

**PENGEMBANGAN INSTRUMEN PENILAIAN BAHASA INDONESIA
BERBANTUAN *GENERATIVE ARTIFICIAL INTELLIGENCE (GenAI)*
UNTUK MENGUKUR KEMAMPUAN KOMUNIKASI PESERTA
DIDIK KELAS IV DI SD NEGERI WINONGO**

Saputra¹, Deri Anggraini²
Universitas PGRI Yogyakarta
Email: putraa20201@gmail.com

Abstrak

Tujuan penelitian ini adalah menghasilkan instrumen penilaian Bahasa Indonesia berbantuan *Generative Artificial Intelligence (GenAI)* yang layak dan praktis digunakan untuk mengukur kemampuan komunikasi peserta didik kelas IV. Penelitian ini menggunakan pendekatan *Research and Development (R&D)* mengacu pada model 4D dari Thiagarajan, dengan 4 tahap yaitu: *Define; Design; Develop; dan Disseminate*. Produk yang dikembangkan berupa instrumen penilaian nontes dalam bentuk rubrik kemampuan komunikasi yang disajikan dalam format buku, dilengkapi dengan soal nontes berbantuan *GenAI* melalui platform *Wordwall*. Hasil penelitian menunjukkan bahwa: 1) instrumen penilaian dinyatakan sangat layak dengan persentase skor ahli materi sebesar 97,5% dan ahli evaluasi sebesar 98,75%, serta rata-rata kelayakan sebesar 98,125%; 2) instrumen penilaian dinyatakan sangat praktis dengan persentase skor kepraktisan guru sebesar 87,5%. Uji validitas menunjukkan seluruh 18 butir instrumen dinyatakan valid (r hitung tertinggi 0,876, terendah 0,483 > r tabel 0,422). Uji reliabilitas memperlihatkan nilai *Cronbach's Alpha* sebesar 0,937 dengan kategori sangat reliabel. Dengan demikian, instrumen penilaian Bahasa Indonesia berbantuan *GenAI* terbukti sangat layak dan sangat praktis digunakan untuk mengukur kemampuan komunikasi pada peserta didik kelas IV SD Negeri Winongo.

Kata kunci: *instrumen penilaian, generative artificial intelligence, kemampuan komunikasi, sekolah dasar.*

Abstract:

The purpose of this study is to produce an Indonesian language assessment instrument assisted by Generative Artificial Intelligence (GenAI) that is feasible and practical to use in measuring the communication skills of fourth-grade students. This research uses a Research and Development (R&D) approach referring to Thiagarajan's 4D model, with four stages, namely: Define; Design;

Develop; and Disseminate (Fajri & Taufiqurrahman, 2017). The product developed is a non-test assessment instrument in the form of a communication skills rubric presented in book format, complemented by GenAI-assisted non-test questions through the Wordwall platform. The results of the study indicate that: 1) the assessment instrument is deemed very feasible with a material expert score percentage of 97.5% and an evaluation expert score of 98.75%, with an average feasibility of 98.125%; 2) the assessment instrument is considered very practical with a teacher practicality score percentage of 87.5%. Validity tests showed that all 18 items of the instrument were declared valid (highest calculated r 0.876, lowest 0.483 > table r 0.422). Reliability tests showed a Cronbach's Alpha value of 0.937 with a very reliable category. Thus, the Indonesian language assessment instrument assisted by GenAI is proven to be very feasible and very practical to use for measuring communication skills in fourth-grade students at SD Negeri Winongo.

Keywords: *assessment instrument, generative artificial intelligence, communication skills, elementary school.*

Pendahuluan

Pendidikan merupakan suatu usaha yang dilakukan dengan sengaja dan terstruktur untuk membangun suasana pembelajaran yang mendukung siswa dalam mengeksplorasi dan mengembangkan bakat yang terdapat dalam diri mereka secara proaktif (Rahman et al., 2022) Dalam konteks pendidikan dasar, pembelajaran Bahasa Indonesia memiliki peran strategis dalam mengembangkan keterampilan berbahasa peserta didik, termasuk kemampuan komunikasi. Komunikasi dapat diartikan sebagai suatu proses pengiriman pesan, gagasan, maupun informasi yang dilakukan oleh seseorang atau kelompok kepada pihak lainnya (Hasanah et al., 2022). Sejalan dengan pendapat tersebut, (Nurmansyah et al., 2023) menyatakan bahwa komunikasi merupakan upaya untuk membuat atau mendapatkan sebuah pendapat maupun ide agar mudah dipahami oleh orang lain. Akan tetapi, melalui serangkaian observasi dan wawancara yang berlangsung di SD Negeri Winongo pada bulan Oktober 2025, ditemukan bahwa kemampuan komunikasi peserta didik kelas IV masih bervariasi, instrumen penilaian yang digunakan guru masih bersifat konvensional, dan belum tersedia rubrik penilaian yang baku untuk mengukur seluruh aspek kemampuan komunikasi secara sistematis. Pernyataan ini diperkuat oleh (Anjelina & Tarmimi, 2022) yang mengungkapkan bahwa keterampilan berbicara siswa sekolah dasar dalam menggunakan Bahasa Indonesia masih tergolong rendah, terutama pada aspek pelafalan, intonasi, kelancaran, isi pembicaraan, dan ekspresi.

Dalam menghadapi permasalahan tersebut, integrasi teknologi *Generative Artificial Intelligence (GenAI)* menjadi salah satu solusi inovatif.

GenAI adalah sistem mesin yang dirancang untuk menduplikasi kecerdasan manusia serta memberikan kemudahan dalam menyelesaikan pekerjaan yang membutuhkan kecerdasan manusia (Hidayat et al., 2025). (Julianto & Ratumanan, 2023) menyatakan bahwa pemanfaatan *GenAI* dalam pembelajaran bahasa untuk siswa SD mempunyai potensi yang tinggi guna meningkatkan efisiensi dan kualitas instrumen penilaian yang dihasilkan, termasuk dalam aspek kemampuan komunikasi. Oleh karena itu, penelitian ini bertujuan untuk menghasilkan instrumen penilaian Bahasa Indonesia berbantuan *GenAI* yang layak dan praktis digunakan dalam mengukur kemampuan komunikasi peserta didik kelas IV SD.

Metode Penelitian

Penelitian ini dilaksanakan dengan menerapkan pendekatan *Research and Development* (R&D) yang berpedoman pada model pengembangan 4D. dari Thiagarajan (Fajri & Taufiqurrahman, 2017). Model 4D yang meliputi 4 tahap , diantaranya: a) *define* (pendefinisian); b) *design* (perancangan); c) *develop* (pengembangan); dan d) *disseminate* (penyebaran). Penelitian dilaksanakan di SD Negeri Winongo, Kabupaten Bantul, Daerah Istimewa Yogyakarta pada bulan Maret 2026. Subjek uji coba dalam penelitian ini meliputi: a) ahli materi yang menilai aspek kelayakan isi dan kebahasaan; b) ahli evaluasi yang menilai aspek tampilan, konstruksi, dan pemrograman; c) guru kelas IV sebagai penilai kepraktisan produk, serta; d) peserta didik kelas IV B sebagai subjek uji coba instrumen dan IV B SD Negeri Winongo sebagai subjek uji coba terbatas.

Pengambilan data dalam penelitian ini dilaksanakan dengan beberapa teknik, seperti wawancara, observasi, penyebaran angket, dan dokumentasi. Instrumen penelitian yang dikembangkan antara lain: pedoman wawancara guru, pedoman observasi pembelajaran, angket kelayakan produk untuk ahli materi dan ahli evaluasi, angket kepraktisan produk untuk guru, serta pedoman observasi kemampuan komunikasi. Seluruh instrumen divalidasi terlebih dahulu oleh ahli. Kriteria penilaian kelayakan produk instrumen penilaian ditentukan berdasarkan hasil angket validasi ahli materi dan ahli evaluasi dengan menggunakan skala Likert 1–4. Data dianalisis menggunakan rumus persentase kelayakan (Hakim et al., 2021). Hasil persentase kemudian dikonversikan ke dalam kategori kelayakan menurut (Sriwijayanti et al.,2020: 98), yaitu: 81%–100% (sangat layak); 61%–80% (layak); 41%–60% (cukup layak); 21%–40% (belum layak); dan 0%–20% (sangat belum layak).

Kriteria penilaian kepraktisan produk instrumen penilaian diperoleh dari hasil angket yang diisi oleh guru kelas IV setelah menggunakan instrumen

yang dikembangkan. Persentase kepraktisan dihitung dengan rumus yang sama dan dikategorikan menurut (Hakim et al., 2021), yaitu: 81%–100% (sangat praktis); 61%–80% (praktis); 41%–60% (cukup praktis); 21%–40% (belum praktis); dan 0%–20% (sangat belum praktis). Analisis data dilaksanakan melalui uji validitas instrumen yang berfungsi untuk mengukur sejauh mana setiap butir pernyataan tepat dalam merepresentasikan indikator kemampuan komunikasi yang telah ditetapkan, sedangkan uji reliabilitas dilaksanakan guna mengetahui seberapa konsisten instrumen tersebut jika digunakan berulang kali pada kondisi yang serupa.

Hasil dan Pembahasan

Hasil Pengembangan Produk

Produk yang dihasilkan pada penelitian pengembangan ini yaitu instrumen penilaian nontes dalam bentuk rubrik kemampuan komunikasi yang disajikan dalam format buku cetak, dilengkapi dengan soal nontes berbantuan *Generative Artificial Intelligence (GenAI)* melalui platform *Wordwall*. Pengembangan produk berpedoman pada model pengembangan 4D yang dirancang Thiagarajan terdiri dari 4 tahapan yaitu: *define* (pendefinisian); *design* (perancangan); *develop* (pengembangan); dan *disseminate* (penyebaran). Pada tahap *define*, peneliti melakukan analisis kebutuhan pembelajaran yang menunjukkan bahwa kemampuan komunikasi peserta didik masih bervariasi dan instrumen penilaian yang digunakan guru masih bersifat konvensional. Selanjutnya pada tahap *design*, peneliti merancang instrumen penilaian dengan tujuh indikator kemampuan komunikasi, yaitu ucapan atau pelafalan, keberanian, pilihan kata, sikap wajar dan tidak kaku, gerak-gerik dan mimik yang tepat, kelancaran berbicara, serta penguasaan topik atau materi (Rahmah et al., 2024).

Buku instrumen penilaian yang dikembangkan memuat berbagai komponen, antara lain sampul, daftar isi, kata pengantar, latar belakang, capaian pembelajaran dan tujuan pembelajaran, indikator kemampuan komunikasi, kisi-kisi soal, rubrik penilaian, pedoman penskoran, serta soal berbasis *Wordwall* yang dilengkapi dengan kode batang (*barcode*). Desain keseluruhan produk dibuat dengan bantuan *Dreamina AI* untuk latar gambar dan *Canva* untuk penataan teks serta tata letak, sehingga menghasilkan produk yang menarik, sistematis, dan siap divalidasi oleh para ahli.

Hasil Kelayakan Produk

Penilaian oleh ahli materi dilaksanakan terhadap Ibu Adeasi Ufita Hemafisah, S.Pd., Gr. pada 17 Februari 2026. Hasil evaluasi memperoleh skor total 78 dari 80, sehingga menghasilkan persentase kelayakan 97,5% yang

termasuk dalam predikat "Sangat Layak". Sementara itu, validasi oleh ahli evaluasi dilakukan kepada Ibu Selly Rahmawati, M.Pd. pada 3 Maret 2026. Skor yang diperoleh mencapai 79 dari 80, dengan persentase kelayakan 98,75% yang juga tergolong "Sangat Layak".

Tabel 1. Rekapitulasi hasil kelayakan validator

No	Validasi Ahli	Skor	Skor Maksimal	Persentase	Kategori
1	Ahli Materi	78	80	97,5%.	Sangat Layak
2	Ahli Evaluasi	79	80	98,75%.	Sangat Layak
Nilai Rata-Rata		78,5	80	98.125%	Sangat Layak

Hasil Uji Validitas

Pengujian validitas dilaksanakan terhadap 22 peserta didik yang berasal dari kelas IV A. Dengan derajat kebebasan (df) sebesar 20 dan taraf signifikansi 5%, diperoleh nilai r tabel 0,422. Hasil pengujian menunjukkan bahwa keseluruhan 18 butir instrumen dinyatakan valid, dengan nilai koefisien korelasi (r hitung) paling tinggi mencapai 0,876 dan paling rendah sebesar 0,483. Selain itu, seluruh nilai yang teramati memiliki signifikansi di bawah 0,05, ini menunjukkan adanya keterkaitan yang berarti antara nilai setiap butir dengan total skor instrumen.

Hasil Uji Reliabilitas

Pengujian reliabilitas instrumen penilaian dilaksanakan dengan metode *Cronbach's Alpha* dan menghasilkan koefisien sebesar 0,937. Apabila merujuk pada kriteria yang dikemukakan oleh (Arikunto, 2021), nilai tersebut termasuk dalam rentang 0,80 hingga 1,00 sehingga dikategorikan sebagai "Sangat Reliabel". Temuan ini mengindikasikan bahwa instrumen penilaian yang dikembangkan memiliki konsistensi internal yang sangat baik, sehingga dapat diaplikasikan secara berulang dan tetap menghasilkan temuan yang relatif sama.

Hasil Kepraktisan Produk

Penilaian terhadap tingkat kepraktisan produk dilaksanakan pada tanggal 10 Maret 2026 bersama Ibu Septi Widowati, S.Pd., yang menjabat sebagai guru kelas IV B. Aspek-aspek yang dinilai mencakup kemudahan dalam penggunaan serta manfaat yang diberikan oleh produk. Hasil penilaian menunjukkan perolehan skor total 70 dari skor maksimal 80, dengan persentase 87,5% yang termasuk dalam kategori "Sangat Praktis". Menurut penilaian guru, instrumen penilaian ini tergolong mudah dioperasikan, dilengkapi petunjuk yang jelas, memiliki kesesuaian materi dan bahasa, menampilkan desain yang

menarik, serta bermanfaat dalam mewujudkan pembelajaran yang berkesadaran, bermakna, dan menyenangkan..

Pembahasan

Berdasarkan hasil penelitian, instrumen penilaian Bahasa Indonesia berbantuan *GenAI* dinyatakan sangat layak dengan rata-rata persentase kelayakan sebesar 98,125%. Temuan ini sejalan dengan penelitian (Riscaputantri & Wening, 2018) yang mengembangkan instrumen penilaian afektif siswa kelas IV SD dengan hasil validitas isi yang sangat valid. Tingginya validasi ahli materi menunjukkan bahwa instrumen penilaian telah memenuhi aspek kelayakan isi, kebahasaan, dan evaluasi. Menurut teori validitas, validitas didefinisikan sebagai bagaimana mana sebuah instrumen penilaian dapat mengukur apa yang harus diukur secara benar (Arikunto, 2021).

Keunggulan produk ini terletak pada pemanfaatan GenAI dalam menyusun konten nontes yang lebih kontekstual dan beragam. (Julianto & Ratumanan, 2023) menegaskan bahwa pemanfaatan GenAI dalam pembelajaran bahasa untuk siswa SD memiliki potensi besar untuk meningkatkan efisiensi dan kualitas instrumen penilaian. Selain itu, integrasi dengan *Wordwall* membuat proses penilaian menjadi lebih menyenangkan, interaktif dan menarik bagi peserta didik (Suarmini & Nurjaya, 2023). Dari segi kepraktisan, persentase sebesar 87,5% (kategori sangat praktis) menunjukkan bahwa produk mudah digunakan oleh guru. Hal ini didukung oleh kelengkapan petunjuk penggunaan, rubrik penilaian yang sistematis, serta efisiensi waktu dan biaya dalam pelaksanaan penilaian.

Penilaian kemampuan komunikasi yang mencakup 7 indikator (Rahmah et al., 2024) memungkinkan guru melakukan observasi secara lebih objektif dan terstruktur. Hasil uji validitas dan reliabilitas yang sangat baik (r hitung tertinggi 0,876, *Cronbach's Alpha* 0,937) membuktikan bahwa instrumen penilaian ini konsisten dan tepat dalam mengukur kemampuan komunikasi peserta didik. Dengan demikian, instrumen penilaian yang dikembangkan dapat digunakan sebagai alternatif alat evaluasi pembelajaran Bahasa Indonesia di sekolah dasar.

Kesimpulan

Mengacu pada temuan dari serangkaian kegiatan penelitian dan pengembangan yang sudah dilaksanakan, maka kesimpulan yang bisa diambil adalah sebagaimana dibawah ini.

1. Kelayakan instrumen penilaian Bahasa Indonesia berbantuan *Generative Artificial Intelligence* (GenAI) dinyatakan sangat layak digunakan untuk mengukur kemampuan

komunikasi peserta didik kelas IV SD Negeri Winongo. Hal ini didukung bukti persentase kelayakan dari ahli materi sebesar 97,5%, ahli evaluasi sebesar 98,75%, serta rata-rata kelayakan sebesar 98,125% yang termasuk pada kategori "Sangat Layak".

2. Kepraktisan instrumen penilaian Bahasa Indonesia berbantuan GenAI dinyatakan sangat praktis digunakan dalam pembelajaran. Hasil penilaian kepraktisan oleh guru kelas IV B memperoleh persentase sebesar 87,5% dengan kategori "Sangat Praktis". Instrumen ini mudah digunakan, memiliki petunjuk yang jelas, materi, bahasa, dan tampilan yang sesuai dengan kebutuhan pembelajaran.
3. Seluruh 18 butir instrumen dinyatakan valid (r hitung 0,483 – 0,876 > r tabel 0,422) dan sangat reliabel (Cronbach's Alpha = 0,937). Dengan demikian, instrumen penilaian ini layak digunakan sebagai alat pengumpulan data dalam penelitian dan dapat dimanfaatkan oleh guru dalam proses penilaian kemampuan komunikasi peserta didik.

Ucapan Terima kasih

Penulis mengucapkan terima kasih kepada Universitas PGRI Yogyakarta yang telah memberikan dukungan pendanaan penelitian melalui dana LPPM Tahun 2025–2026. Bantuan pendanaan tersebut sangat membantu kelancaran pelaksanaan penelitian, mulai dari proses persiapan, pengumpulan data, hingga penyelesaian luaran penelitian

Daftar Pustaka

- Anjelina, N., & Tarmini, W. (2022). Keterampilan berbicara siswa sekolah dasar pada pembelajaran bahasa indonesia. *Jurnal Basicedu*, 6(4), 7327–7333.
- Arikunto, S. (2021). **Dasar-dasar evaluasi pendidikan edisi 3**. Bumi Aksara.
- Fajri, N., & Taufiqurrahman, T. (2017). Pengembangan media pembelajaran menggunakan model 4D pada mata pelajaran IPA. *Jurnal Ilmiah Pendidikan Teknik dan Kejuruan*, 10(2), 65–72.
- Hakim, L., dkk. (2021). *Metode penelitian dan pengembangan*. LPPM Universitas Brawijaya.
- Hasanah, H., dkk. (2022). Komunikasi sebagai sarana interaksi sosial. *Jurnal Komunikasi*, 8(2), 270–278.
- Hidayat, T., dkk. (2025). Pemanfaatan generative AI dalam dunia pendidikan. *Jurnal Teknologi Pendidikan*, 12(1), 45–55.

- Julianto, T. S., & Ratumanan, S. (2023). Pemanfaatan generatif AI dalam pembelajaran bahasa untuk peserta didik SD: Pendekatan inovatif dalam meningkatkan kemampuan menulis. *Jurnal Inovasi Pendidikan Dasar*, 7(2), 112–125.
- Nurmansyah, A., dkk. (2023). Permasalahan komunikasi yang kerap terjadi pada penyandang disabilitas. *Jurnal Komunikasi dan Sosial*, 2(2), 203–212.
- Rahmah, M. F., dkk. (2024). Pengaruh model role playing terhadap kemampuan komunikasi. *Elementary School Education*, 8(3), 32–40.
- Rahman, Yumriani, Munandar, S. A., Fitriani, A., & Karlina, Y. (2022). Pengertian pendidikan, ilmu pendidikan dan unsur-unsur pendidikan. *Jurnal Pendidikan*, 2(1), 1–8.
- Riscaputantri, A., & Wening, S. (2018). Pengembangan instrumen penilaian afektif siswa kelas IV sekolah dasar di Kabupaten Klaten. *Jurnal Penelitian dan Evaluasi Pendidikan*, 22(2), 231–242.
- Sriwijayanti, R. P., Qomariyah, R. S., & Nurma, I. F. (2020). Pengembangan media adobe flash berbasis PAKEM di sekolah dasar. *Pedagogy: Jurnal Ilmiah Ilmu Pendidikan*, 7(2), 92–105.
- Suarmini, N. K., & Nurjaya, I. G. (2023). Pemanfaatan media wordwall dalam pembelajaran interaktif bahasa indonesia. *Jurnal Ilmiah Pendidikan Profesi Guru*, 6(2), 328–339.