahli *game design*	Masih ditemukan *bug* teknis	Perbaikan *bug* pada versi berikutnya
Ahli *game design*	Perlu penambahan *sound effect* untuk pengalaman bermain	Penambahan efek suara pada versi berikutnya
Uji coba awal	Level 7 dan 10: urutan aksi kurang jelas dan tombol tanpa penanda angka	Perjelas instruksi dan tambahkan penanda angka

F. Hasil Pre-Test dan Post-Test

Tabel 8 memuat skor setiap responden. Lima responden (R1, R2, R3, R5, R10) telah memperoleh skor 100 pada *pre-test*, sehingga penyebut Maks − Pre sama dengan nol dan N-Gain individunya tidak terdefinisi. Kondisi ini adalah *ceiling effect*: instrumen tidak mampu mendeteksi peningkatan pada responden yang sudah mencapai skor tertinggi. Dengan demikian, N-Gain individu hanya dapat dihitung untuk lima responden. Nilai pada penelitian ini telah dihitung ulang dengan (3) dan (4).

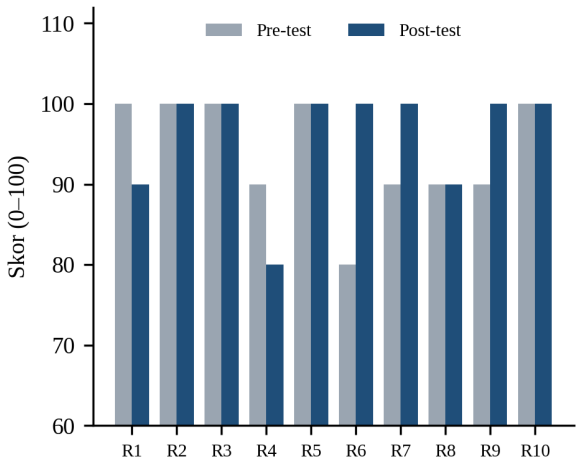
TABEL VIII
HASIL PRE-TEST, POST-TEST, DAN N-GAIN

Resp.	Pre	Post	Selisih	N-Gain	Kategori
R1	100	90	−10	*Tidak terdefinisi*	
R2	100	100	0	*Tidak terdefinisi*	
R3	100	100	0	*Tidak terdefinisi*	
R4	90	80	−10	−1,00	Rendah
R5	100	100	0	*Tidak terdefinisi*	
R6	80	100	+20	1,00	Tinggi
R7	90	100	+10	1,00	Tinggi
R8	90	90	0	0,00	Rendah
R9	90	100	+10	1,00	Tinggi
R10	100	100	0	*Tidak terdefinisi*	
Rata-rata kelas (n=10)	**94,0**	**96,0**	**+2,0**	**0,33**	**Sedang**
Rata-rata N-Gain individu (n=5)				**0,40**	**Sedang**

Rata-rata skor kelas naik dari 94,0 menjadi 96,0, sehingga N-Gain kelas menurut (4) sebesar 0,33 (kategori sedang). N-Gain individu pada responden yang dapat dihitung adalah R4 (−1,00), R6 (1,00), R7 (1,00), R8 (0,00), R9 (1,00), dengan rata-rata 0,40 (kategori sedang). Dari 10 responden, tiga mengalami kenaikan skor, lima tetap, dan dua turun.

N-Gain kelas dan rata-rata N-Gain individu sama-sama berkategori sedang, tetapi estimasi tersebut bertumpu pada lima responden yang masih memiliki ruang peningkatan: tiga di antaranya memperoleh N-Gain 1,00, satu tidak berubah, dan satu menurun. Separuh responden telah mencapai skor maksimal pada *pre-test*, dan nilai N-Gain yang besar pada beberapa responden bertumpu pada pembagi kecil (selisih 10 sampai 20 poin) sehingga sangat sensitif terhadap satu jawaban. Hasil ini baru dapat disebut indikasi awal, bukan bukti efektivitas.

Sebagai contoh perhitungan, R6 memperoleh skor 80 pada *pre-test* dan 100 pada *post-test*, sehingga menurut (3) $g = (100 - 80) / (100 - 80) = 1{,}00$. R4 turun dari 90 menjadi 80, sehingga $g = (80 - 90) / (100 - 90) = -1{,}00$. Untuk kelas, (4) menghasilkan $g = (96{,}0 - 94{,}0) / (100 - 94{,}0) = 0{,}33$. Karena penyebut kelas hanya 6,0, perubahan satu jawaban pada satu responden (10 poin) menggeser N-Gain kelas sebesar 0,17. Sensitivitas ini menunjukkan bahwa nilai N-Gain pada data dengan skor awal tinggi tidak stabil dan perlu ditafsirkan dengan sangat hati-hati.



Gambar. 4. Skor pre-test dan post-test setiap responden.

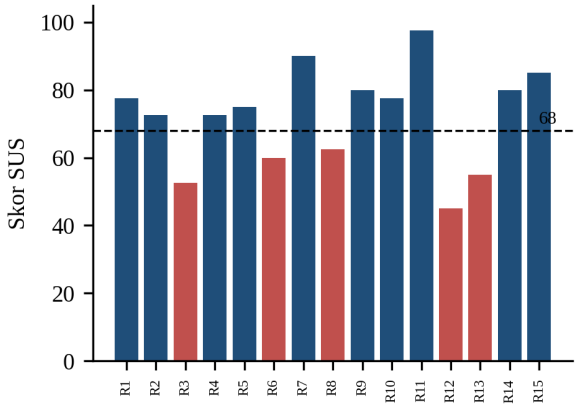
G. Hasil Uji Usability (SUS)

Kuesioner SUS diisi oleh 15 responden dan ringkasannya disajikan pada Tabel 9. Rata-rata skor 72,17 berada di atas acuan 68 [22] dan setara dengan penilaian "baik" [23], dengan 10 dari 15 responden (66,7%) memberi skor di atas 68. Rentang yang lebar (45,0 sampai 97,5; simpangan baku 14,57) menunjukkan pengalaman pengguna yang tidak seragam, termasuk satu responden dengan skor 45,0. Temuan ini sejalan dengan masukan ahli tentang *bug* dan efek suara. SUS menilai kemudahan penggunaan, bukan efektivitas pembelajaran [27].

TABEL IX
SKOR SUS PER RESPONDEN

Resp.	Total	Skor	Resp.	Total	Skor
R1	31	77,5	R9	32	80,0
R2	29	72,5	R10	31	77,5
R3	21	52,5	R11	39	97,5
R4	29	72,5	R12	18	45,0
R5	30	75,0	R13	22	55,0
R6	24	60,0	R14	32	80,0
R7	36	90,0	R15	34	85,0
R8	25	62,5	**Rata-rata**		**72,17**

Median skor adalah 75,0, dengan 5 responden (R3, R6, R8, R12, R13) memberi skor di bawah 68. Sebaran ini menunjukkan bahwa penilaian kebergunaan baik secara umum, tetapi masih ada pengguna yang mengalami hambatan sehingga hasil rata-rata tidak boleh dibaca sebagai kepuasan seragam.



Gambar. 5. Skor SUS setiap responden; batang merah berada di bawah acuan 68 (garis putus-putus).

H. Pembahasan dan Keterbatasan

Hasil penelitian menunjukkan bahwa Nalar.exe layak secara fungsional dan menurut ahli, serta dinilai cukup mudah digunakan. Bukti efektivitasnya, sebaliknya, masih lemah. Temuan ini belum dapat mengonfirmasi manfaat permainan teka-teki yang dilaporkan penelitian lain [5], [10] karena instrumen tes tidak mampu membedakan kemampuan responden. Soal berbasis skenario dengan opsi pengecoh yang jelas keliru tampaknya terlalu mudah, sehingga sebagian besar responden sudah memperoleh skor maksimal sebelum bermain. Penyempurnaan butir agar lebih menantang, misalnya dengan lebih banyak langkah penalaran dan opsi pengecoh yang masuk akal, perlu dilakukan sebelum instrumen dipakai kembali.

Ketercapaian tujuan penelitian dirangkum pada Tabel 10.

TABEL X
KETERCAPAIAN TUJUAN PENELITIAN

Tujuan	Indikator	Hasil	Status
Mengembangkan Nalar.exe	10 level dalam tiga fase	Seluruh level berfungsi (10/10)	Tercapai
Menilai kelayakan	Uji fungsional dan validasi ahli	100%; 76,67%; 75,00%	Tercapai, revisi kecil
Menilai *usability*	Skor SUS	72,17 (acuan 68)	Tercapai, kategori baik
Indikasi peningkatan kemampuan	N-Gain kelas dan individu	0,33; 0,40 (sedang)	Indikasi awal terbatas

Ditinjau menurut tujuan penelitian, pertama, pengembangan dengan ADDIE menghasilkan produk yang utuh dan dapat dijalankan, dan alur linier dengan umpan balik tanpa penalti memberi ruang bagi pemain untuk mencoba kembali. Kedua, persentase kelayakan sekitar 75% pada kedua ahli menunjukkan produk layak, namun belum berada pada kategori sangat layak, sejalan dengan catatan tentang *bug* dan efek suara. Ketiga, skor SUS di atas 68 menunjukkan bahwa mekanik permainan dapat dipahami oleh sebagian besar pengguna, meskipun tiga dari lima responden dengan skor rendah (R3, R12, R13) berada di bawah 56, sehingga kejelasan instruksi pada level yang lebih sulit tetap perlu diperbaiki. Keempat, pengukuran peningkatan kemampuan belum dapat dipertanggungjawabkan, sebagaimana diuraikan di atas.

Implikasi praktisnya adalah bahwa pengembang permainan edukatif perlu merancang instrumen ukur bersamaan dengan merancang permainan. Permainan yang menarik dan mudah digunakan tidak otomatis menghasilkan peningkatan yang terukur apabila tes awal tidak memberi ruang bagi peningkatan. Perbaikan instrumen, penambahan kelompok pembanding, dan pengulangan uji pada sampel yang lebih besar merupakan langkah yang lebih penting daripada penambahan level baru pada tahap berikutnya.

Beberapa ancaman terhadap validitas perlu diakui. Penggunaan instrumen yang sama pada *pre-test* dan *post-test* dapat menimbulkan efek pengulangan tes. Tanpa kelompok kontrol, perubahan skor tidak dapat dipisahkan dari faktor lain seperti kematangan atau kebiasaan mengerjakan soal. Ukuran dan cara pemilihan sampel yang terbatas dapat membuat responden tidak mewakili populasi mahasiswa. Selain itu, jumlah responden yang kecil membuat setiap perubahan satu jawaban berdampak besar pada N-Gain individu.

Keterbatasan lain meliputi jumlah responden yang kecil, desain satu kelompok tanpa kelompok kontrol, validasi oleh satu ahli untuk setiap bidang, dan instrumen yang belum diuji validitas serta reliabilitasnya. Selain itu, masukan ahli belum diterapkan pada versi yang diuji. Penelitian lanjutan perlu menggunakan soal yang lebih sulit dan bertingkat dengan analisis tingkat kesukaran dan daya beda, menambah jumlah soal dan responden, melibatkan kelompok kontrol, serta memperbaiki permainan sesuai Tabel 7.

I. Keterkaitan dengan Penelitian Terdahulu

Weng dkk. [10] melaporkan peningkatan kemampuan pemecahan masalah yang bermakna setelah peserta memainkan permainan teka-teki logika, dan meta-analisis Gui dkk. [5] menunjukkan manfaat permainan digital pada pembelajaran STEM. Penelitian ini tidak dapat mereplikasi atau membantah temuan tersebut karena jumlah responden yang kecil dan instrumen yang tidak sensitif. Dibandingkan dengan permainan fisik pada penelitian Wardani dkk. [11], Nalar.exe menyediakan umpan balik otomatis dan pengacakan kondisi, tetapi belum diuji efektivitasnya dengan rancangan yang setara. Sementara itu, Azhar dkk. [12] juga memakai Roblox dan ADDIE, namun berfokus pada materi etnomatematika; penelitian ini menambah contoh pemakaian Roblox untuk konsep algoritma logika yang tidak terikat pada materi tertentu. Pada tahap awal pengembangan seperti ini, validasi ahli dan uji kebergunaan baru memberi bukti kelayakan, sehingga bukti efektivitas sebaiknya diposisikan secara hati-hati.

J. Rekomendasi Penyempurnaan Instrumen

Berdasarkan temuan ini, instrumen tes pada penelitian lanjutan perlu diperkuat dalam beberapa hal. Pertama, jumlah butir ditambah dan tingkat kesukaran dibuat beragam, termasuk butir yang menuntut beberapa langkah penalaran. Kedua, opsi pengecoh dirancang masuk akal sehingga tidak dapat dieliminasi tanpa memahami konsep. Ketiga, butir diuji coba lebih dahulu untuk memperoleh indeks kesukaran, daya beda, dan koefisien reliabilitas. Keempat, bentuk tes awal dan tes akhir dapat dibuat paralel untuk mengurangi

JUTEI Edisi Volume. 10 No.2 Oktober 2026
ISSN 2579-3675, e-ISSN-5538
DOI 10.21460/jutei.2026.102.538

efek pengulangan tes. Kelima, kelompok kontrol dan tes tertunda dapat ditambahkan agar perubahan skor dapat ditafsirkan lebih tegas.

V. KESIMPULAN

Nalar.exe berhasil dikembangkan dengan model ADDIE sebagai permainan 3D di Roblox Studio yang berisi 10 level dalam tiga fase logika, dan seluruh levelnya berfungsi sesuai rancangan. Ahli menilai permainan ini layak dengan persentase 76,67% (algoritma logika) dan 75,00% (*game design*), sedangkan rata-rata skor SUS 72,17 menunjukkan kebergunaan yang baik. Mengenai efektivitas, N-Gain kelas sebesar 0,33 dan rata-rata N-Gain individu sebesar 0,40 berkategori sedang, tetapi lima dari 10 responden telah memperoleh skor maksimal pada *pre-test* sehingga estimasi bertumpu pada sedikit responden. Dengan demikian, penelitian ini hanya memberikan indikasi awal, bukan bukti efektivitas. Penelitian lanjutan disarankan menggunakan instrumen yang lebih menantang dan telah diuji validitas serta reliabilitasnya, sampel yang lebih besar, kelompok kontrol, serta perbaikan permainan berdasarkan masukan ahli.

Bagi pengembang media serupa, hasil ini menegaskan bahwa kelayakan teknis dan kebergunaan yang baik perlu dilengkapi dengan alat ukur yang peka terhadap perubahan. Bagi pendidik, Nalar.exe dapat dipertimbangkan sebagai latihan pendukung, tetapi belum dapat diandalkan sebagai bukti peningkatan kemampuan pemecahan masalah.

UCAPAN TERIMA KASIH

Penulis mengucapkan terima kasih kepada para ahli validator dan seluruh mahasiswa yang berpartisipasi sebagai responden dalam penelitian ini.

KONFLIK KEPENTINGAN

Para penulis menyatakan tidak ada konflik kepentingan terkait dengan publikasi artikel ini.

PENDANAAN

Penelitian ini tidak mendapatkan dukungan keuangan spesifik dari lembaga pendanaan sektor publik, komersial, atau nirlaba.

DEKLARASI PENGGUNAAN ALAT AI

Dalam penelitian ini, penulis menggunakan Gemini (Google) untuk membantu penulisan kode program Lua pada pengembangan permainan Nalar.exe di Roblox Studio, serta Claude (Anthropic) untuk membantu menyusun ulang isi skripsi ke dalam format artikel, menyunting bahasa, dan menghitung ulang N-Gain. Setelah menggunakan kedua alat tersebut, penulis meninjau, menguji, dan merevisi seluruh keluarannya sebagaimana diperlukan, dan bertanggung jawab penuh atas keseluruhan isi publikasi.

KETERSEDIAAN DATA

Skor *pre-test*, *post-test*, dan SUS setiap responden disajikan pada Tabel 8 dan Tabel 9. Data lengkap dan instrumen penelitian tersedia dari penulis korespondensi atas permintaan yang wajar.

KONTRIBUSI PENULIS

AADM merancang penelitian, merancang desain permainan, dan melakukan pengkodean permainan Nalar.exe, serta mengumpulkan dan menganalisis data dan menyusun draf naskah. SHW membantu pembuatan sistem dan mendukung proses pengkodean permainan. MI membantu perancangan desain permainan, termasuk penataan tampilan dan alur antarlevel. AKH melakukan supervisi dan penelaahan kritis terhadap penelitian dan naskah. Semua penulis telah membaca dan menyetujui versi final naskah ini.

DAFTAR PUSTAKA

[1] D. Indraswati *et al.*, "Critical thinking dan problem solving dalam pembelajaran IPS untuk menjawab tantangan abad 21," *Sosial Horizon: Jurnal Pendidikan Sosial*, vol. 7, no. 1, hal. 12–28, 2020.

[2] Y. Liando, T. Sihotang, dan H. Kuswanto, "Effects of problem solving skills in school learning: A systematic literature review," *Indonesian Journal of Educational Research and Review*, vol. 8, hal. 70–79, 2025.

[3] A. Latip *et al.*, "Implementasi dan implikasi media berbasis game edukasi pada pembelajaran: A systematic literature review," *Paedagoria: Jurnal Kajian, Pendidikan dan Pengembangan Kependidikan*, hal. 255–262, 2024.

[4] F. N. Ulumudin dan B. Sujatmiko, "Pengembangan media pembelajaran berbasis game edukasi Running Maze pemrograman dasar," *IT-Edu: Jurnal Information Technology and Education*, vol. 8, 2023.

[5] Y. Gui *et al.*, "Effectiveness of digital educational game and game design in STEM learning: A meta-analytic review," *International Journal of STEM Education*, 2023, doi: 10.1186/s40594-023-00424-9.

[6] K. Alhasan, K. Alhasan, dan S. Al Hashimi, "Roblox in higher education: Opportunities, challenges, and future directions for multimedia learning," *International Journal of Emerging Technologies in Learning*, vol. 18, no. 19, hal. 32–46, 2023, doi: 10.3991/ijet.v18i19.43133.

[7] J. Han, G. Liu, dan Y. Gao, "Learners in the metaverse: A systematic review on the use of Roblox in learning," *Education Sciences*, vol. 13, no. 3, hal. 296, 2023, doi: 10.3390/educsci13030296.

[8] S. A. Eltomi dan R. S. Mottaleb, "Video game development on Roblox platform using Lua programming language," *Gharyan Journal of Technology*, no. 10, hal. 97–122, 2024.

[9] O. Yolal, "Roblox Studio ile mühendislik eğitimi için deneyim geliştirme," *International Journal of Multidisciplinary Studies and Innovative Technologies*, vol. 6, no. 1, hal. 51–57, 2022, doi: 10.36287/ijmsit.6.1.51.

[10] T. Weng *et al.*, "Enhancing problem-solving ability through a puzzle-type logical thinking game," *Scientific Programming*, vol. 2022, 2022, doi: 10.1155/2022/7481798.

[11] S. S. Wardani, R. D. Susanti, dan M. Taufik, "Implementasi pendekatan computational thinking melalui game Jungle Adventure terhadap kemampuan problem solving," *SJME (Supremum Journal of Mathematics Education)*, vol. 6, no. 1, hal. 1–13, 2022.

[12] F. Azhar, N. Agustyaningrum, dan Y. Franita, "Development of Geoarcala Roblox game integrating ethnomathematics of Magelang temples to enhance critical thinking ability," *Prisma Sains: Jurnal Pengkajian Ilmu dan Pembelajaran Matematika dan IPA IKIP Mataram*, vol. 14, no. 2, hal. 535–554, 2026.

[13] M. F. A. Romadhan *et al.*, "Penerapan model project-based learning (PjBL) dalam pengembangan modul pembelajaran pembuatan game Obby berbasis Roblox Studio," *Jurnal Media Akademik (JMA)*, vol. 3, no. 12, 2025.

[14] J. M. Wing, "Computational thinking," *Communications of the ACM*, vol. 49, no. 3, hal. 33–35, Mar. 2006.

[15] D. Ariani *et al.*, "Urgensi pengembangan instrumen tes keterampilan berpikir komputasional bertema biologi bagi siswa madrasah aliyah," *Jurnal Jeumpa: Jurnal Pendidikan Sains dan Biologi*, vol. 11, no. 2, hal. 337–348, 2025, doi: 10.33059/jj.v11i2.10831.

[16] M. Saleh *et al.*, "Meningkatkan keterampilan pemecahan masalah mahasiswa melalui penggunaan simulasi dan game edukasi pada matematika dan IPA," *Indo Green Journal*, vol. 4, hal. 128–135, 2026.
[17] G. Polya, *How to Solve It: A New Aspect of Mathematical Method*. Princeton, NJ, USA: Princeton Univ. Press, 1945.
[18] R. R. Hake, "Interactive-engagement versus traditional methods: A six-thousand-student survey of mechanics test data for introductory physics courses," *American Journal of Physics*, vol. 66, no. 1, hal. 64–74, Jan. 1998.
[19] R. M. Branch, *Instructional Design: The ADDIE Approach*. New York, NY, USA: Springer, 2009.
[20] Akhyanto, N. Suarna, dan A. I. Purnamasari, "Game edukasi ilmu tajwid berbasis Android menggunakan metode ADDIE untuk meningkatkan minat belajar siswa," *Jurnal Informatika Terpadu*, vol. 8, no. 2, hal. 117–126, 2022.
[21] J. Brooke, "SUS: A 'quick and dirty' usability scale," dalam *Usability Evaluation in Industry*, P. W. Jordan, B. Thomas, B. A. Weerdmeester, dan I. L. McClelland, Ed. London, U.K.: Taylor & Francis, 1996, hal. 189–194.
[22] J. Sauro dan J. R. Lewis, *Quantifying the User Experience: Practical Statistics for User Research*, edisi ke-2. Cambridge, MA, USA: Morgan Kaufmann, 2016.
[23] A. Bangor, P. T. Kortum, dan J. T. Miller, "An empirical evaluation of the System Usability Scale," *International Journal of Human–Computer Interaction*, vol. 24, no. 6, hal. 574–594, 2008.
[24] Okpatrioka, "Research and development (R&D) penelitian yang inovatif dalam pendidikan," *Dharma Acariya Nusantara: Jurnal Pendidikan, Bahasa dan Budaya*, vol. 1, no. 1, 2023.
[25] A. R. Bahtiar dan M. A. Gustalika, "Penerapan metode System Usability Scale dalam pengujian rancangan mobile apps gamification tari rakyat di Indonesia," *MIB*, vol. 6, hal. 491–499, 2022, doi: 10.30865/mib.v6i1.3510.
[26] T. Asrori, D. Isyuniandri, dan M. Imron, "Model konseptual value-driven procedural content generation (PCG) untuk pengembangan game edukasi karakter," *Jurnal Aplikasi Teknologi Informasi dan Manajemen (JATIM)*, vol. 6, no. 2, hal. 49–57, 2025.
[27] O. Suria, "A statistical analysis of System Usability Scale (SUS) evaluations in online learning platform," *Journal of Information System and Informatics*, vol. 6, no. 2, hal. 992–1007, 2024, doi: 10.51519/journalisi.v6i2.750.

JUTEI Edisi Volume. 10 No.2 Oktober 2026
ISSN 2579-3675, e-ISSN-5538
DOI 10.21460/jutei.2026.102.538