Integrative Perspectives of Social and Science Journal (IPSSJ)
ipssj.com ippsj2024@gmail.com E-ISSN: 3064-4011

EVALUASI PELAKSANAAN DAN ANALISIS HASIL TES KEMAMPUAN AKADEMIK PADA JENJANG SEKOLAH DASAR

Rifa Nugraha [1*]
[1] SDN Sambiroto, Kabupaten Rembang
*Corresponding email: rifanugraga@gmail.com

Abstrak. Penelitian ini bertujuan mengevaluasi pelaksanaan dan menganalisis hasil Tes Kemampuan Akademik 30 siswa kelas VI tahun ajaran 2025/2026. Pendekatan kuantitatif deskriptif dengan teknik dokumentasi digunakan. Data berupa nilai Matematika dan Bahasa Indonesia, surat keputusan panitia, daftar hadir, dan berita acara dianalisis melalui statistik deskriptif, uji normalitas, uji t berpasangan, korelasi Pearson, ukuran efek, dan analisis ketuntasan pada batas 60. Tes dilaksanakan 27–28 April 2026 dengan kehadiran 30 siswa atau 100%, pengawas dari sekolah lain, dan tanpa gangguan teknis, namun surat edaran tidak ditemukan dalam arsip. Hasil menunjukkan rata-rata Matematika (60,67) signifikan lebih tinggi daripada Bahasa Indonesia (43,21; $t = 6,29$; $p < 0,001$; $d = 1,15$) dengan korelasi positif sedang ($r = 0,537$; $p = 0,002$). Hanya 33,3% siswa mencapai batas ketuntasan dan ketuntasan Bahasa Indonesia hanya 6,7%. Simpulan penelitian ini adalah pelaksanaan tes tergolong baik pada dimensi input dan proses namun lemah pada dokumentasi, sementara penguasaan mata pelajaran tidak merata. Temuan ini berimplikasi pada perlunya program literasi terstruktur, perbaikan pengarsipan, dan pemanfaatan hasil tes sebagai umpan balik diagnostik.
Kata Kunci: asesmen; evaluasi; ketuntasan; sekolah dasar; tes kemampuan akademik

Abstrak. *This study aims to evaluate the implementation and analyze the results of the Academic Ability Test for 30 sixth-grade students of the 2025/2026 academic year. A quantitative descriptive approach with documentary technique was employed. Data consisted of test scores in Mathematics and Indonesian Language, committee decree, attendance list, and official report, analyzed through descriptive statistics, normality test, paired sample t-test, Pearson correlation, effect size, and mastery analysis at a threshold of 60. The test was held on 27–28 April 2026 with all 30 students present, cross-school invigilators, and no reported technical disruption, yet the circular letter was not found in the archive. The results show that the mean Mathematics score (60.67) was significantly higher than Indonesian Language (43.21; $t = 6.29$; $p < 0.001$; $d = 1.15$), with a moderate positive correlation ($r = 0.537$; $p = 0.002$). Only 33.3% of students reached the mastery threshold, and Indonesian Language mastery was merely 6.7%. The study concludes that the test was well executed in its input and process dimensions but weak in documentation, while subject mastery remained uneven. These findings imply the need for a structured literacy program, better archival management, and the use of test results as diagnostic feedback.*
Keywords: *academic ability test; assessment; evaluation; mastery; primary school*

PENDAHULUAN

Mutu pendidikan sangat ditentukan oleh kualitas asesmen yang digunakan untuk mengukur capaian belajar siswa. Asesmen tidak berfungsi semata sebagai penilaian akhir, tetapi juga sebagai sumber informasi untuk pengambilan keputusan pedagogis (Black & Wiliam, 1998; Shute, 2008). Melalui asesmen, guru memperoleh gambaran tentang kompetensi yang telah dikuasai dan yang masih perlu diperbaiki (Bennett, 2011; Hattie & Timperley, 2007). Evaluasi formatif yang berjalan baik akan memberikan umpan balik selama pembelajaran, sedangkan evaluasi sumatif mengukur capaian akhir periode (Felix & Pinem, 2025). Namun, praktik di lapangan sering kali evaluasi masih didominasi oleh tes tertulis sumatif, sehingga kesulitan belajar siswa baru terdeteksi di akhir (Wahyuni et al., 2026).

Salah satu bentuk asesmen yang diselenggarakan satuan pendidikan adalah Tes Kemampuan Akademik (TKA). Kedua mata pelajaran inti, Matematika dan Bahasa Indonesia, menjadi fondasi literasi numerasi dan literasi bahasa (Popham, 2011). Untuk memetakan kesiapan siswa, diperlukan instrumen asesmen diagnostik kognitif yang valid, salah satunya dapat berbantuan platform digital (Syahroh & Reinita, 2024). Pada jenjang sekolah dasar, penguasaan kedua kompetensi tersebut menjadi prasyarat krusial bagi kesiapan siswa memasuki pendidikan menengah.

Keberhasilan sebuah tes tidak hanya ditentukan oleh mutu instrumen, tetapi juga oleh mutu pelaksanaannya. Penyelenggaraan tes yang tertib, pengawasan yang objektif, dan

dokumentasi yang lengkap merupakan bagian dari jaminan mutu asesmen (AERA, APA, & NCME, 2014). Sebaliknya, kelemahan pada aspek dokumentasi dapat menyulitkan penelusuran dan evaluasi program pada periode berikutnya.

Secara ideal, hasil tes kemampuan akademik seharusnya mencerminkan pemerataan penguasaan kompetensi sekaligus menjadi dasar perbaikan pembelajaran. Kenyataannya, hasil tes di banyak satuan pendidikan masih menunjukkan ketimpangan antar mata pelajaran. Capaian Bahasa Indonesia sering tertinggal dibandingkan Matematika, sementara pemanfaatan hasil tes sebagai umpan balik pembelajaran belum berjalan optimal. Kesenjangan antara harapan dan kenyataan inilah yang menjadi fokus penelitian ini.

Penelitian terdahulu mengenai hasil belajar umumnya membahas satu mata pelajaran secara terpisah atau menilai program secara normatif, tanpa membandingkan pola capaian antar mata pelajaran pada kelompok siswa yang sama (Stufflebeam & Coryn, 2014). Kebaruan penelitian ini terletak pada integrasi evaluasi pelaksanaan berbasis dokumen dengan analisis komparatif, korelasional, dan ketuntasan dua mata pelajaran inti dalam satu kelompok siswa, sehingga diperoleh profil capaian yang lebih utuh dan dapat langsung diterjemahkan menjadi rekomendasi pembelajaran.

Tujuan penelitian ini adalah mengevaluasi pelaksanaan Tes Kemampuan Akademik dan menganalisis hasilnya pada mata pelajaran Matematika dan Bahasa Indonesia, mencakup gambaran deskriptif capaian, perbedaan capaian antar kedua mata pelajaran, hubungan antar keduanya, serta tingkat ketuntasannya, sebagai dasar perumusan rekomendasi peningkatan mutu pembelajaran.

LANDASAN TEORI

Asesmen dalam konteks pendidikan dipahami sebagai proses sistematis untuk mengumpulkan dan menafsirkan bukti belajar siswa. Messick (1995) menegaskan bahwa validitas asesmen tidak berhenti pada ketepatan skor, tetapi juga mencakup konsekuensi penggunaan hasilnya terhadap keputusan pembelajaran. Pandangan tersebut diperkuat Kane (2013) yang menempatkan validasi sebagai rangkaian argumen yang menghubungkan skor dengan tafsir dan tindakan yang didasarkan padanya. Oleh karena itu, hasil asesmen seharusnya menjadi umpan balik yang memandu perbaikan, bukan hanya angka yang dilaporkan (Shute, 2008).

Fungsi umpan balik dalam asesmen dijelaskan Hattie dan Timperley (2007) melalui tiga pertanyaan pokok, yaitu ke mana arah pembelajaran, seberapa jauh siswa telah mencapai tujuan, dan langkah apa yang perlu dilakukan selanjutnya. Black dan Wiliam (1998) serta Bennett (2011) menegaskan bahwa asesmen formatif yang dijalankan secara konsisten memberikan kontribusi lebih besar terhadap peningkatan hasil belajar dibandingkan asesmen yang semata-mata bersifat sumatif.

Tes Kemampuan Akademik merupakan instrumen untuk mengukur penguasaan kompetensi pada mata pelajaran inti. Popham (2011) menekankan bahwa interpretasi hasil tes perlu disertai pertimbangan tentang standar ketuntasan. Dalam Kurikulum Merdeka, standar ini diwujudkan melalui Kriteria Ketercapaian Tujuan Pembelajaran (KKTP). Penetapan KKTP yang tepat terbukti mampu meningkatkan ketuntasan klasikal siswa, seperti yang ditunjukkan pada perbaikan hasil belajar Bahasa Indonesia (Yani et al., 2026). Standar profesi asesmen juga menuntut agar prosedur administrasi, termasuk dokumentasi, dijalankan secara konsisten sebagai bagian dari jaminan mutu (AERA, APA, & NCME, 2014).

Evaluasi pelaksanaan program asesmen dapat didekati dengan model CIPP yang mencakup dimensi context, input, process, dan product (Stufflebeam & Coryn, 2014). Dimensi context menelaah dasar dan tujuan penyelenggaraan, dimensi input menelaah kesiapan sumber daya, dimensi process menelaah jalannya pelaksanaan, dan dimensi product menelaah hasil

yang diperoleh. Model ini dipilih karena mampu menghubungkan mutu penyelenggaraan dengan mutu hasil secara terpadu.

Secara statistik, perbandingan capaian antar mata pelajaran pada kelompok siswa yang sama dianalisis menggunakan uji berpasangan, sedangkan hubungan antar mata pelajaran dianalisis melalui korelasi (Field, 2018). Besarnya perbedaan tidak cukup dilaporkan melalui nilai p, tetapi perlu disertai ukuran efek agar maknanya praktis (Cohen, 1988). Dengan demikian, analisis hasil tes dalam penelitian ini diarahkan tidak hanya pada pengujian signifikansi, tetapi juga pada penilaian besarnya perbedaan dan proporsi varians bersama antar mata pelajaran.

Beberapa penelitian terdahulu menunjukkan bahwa capaian Matematika dan Bahasa Indonesia cenderung berkorelasi positif karena keduanya menuntut kemampuan penalaran dan pemahaman bacaan. Namun, pola tersebut tidak selalu diikuti tingkat ketuntasan yang merata, terutama pada mata pelajaran yang menuntut penguasaan literasi kompleks (Popham, 2011). Berdasarkan kajian tersebut, penelitian ini mengajukan dua hipotesis, yaitu terdapat perbedaan capaian yang signifikan antara Matematika dan Bahasa Indonesia, dan terdapat hubungan positif yang signifikan antara keduanya.

METODE

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan desain ex post facto, yaitu analisis terhadap data nilai yang telah tersedia tanpa memberikan perlakuan kepada subjek, serta dilengkapi studi dokumentasi untuk menelaah pelaksanaan tes. Subjek penelitian adalah 30 siswa kelas VI SDN Sambiroto tahun ajaran 2025/2026 yang seluruhnya mengikuti Tes Kemampuan Akademik pada tanggal 27 dan 28 April 2026. Data dikumpulkan melalui teknik dokumentasi terhadap nilai Matematika dan Bahasa Indonesia, surat keputusan panitia penyelenggara, daftar hadir peserta, dan berita acara pelaksanaan, sedangkan NA TKA merupakan rata-rata kedua nilai tersebut sehingga tidak dianalisis sebagai variabel terpisah untuk menghindari kolinearitas sempurna. Analisis data kuantitatif dilakukan dengan statistik deskriptif (mean, median, modus, standar deviasi, nilai minimum dan maksimum), uji normalitas Shapiro-Wilk, uji t berpasangan sebagai uji utama dengan uji Wilcoxon signed-rank sebagai pembanding nonparametrik, korelasi Pearson, penghitungan ukuran efek Cohen's d dan koefisien determinasi, kategorisasi mean ± standar deviasi, serta analisis ketuntasan dengan batas nilai 60, sedangkan data dokumen dianalisis secara deskriptif dengan kerangka CIPP yang mencakup dimensi context, input, process, dan product. Seluruh analisis kuantitatif dilakukan pada taraf signifikansi 5%, dan identitas siswa dianonimkan dengan kode S1–S30 untuk menjaga etika penelitian.

HASIL DAN PEMBAHASAN

Hasil telaah dokumen menunjukkan bahwa Tes Kemampuan Akademik di SDN Sambiroto dilaksanakan pada tanggal 27 dan 28 April 2026 dengan Kepala Sekolah, Moh. Sakroni, S.Pd., M.A.P., bertindak sebagai penanggung jawab sebagaimana tercantum dalam surat keputusan panitia. Susunan kepanitiaan mencakup proktor Rifa Nugraha dan teknisi Meiriska Dwiariyanti, sedangkan pengawasan pelaksanaan dilakukan oleh guru dari sekolah lain yang bertugas pada dua hari berbeda. Daftar hadir menunjukkan seluruh 30 siswa peserta hadir tanpa absen, sehingga tingkat kehadiran mencapai 100%. Berita acara pelaksanaan menyatakan bahwa seluruh rangkaian kegiatan terlaksana dengan lancar tanpa gangguan teknis maupun kendala berarti.

Berdasarkan kerangka CIPP, dimensi input dan process pelaksanaan tes tergolong baik. Penetapan penanggung jawab, proktor, dan teknisi menunjukkan kesiapan sumber daya manusia yang jelas, sementara penggunaan pengawas dari sekolah lain pada dua hari yang berbeda merupakan praktik yang mendukung objektivitas penilaian karena mengurangi potensi

konflik kepentingan antara pengawas dan peserta. Kehadiran peserta yang mencapai 100% juga mencerminkan kesiapan siswa dan efektivitas koordinasi sekolah. Pada dimensi process, pernyataan berita acara bahwa kegiatan berlangsung tanpa gangguan mengindikasikan bahwa prosedur pelaksanaan berjalan sesuai rencana.

Meskipun demikian, ditemukan kelemahan pada dimensi context, yaitu surat edaran pelaksanaan tidak ditemukan dalam arsip sekolah. Kelemahan ini bersifat administratif dan tidak memengaruhi jalannya tes, tetapi berdampak pada kelengkapan bukti dokumentasi. Dalam kerangka jaminan mutu asesmen, kelengkapan dokumen merupakan bagian yang menopang akuntabilitas penyelenggaraan, sehingga ketiadaan surat edaran menyulitkan penelusuran dasar pelaksanaan pada periode berikutnya (AERA, APA, & NCME, 2014).

Hasil analisis deskriptif terhadap 30 siswa disajikan pada Tabel 1. Rata-rata nilai Matematika sebesar 60,67 lebih tinggi daripada Bahasa Indonesia sebesar 43,21, sedangkan NA TKA sebagai rata-rata kedua mata pelajaran sebesar 51,94. Sebaran nilai Matematika lebih heterogen dengan standar deviasi 17,99 dan rentang 30,00 sampai 90,00, sedangkan Bahasa Indonesia memiliki standar deviasi 10,31 dengan rentang 16,67 sampai 60,00. Median Matematika sebesar 61,67 dan median Bahasa Indonesia sebesar 43,33 menunjukkan bahwa separuh siswa memperoleh nilai di bawah angka tersebut pada masing-masing mata pelajaran.

Tabel 1. Statistik Deskriptif Nilai Tes Kemampuan Akademik

Statistik	Matematika	Bahasa Indonesia	NA TKA
N	30	30	30
Mean	60,67	43,21	51,94
Median	61,67	43,33	53,33
Modus	76,67	43,33	35 dan 45
Standar deviasi	17,99	10,31	12,54
Varians	323,69	106,26	157,31
Minimum	30,00	16,67	31,67
Maksimum	90,00	60,00	73,34

Sebelum pengujian hipotesis, dilakukan uji asumsi normalitas terhadap skor kedua mata pelajaran menggunakan uji Shapiro-Wilk. Hasil menunjukkan bahwa data berdistribusi normal, dengan nilai signifikansi Matematika sebesar 0,265 (W = 0,957) dan Bahasa Indonesia sebesar 0,366 (W = 0,963). Karena asumsi normalitas terpenuhi, uji t berpasangan digunakan sebagai uji utama.

Uji t berpasangan menunjukkan selisih rata-rata sebesar 17,46 dengan standar deviasi selisih 15,19; t = 6,29; df = 29; p < 0,001, sehingga hipotesis pertama diterima. Interval kepercayaan 95% bagi selisih rata-rata berada pada rentang 11,79 sampai 23,13. Ukuran efek Cohen's d sebesar 1,15 tergolong besar. Arah perbedaan ini konsisten pada tingkat individu, yaitu 24 siswa memperoleh nilai Matematika lebih tinggi, dua siswa memperoleh nilai sama, dan hanya empat siswa memperoleh nilai Bahasa Indonesia lebih tinggi. Sebagai uji ketahanan, uji Wilcoxon signed-rank (n efektif = 28) menghasilkan W = 17; Z = -4,24; p < 0,001, yang mengonfirmasi arah dan signifikansi simpulan yang sama.

Tabel 2. Hasil Uji t Berpasangan dan Ukuran Efek

Perbandingan	Selisih Mean	SD Selisih	95% CI	t	df	p	Cohen's d

Matematika – Bahasa Indonesia	17,46	15,19	11,79–23,13	6,29	29	< 0,001	1,15

Korelasi Pearson antara Matematika dan Bahasa Indonesia sebesar r = 0,537 dengan p = 0,002, sehingga hipotesis kedua diterima. Koefisien determinasi sebesar 0,288 menunjukkan bahwa sekitar 28,8% varians capaian Bahasa Indonesia dapat dijelaskan oleh capaian Matematika, sedangkan sisanya ditentukan faktor lain. Hubungan positif sedang ini mengindikasikan bahwa siswa yang tinggi pada Matematika cenderung tinggi pula pada Bahasa Indonesia, tetapi hubungan tersebut tidak sempurna sehingga masih terdapat siswa dengan profil capaian yang tidak seimbang.

Berdasarkan kategorisasi mean ± standar deviasi, sebagian besar siswa berada pada kategori sedang pada ketiga variabel, sebagaimana disajikan pada Tabel 3. Namun, analisis ketuntasan dengan batas nilai 60 memperlihatkan pola yang berbeda: sebanyak 16 siswa atau 53,3% tuntas pada Matematika, tetapi hanya dua siswa atau 6,7% yang tuntas pada Bahasa Indonesia, dan 10 siswa atau 33,3% tuntas pada NA TKA. Selain itu, tujuh siswa atau 23,3% memperoleh nilai Bahasa Indonesia di bawah 40, dan 12 siswa atau 40% memiliki NA TKA di bawah 50.

Tabel 3. Kategorisasi Nilai dan Tingkat Ketuntasan

Variabel	Tinggi	Sedang	Rendah	Tuntas (≥60)	Persentase Tuntas
Matematika	5	20	5	16	53,3%
Bahasa Indonesia	5	20	5	2	6,7%
NA TKA	5	18	7	10	33,3%

Temuan utama penelitian ini adalah ketidakmerataan penguasaan antar mata pelajaran. Capaian Matematika secara signifikan lebih tinggi daripada Bahasa Indonesia dengan efek besar, sementara ketuntasan Bahasa Indonesia hanya 6,7%. Pola ini menunjukkan bahwa persoalan bukan terletak pada kemampuan akademik siswa secara umum, melainkan pada penguasaan kompetensi literasi bahasa yang belum optimal. Rendahnya literasi membaca sering kali dipicu oleh faktor internal dan eksternal, seperti minat baca siswa yang rendah dan baru membaca ketika diperintah guru (Ihwan et al., 2026; Putri & Elvina, 2025). Di sisi lain, penguatan literasi numerasi juga masih menjadi tantangan, di mana kemampuan numerasi siswa SD secara umum masih tergolong rendah (Erlyana et al., 2023; Suryaningsih et al., 2023), padahal kemampuan ini memiliki hubungan positif yang signifikan dengan hasil belajar Matematika siswa (Rosmalah et al., 2023).

Korelasi positif sedang sebesar 0,537 menunjukkan bahwa kedua mata pelajaran berbagi sebagian kemampuan dasar, terutama penalaran dan pemahaman bacaan, namun masing-masing tetap menuntut kompetensi spesifik. Dengan demikian, peningkatan capaian Bahasa Indonesia tidak dapat sepenuhnya diharapkan sebagai efek samping dari perbaikan Matematika, melainkan memerlukan intervensi yang dirancang khusus (Bennett, 2011; Popham, 2011).

Dari sisi pelaksanaan, hasil telaah dokumen menunjukkan bahwa dimensi input dan process telah berjalan baik, tercermin dari kejelasan struktur kepanitiaan, penggunaan pengawas silang antar sekolah, kehadiran peserta 100%, serta pernyataan berita acara bahwa kegiatan berlangsung tanpa gangguan. Pengawasan silang merupakan praktik yang sejalan dengan tuntutan objektivitas asesmen, karena memisahkan peran pengawas dari kepentingan penilaian peserta didiknya sendiri (AERA, APA, & NCME, 2014). Namun, ketiadaan surat edaran dalam arsip menunjukkan bahwa dimensi context belum terdokumentasi dengan tertib. Temuan ini penting karena dokumentasi bukan sekadar kelengkapan administratif, melainkan

bagian dari akuntabilitas yang memungkinkan evaluasi program dilakukan secara berkelanjutan (Stufflebeam & Coryn, 2014).

Rendahnya ketuntasan Bahasa Indonesia dapat ditafsirkan sebagai indikasi lemahnya pemanfaatan hasil asesmen sebagai umpan balik pembelajaran. Sebagaimana ditegaskan Black dan Wiliam (1998) serta Hattie dan Timperley (2007), asesmen baru berdampak pada hasil belajar apabila hasilnya dikembalikan kepada siswa dalam bentuk informasi yang jelas mengenai kekuatan, kelemahan, dan langkah perbaikan. Apabila nilai tes hanya berhenti sebagai laporan administratif, kesenjangan capaian antar mata pelajaran cenderung bertahan (Shute, 2008).

Secara metodologis, penggunaan ukuran efek dan interval kepercayaan dalam penelitian ini memperkuat interpretasi hasil. Signifikansi statistik semata tidak menunjukkan besarnya persoalan, sedangkan nilai Cohen's d sebesar 1,15 menegaskan bahwa perbedaan capaian antar kedua mata pelajaran bersifat substansial dan relevan bagi pengambilan keputusan pembelajaran (Cohen, 1988; Field, 2018). Selain itu, konsistensi arah perbedaan pada tingkat individu, yaitu 24 dari 30 siswa, menunjukkan bahwa temuan ini bukan disebabkan oleh sejumlah kecil nilai ekstrem.

Temuan penelitian ini sejalan dengan kajian terdahulu yang menempatkan asesmen sebagai bagian dari proses pembelajaran, bukan sebagai tahap akhir yang terpisah (Stufflebeam & Coryn, 2014). Kebaruan penelitian ini terletak pada integrasi evaluasi pelaksanaan berbasis dokumen dengan analisis komparatif, korelasional, dan ketuntasan dalam satu kelompok siswa, sehingga diperoleh profil capaian yang lebih utuh. Implikasi praktisnya, satuan pendidikan perlu memprioritaskan program penguatan literasi bahasa, menjadikan hasil tes sebagai dasar diagnosis awal pembelajaran remedial, menetapkan target ketuntasan yang terukur, serta menertibkan pengarsipan dokumen pelaksanaan asesmen.

Keterbatasan penelitian ini adalah cakupan subjek yang terbatas pada satu satuan pendidikan dan satu tahun ajaran, sehingga hasilnya tidak dapat digeneralisasikan secara luas. Analisis pelaksanaan juga terbatas pada bukti dokumen tanpa wawancara mendalam dengan panitia dan peserta, sehingga aspek pengalaman subjektif pelaksana tidak tergali. Penelitian ini belum mempertimbangkan faktor latar belakang siswa, kualitas pembelajaran di kelas, maupun karakteristik butir soal. Oleh karena itu, penelitian selanjutnya disarankan melibatkan sampel yang lebih besar, melengkapi data wawancara, dan mengaitkan capaian tes dengan faktor kontekstual siswa dan pembelajaran.

SIMPULAN

Berdasarkan hasil dan pembahasan, dapat disimpulkan bahwa pelaksanaan Tes Kemampuan Akademik di SDN Sambiroto pada tanggal 27 dan 28 April 2026 tergolong baik pada dimensi input dan process, tercermin dari kejelasan struktur kepanitiaan dengan penanggung jawab kepala sekolah, keberadaan proktor dan teknisi, penggunaan pengawas dari sekolah lain pada dua hari berbeda, kehadiran peserta 100%, serta berita acara yang menyatakan kegiatan berlangsung lancar tanpa gangguan, namun masih lemah pada dimensi context karena surat edaran pelaksanaan tidak ditemukan dalam arsip. Dari sisi hasil, capaian Matematika dengan rata-rata 60,67 signifikan lebih tinggi daripada Bahasa Indonesia dengan rata-rata 43,21, ditunjukkan oleh $t = 6,29$ dan $p < 0,001$ dengan ukuran efek besar sebesar 1,15, serta keduanya berkorelasi positif sedang sebesar 0,537. Meskipun sebagian besar siswa berada pada kategori sedang, tingkat ketuntasan tidak merata, yaitu 53,3% pada Matematika, 33,3% pada NA TKA, dan hanya 6,7% pada Bahasa Indonesia. Ketidakmerataan ini menunjukkan bahwa persoalan utama bukan pada kemampuan akademik secara umum, melainkan pada penguasaan literasi bahasa yang perlu diprioritaskan. Oleh karena itu, satuan pendidikan disarankan menyusun program penguatan literasi yang terstruktur, memanfaatkan hasil tes sebagai umpan balik diagnostik bagi siswa dan guru, menertibkan pengarsipan dokumen

pelaksanaan asesmen, serta melakukan pemantauan ketuntasan secara berkala. Penelitian selanjutnya disarankan memperluas cakupan sampel, melengkapi data wawancara, dan mengaitkan hasil tes dengan faktor latar belakang siswa maupun proses pembelajaran.

DAFTAR PUSTAKA

AERA, APA, & NCME. (2014). Standards for educational and psychological testing. American Educational Research Association.

Bennett, R. E. (2011). Formative assessment: A critical review. Assessment in Education: Principles, Policy & Practice, 18(1), 5–25. https://doi.org/10.1080/0969594X.2010.513678

Black, P., & Wiliam, D. (1998). Assessment and classroom learning. Assessment in Education: Principles, Policy & Practice, 5(1), 7–74. https://doi.org/10.1080/0969595980050102

Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.). Lawrence Erlbaum Associates.

Erlyana, R., Nugraheni, P., & Yuzianah, D. (2023). Deskripsi kemampuan literasi numerasi siswa SD. Jurnal Theorems (The Original Research of Mathematics), 7(2), 193–200.

Felix, F., & Pinem, I. (2025). Peran evaluasi formatif dan sumatif dalam meningkatkan efektivitas pembelajaran di sekolah dasar. Jurnal Pendidikan Tambusai, 9(2), 1234–1245. https://doi.org/10.31004/jptam.v9i2.29831

Field, A. (2018). Discovering statistics using IBM SPSS Statistics (5th ed.). SAGE.

Hattie, J., & Timperley, H. (2007). The power of feedback. Review of Educational Research, 77(1), 81–112. https://doi.org/10.3102/003465430298487

Ihwan, I., Dantes, N., & Parmiti, D. P. (2026). Analisis faktor penyebab rendahnya kemampuan literasi membaca siswa kelas V SDN 1 Suntalangu. SENTRI: Jurnal Riset Ilmiah, 5(1), 992–1002. https://doi.org/10.55681/sentri.v5i1.5502

Kane, M. T. (2013). Validating the interpretations and uses of test scores. Journal of Educational Measurement, 50(1), 1–73. https://doi.org/10.1111/jedm.12000

Messick, S. (1995). Validity of psychological assessment: Validation of inferences from persons' responses and performances as scientific inquiry into score meaning. American Psychologist, 50(9), 741–749. https://doi.org/10.1037/0003-066X.50.9.741

Popham, W. J. (2011). Classroom assessment: What teachers need to know (6th ed.). Pearson.

Putri, D. O., & Elvina, E. (2025). Analisis kemampuan literasi membaca siswa di SDN 29 Rantau Batu Pasar. Jurnal Kepemimpinan dan Pengurusan Sekolah, 10(3), 1206–1211. https://doi.org/10.34125/jkps.v10i3.850

Rosmalah, Sudarto, & Hur'ainun, K. (2023). Hubungan antara kemampuan literasi numerasi dengan hasil belajar matematika siswa kelas tinggi. Jurnal Pendidikan & Pembelajaran Sekolah Dasar, 2(4), 334–342.

Shute, V. J. (2008). Focus on formative feedback. Review of Educational Research, 78(1), 153–189. https://doi.org/10.3102/0034654307313795

Stufflebeam, D. L., & Coryn, C. L. S. (2014). Evaluation theory, models, and applications (2nd ed.). Jossey-Bass.

Suryaningsih, T., Edwita, & Yarmi, G. (2023). Pengaruh media realia berbasis realistic mathematic education terhadap kemampuan literasi numerasi. AULADUNA: Jurnal Pendidikan Dasar Islam, 10(2), 193–204. https://doi.org/10.24252/auladuna.v10i2a6.2023

Syahroh, K., & Reinita, R. (2024). Validitas asesmen diagnostik kognitif berbantuan Quizizz pada pembelajaran pendidikan pancasila di sekolah dasar. Jurnal Inovasi Pendidikan dan Pembelajaran Sekolah Dasar, 8(1), 331–344. https://doi.org/10.24036/jippsd.v8i1.128705

Wahyuni, S., Nurindah, N., Sari, A. F., Hasanah, S. N., Rahantan, F. A. R., & Hamid, R. J. (2026). Pelaksanaan evaluasi formatif pembelajaran: Studi deskriptif di UPT SPF SD Inpres Monginsidi. Tunas Cendekia: Jurnal Program Studi Pendidikan Anak Usia Dini, 9(1), 81–90. https://doi.org/10.24256/tunascendekia.v9i1.12240

Yani, J. F., Rahmawati, E., Putri, S. A., Yuda, R. A., Saputra, A., & Hastuti. (2026). Pemanfaatan video infografis animasi untuk meningkatkan pemahaman dan hasil belajar teks prosedur pada mata pelajaran Bahasa Indonesia siswa kelas V SDN Sidodadi Asri 2025/2026. Lingua Rima: Jurnal Pendidikan Bahasa dan Sastra Indonesia, 15(1), 143–151. https://jurnal.umt.ac.id/index.php/lgrm/article/view/15875