



# SCHOOL EDUCATION JOURNAL PGSD FIP UNIMED

Volume 16 No. 2 Juni 2026

The journal contains the result of education research, learning research, and service of the public at primary school, elementary school, senior high school and the university  
<https://jurnal.unimed.ac.id/2012/index.php/school>



## PENGEMBANGAN INSTRUMEN PENILAIAN BAHASA INDONESIA BERBANTUAN GENERATIVE ARTIFICIAL INTELLIGENCE (GenAI) UNTUK MENGUKUR KREATIVITAS PESERTA DIDIK KELAS IV SEKOLAH DASAR

Ana Afifatur Rakhmah<sup>1</sup>, Deri Anggraini<sup>2</sup>

Program Studi Guru Sekolah Dasar Fakultas Keguruan dan Ilmu Pendidikan Universitas  
PGRI Yogyakarta<sup>1,2</sup>

Surel: [anaarkhmh@gmail.com](mailto:anaarkhmh@gmail.com)

### ABSTRACT

This study aimed to develop a Generative Artificial Intelligence (GenAI)-assisted Indonesian language creativity assessment instrument for fourth-grade elementary school students. The developed product consists of *JurnalMembaca* as a digital assessment platform for students and *JurnalKreasi* as a web-based assessment application for teachers. This study employed the Research and Development (R&D) method using the Four-D model (Define, Design, Develop, and Disseminate). The subjects were 46 fourth-grade students of SD Negeri Panggang, Bantul. Data were collected through expert validation, instrument testing, limited trials, and practicality questionnaires. Data were analyzed using percentage analysis, Pearson Product-Moment correlation, Cronbach's Alpha, item difficulty analysis, and discrimination index analysis. The results showed feasibility scores of 95.24% from the educational assessment expert and 92.71% from the subject matter expert, both categorized as highly feasible. Instrument testing indicated that 17 out of 20 items were valid, with a Cronbach's Alpha reliability coefficient of 0.954, categorized as very high. The practicality scores reached 87.50% from the teacher and 88.12% from the students, both categorized as highly practical. Therefore, the developed instrument was found to be feasible, valid, reliable, and practical for measuring elementary school students' creativity.

**Keywords:** Creativity, Assessment Instrument, GenAI

### ABSTRAK

Penelitian ini bertujuan mengembangkan instrumen penilaian kreativitas Bahasa Indonesia berbantuan Generative Artificial Intelligence (GenAI) untuk peserta didik kelas IV sekolah dasar. Produk yang dikembangkan terdiri atas *JurnalMembaca* sebagai platform penilaian digital peserta didik dan *JurnalKreasi* sebagai aplikasi penilaian berbasis web untuk guru. Penelitian ini menggunakan metode Research and Development (R&D) dengan model Four-D (Define, Design, Develop, dan Disseminate). Subjek penelitian berjumlah 46 peserta didik kelas IV SD Negeri Panggang, Bantul. Data dikumpulkan melalui validasi ahli, uji instrumen, uji coba terbatas, dan angket kepraktisan. Analisis data menggunakan persentase, korelasi Product Moment Pearson, Alpha Cronbach, analisis tingkat kesukaran, dan daya pembeda. Hasil penelitian menunjukkan persentase kelayakan sebesar 95,24% dari ahli evaluasi pembelajaran dan 92,71% dari ahli materi dengan kategori sangat layak. Hasil uji instrumen menunjukkan 17 dari 20 butir soal valid dengan koefisien reliabilitas Alpha Cronbach sebesar 0,954 yang termasuk kategori sangat tinggi. Kepraktisan instrumen memperoleh persentase 87,50% dari guru dan 88,12% dari peserta didik dengan kategori sangat praktis. Dengan demikian, instrumen yang dikembangkan dinyatakan layak, valid, reliabel, dan praktis untuk digunakan dalam mengukur kreativitas peserta didik sekolah dasar.

**Kata Kunci:** Kreativitas, Instrumen Penilaian, GenAI

Copyright (c) 2026 Ana Afifatur Rakhmah<sup>1</sup>, Deri Anggraini<sup>2</sup>

✉ Corresponding author:

Email : [anaarkhmh@gmail.com](mailto:anaarkhmh@gmail.com)

HP : 081225638641

ISSN 2355-1720 (Media Cetak)

ISSN 2407-4926 (Media Online)

Received 20 May 2026, Accepted 16 June 2026, Published 17 June 2026

DOI : [10.24114/82nx6k84](https://doi.org/10.24114/82nx6k84)

## PENDAHULUAN

Pendidikan abad ke-21 mengalami perubahan cara pandang, dari yang semula berfokus pada penyampaian pengetahuan menjadi pembelajaran yang berorientasi pada pengembangan kompetensi. Kreativitas menjadi salah satu kompetensi penting yang perlu ditanamkan sejak jenjang pendidikan dasar (Setiani, 2022). Dalam kerangka kompetensi 4C yang meliputi berpikir kritis, komunikasi, kolaborasi, dan kreativitas, setiap individu dituntut mampu menyesuaikan diri terhadap perubahan yang cepat serta menciptakan gagasan inovatif (Nurhayati et al., 2024). Bahasa Indonesia menjadi salah satu mata pelajaran yang memiliki ruang luas untuk menumbuhkan kreativitas, karena kreativitas tidak muncul begitu saja melainkan bermula dari proses berpikir dalam penyampaian gagasan yang terencana dan sesuai konteks (Irawan & Mukhlis, 2023).

Namun pada kenyataannya, penilaian kreativitas peserta didik di sekolah dasar masih menghadapi keterbatasan mendasar. Instrumen yang tersedia umumnya disajikan dalam bentuk rubrik berdasarkan indikator kompetensi dasar secara umum, sehingga belum menyentuh aspek kreativitas secara mendalam (Aprilia et al., 2021). Penilaian yang berfokus pada hasil akhir juga belum mampu memberikan gambaran utuh tentang proses kreativitas peserta didik (Heryahya et al., 2022). Kondisi ini terkonfirmasi melalui hasil wawancara dengan guru kelas IVA dan IVB SD Negeri Panggang, yang menyatakan bahwa instrumen penilaian yang digunakan cenderung tidak diperbarui, masih bersifat umum seperti tes tulis biasa, dan belum pernah memanfaatkan *GenAI* untuk keperluan penilaian.

Kebutuhan akan instrumen penilaian

yang lebih adaptif sejalan dengan kemajuan *Generative Artificial Intelligence (GenAI)*, yang merupakan cabang kecerdasan buatan yang berfokus pada kemampuan mesin menghasilkan konten baru berdasarkan pola yang dipelajari dari data (Andriyani et al., 2024). Pemanfaatan *GenAI* memungkinkan pendidik merancang pendekatan evaluasi yang tidak hanya mengukur capaian kognitif, tetapi juga mengidentifikasi dan menstimulasi pengembangan keterampilan peserta didik (Xian-feng, 2025). Meskipun sejumlah penelitian telah mengkaji pemanfaatan *GenAI* dalam pembelajaran, belum ada yang secara khusus mengembangkan instrumen penilaian kreativitas pada mata pelajaran Bahasa Indonesia di sekolah dasar berbantuan *GenAI*. Oleh karena itu, penelitian ini mengembangkan instrumen penilaian kreativitas Bahasa Indonesia berbantuan *GenAI* melalui model 4D (*Define, Design, Develop, Disseminate*) yang mencakup empat dimensi kreativitas: *fluency* (kelancaran), *flexibility* (keluwesan), *originality* (keaslian), dan *elaboration* (elaborasi).

## METODE PENELITIAN

Penelitian ini menggunakan metode *Research and Development (R&D)* dengan model pengembangan *Four-D (4D)* yang terdiri dari empat tahap: *Define, Design, Develop, dan Disseminate*. Lokasi penelitian adalah SD Negeri Panggang, Bantul. Subjek penelitian berjumlah 46 peserta didik kelas IV (kelas IVA dan IVB), dengan 22 peserta didik sebagai subjek uji instrumen dan 24 peserta didik sebagai subjek uji coba terbatas. Penelitian dilaksanakan pada tahun ajaran 2025/2026.

Pengumpulan data dilakukan melalui wawancara terstruktur, observasi kelas, validasi ahli (ahli evaluasi

pembelajaran dan ahli materi), uji instrumen, uji coba terbatas, serta angket kepraktisan guru dan peserta didik. Analisis data menggunakan analisis persentase kelayakan dengan mengacu pada kriteria Akbar (2013), uji validitas butir soal dengan korelasi Product Moment Pearson ( $r$  tabel = 0,4227 untuk  $n=22$ ), uji reliabilitas dengan koefisien Alpha Cronbach, analisis tingkat kesukaran, dan analisis daya pembeda.

Produk yang dikembangkan terdiri dari dua komponen utama sebagai berikut. (1) JurnalMembaca, platform penilaian digital berbasis *Genially* dengan konsep paspor petualangan untuk peserta didik; dan (2) JurnalKreasi, aplikasi penilaian berbasis web yang dibangun menggunakan *HTML*, *CSS*, dan *JavaScript* dengan integrasi *GenAI* (ClaudeAI dan *Gemini*) untuk membantu guru melakukan penilaian kreativitas. Dalam penelitian ini, *GenAI* berperan sebagai teknologi pendukung pengembangan produk, meliputi pembuatan ilustrasi visual, pengembangan konten, dan integrasi sistem digital, sedangkan perancangan isi, indikator penilaian, dan mekanisme penilaian tetap disusun oleh peneliti berdasarkan landasan teori.

## HASIL PENELITIAN DAN PEMBAHASAN

### Proses Pengembangan Produk

Tahap *Define* (Pendefinisian) dilakukan melalui wawancara terstruktur dan observasi langsung di kelas IV SD Negeri Panggang. Hasil wawancara mengungkapkan bahwa guru menghadapi keterbatasan waktu dalam memperbarui instrumen penilaian, rubrik yang digunakan masih bersifat umum, dan pemanfaatan *GenAI* untuk penilaian belum pernah dilakukan. Hasil observasi menunjukkan bahwa pembelajaran masih

didominasi penjelasan guru, instrumen yang digunakan masih berupa tes tulis sumatif, dan variasi jawaban peserta didik masih terbatas. Temuan ini menunjukkan kesenjangan antara kebutuhan penilaian kreativitas dengan instrumen yang tersedia.

Tahap *Design* (Perancangan) menghasilkan rancangan dua komponen terintegrasi. JurnalMembaca dirancang dengan konsep paspor petualangan menggunakan platform *Genially*, yang memungkinkan integrasi elemen visual, animasi, dan kuis interaktif. Konsep gamifikasi ini dipilih berdasarkan temuan Triwahyu et al. (2024) yang menyimpulkan bahwa gamifikasi berpotensi meningkatkan motivasi dan keterlibatan peserta didik. JurnalKreasi dirancang sebagai platform berbasis web yang mengintegrasikan rubrik analitik, kisi-kisi, dan jawaban ideal dalam satu sistem untuk membantu guru melakukan penilaian secara efisien. Rubrik penilaian disusun berdasarkan empat dimensi kreativitas: *fluency*, *flexibility*, *originality*, dan *elaboration* (Aprilia et al., 2021).

Tahap *Develop* (Pengembangan) meliputi penilaian kelayakan oleh para ahli dan uji coba lapangan. Hasil penilaian menunjukkan persentase kelayakan sebesar 95,24% dari ahli evaluasi pembelajaran dan 92,71% dari ahli materi, keduanya termasuk kategori Sangat Layak berdasarkan kriteria Akbar (2013). Revisi produk dilakukan berdasarkan masukan validator, terutama pada kejelasan indikator rubrik penilaian dan deskripsi kriteria penilaian agar lebih operasional dan mudah dipahami.

Tahap *Disseminate* (Penyebaran) dilakukan secara terbatas di SD Negeri Panggang melalui pengenalan instrumen kepada guru disertai pendampingan penggunaan sistem, mulai dari pengunggahan

data peserta didik, pelaksanaan penilaian, hingga ekspor hasil dalam bentuk PDF.

### Kelayakan Produk

Penilaian Ahli Evaluasi Pembelajaran menghasilkan persentase kelayakan 95,24% (skor 80 dari maksimal 84) yang mencakup aspek tampilan, konstruksi, dan pemrograman. Seluruh indikator aspek tampilan memperoleh skor maksimal, menunjukkan bahwa instrumen telah memenuhi kriteria kejelasan visual, kesesuaian warna, dan kesesuaian dengan karakteristik peserta didik kelas IV SD (Arsyad, 2013). Pada aspek konstruksi, beberapa indikator memperoleh skor 3 namun tetap dalam kategori layak. Pada aspek pemrograman, hampir seluruh indikator memperoleh skor maksimal kecuali indikator keandalan program.

Penilaian Ahli Materi menghasilkan persentase kelayakan 92,71% (skor 89 dari maksimal 96) yang mencakup aspek materi dan bahasa. Pada aspek materi, indikator kesesuaian dengan Capaian Pembelajaran, ketepatan kisi-kisi kreativitas, dan kesesuaian rubrik penilaian dengan indikator kreativitas memperoleh skor tinggi. Seluruh indikator aspek bahasa memperoleh skor maksimal, menunjukkan bahwa bahasa yang digunakan sesuai kaidah, komunikatif, dan mudah dipahami peserta didik kelas IV SD.

Analisis Kualitas Butir Soal dilakukan terhadap 20 butir soal yang diujicobakan pada 22 peserta didik. Hasil analisis tingkat kesukaran menunjukkan 16 butir soal (80%) berkategori sedang dan 4 butir soal (20%) berkategori mudah, tanpa ada butir soal berkategori sukar. Hasil analisis daya pembeda menunjukkan 11 butir soal berkategori cukup (S05, S06, S07, S10, S11, S13, S14, S15, S16, S19, S20) dan 9 butir soal

berkategori buruk. Hasil uji validitas menunjukkan 17 dari 20 butir soal valid ( $r$  hitung  $>$   $r$  tabel 0,4227), sedangkan 3 butir soal (S02, S08, S09) dinyatakan tidak valid. Berdasarkan pertimbangan gabungan validitas, tingkat kesukaran, dan daya pembeda, ditetapkan 11 butir soal sebagai instrumen final. Uji reliabilitas terhadap 11 butir soal final menghasilkan nilai Cronbach's Alpha sebesar 0,954, termasuk kategori sangat tinggi. Rangkuman kelayakan produk disajikan pada Tabel 1.

Tabel 1. Rekap Kelayakan Instrumen

Aspek Kelayakan	Persentase	Kategori
Ahli Evaluasi Pembelajaran	95,24%	Sangat Layak
Ahli Materi	92,71%	Sangat Layak
Validitas Butir Soal	17/20 valid	Valid
Reliabilitas (Alpha Cronbach)	0,954	Sangat Tinggi
Tingkat Kesukaran	80% sedang	Memadai
Daya Pembeda (instrumen final)	11 butir cukup	Dapat Digunakan

### Kepraktisan Produk

Penilaian Guru menghasilkan persentase kepraktisan total sebesar 87,50% (kategori Sangat Praktis), dengan aspek kemudahan penggunaan 85,71% dan aspek kebermanfaatan 91,67%. Tingginya kepraktisan menunjukkan bahwa instrumen dapat digunakan tanpa pelatihan khusus, serta sistem penilaian terintegrasi membantu mengurangi beban guru dalam proses koreksi dan penilaian (Mittal et al., 2024).

Penilaian Peserta Didik menghasilkan persentase kepraktisan sebesar 88,12% (kategori Sangat Praktis), dengan aspek kemudahan penggunaan 88,39% dan aspek kebermanfaatan 87,50%. Sebanyak 20 dari 24 peserta didik (83,33%) berada pada kategori Sangat Praktis dan 4 peserta didik pada

kategori Praktis. Penggunaan konsep paspor petualangan pada JurnalMembaca membantu menciptakan pengalaman belajar yang menyenangkan dan mendorong keterlibatan aktif peserta didik, sejalan dengan pendapat Piaget bahwa peserta didik sekolah dasar

berada pada tahap operasional konkret yang membutuhkan media visual dan dekat dengan pengalaman nyata. Hasil kepraktisan produk disajikan pada Tabel 2.

**Tabel 2. Rekap Kepraktisan Produk**

Penilai	Kemudahan Penggunaan	Kebermanfaatan	Total
Guru	85,71%	91,67%	87,50%
Peserta Didik	88,39%	87,50%	88,12%

Secara keseluruhan, instrumen penilaian kreativitas Bahasa Indonesia berbantuan *GenAI* dinyatakan layak, valid, reliabel, dan praktis. Integrasi teknologi *GenAI* dalam instrumen tidak hanya menghasilkan produk yang layak secara teoritis, tetapi juga mudah diterapkan dalam kondisi pembelajaran nyata di sekolah dasar.

## SIMPULAN

Berdasarkan hasil penelitian, dapat disimpulkan tiga hal berikut. Pertama, proses pengembangan instrumen dilakukan melalui model 4D: tahap *define* mengidentifikasi kesenjangan antara kebutuhan penilaian kreativitas dan instrumen yang tersedia; tahap *design* menghasilkan rancangan JurnalMembaca dan JurnalKreasi; tahap *develop* mencakup validasi ahli dan uji coba lapangan; serta tahap *disseminate* dilakukan secara terbatas di SD Negeri Panggang. Kedua, instrumen dinyatakan sangat layak dengan persentase 95,24% dari ahli evaluasi pembelajaran dan 92,71% dari ahli materi, serta memenuhi kriteria kualitas butir soal dengan 17 dari 20 butir valid, nilai reliabilitas Alpha Cronbach 0,954 (sangat tinggi), dan 11 butir soal terpilih sebagai instrumen final. Ketiga, instrumen dinyatakan sangat praktis

dengan persentase 87,50% dari guru dan 88,12% dari peserta didik, menunjukkan bahwa instrumen mudah digunakan, bermanfaat, dan mendukung keterlibatan aktif peserta didik dalam pembelajaran Bahasa Indonesia kelas IV sekolah dasar.

## DAFTAR RUJUKAN

- Andriyani, W., et al. (2024). *AI Generatif dan Mutu Pendidikan*. Bandung: Penerbit Widina.
- Aprilia, M., Lidinillah, D. A. M., & Giyartini, R. (2021). *Pengembangan instrumen penilaian kreativitas siswa melalui analisis Rasch model di sekolah dasar*. *Jurnal Basicedu*, 5(4), 2302–2310. <https://doi.org/10.31004/basicedu.v5i4.1199>
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Haryanti, Y. D., & Saputra, D. S. (2019). *Instrumen penilaian berpikir kreatif pada pendidikan abad 21*. *Jurnal Cakrawala Pendas*, 5(2). <https://doi.org/10.31949/jcp.v5i2.1350>
- Indrawaty, R., Rohaeti, E. E., & Mahardika, R. Y. (2025). *Effectiveness of Genially interactive media in improving presentation competence of elementary*

- school students*. *Inovasi Kurikulum*, 22(4), 2161–2172.
- Legi, M. B., Sengkey, D. F., & Sambul, A. M. (2024). *An application of generative artificial intelligence for automated rubric-based grading*. *Jurnal Teknik Informatika*, 19(3), 183–192.
- Mittal, U., et al. (2024). *A comprehensive review on generative AI for education*. *IEEE Access*, 12, 142733–142759. <https://doi.org/10.1109/access.2024.3468368>
- Naseer, F., et al. (2024). *Automated assessment and feedback in higher education using generative AI*. *IGI Global*. <https://doi.org/10.4018/979-8-3693-1351-0.ch021>
- Nurhayati, I., Pramono, K. S. E., & Farida, A. (2024). *Keterampilan 4C dalam pembelajaran IPS untuk menjawab tantangan abad 21*. *Jurnal Basicedu*, 8(1), 36–43. <https://doi.org/10.31004/basicedu.v8i1.6842>
- Setiani, A. E. (2022). *Pengembangan instrumen penilaian HOTS dari buku tematik menggunakan Quizizz di sekolah dasar*. *El-Ibtidaiy Journal of Primary Education*, 5(1), 90. <https://doi.org/10.24014/ejpe.v5i1.14324>
- Triwahyu, A., et al. (2024). *Perkembangan Gamifikasi di Bidang Pendidikan*. *Scholaria*, vol. 14 no. 2. <https://doi.org/10.24246/j.js.2024.v14.i2.p177-186>
- Waruwu, M. (2024). *Metode penelitian dan pengembangan (R&D): konsep, jenis, tahapan dan kelebihan*. *Jurnal Ilmiah Profesi Pendidikan*, 9(2), 1220–1230. <https://doi.org/10.29303/jipp.v9i2.2141>