

Analisis Kualitas Butir Soal Berbasis AKM dalam Mengukur Kemampuan Literasi dan Numerasi Siswa SMP pada Implementasi Kurikulum Merdeka.

Syahrul¹, Fathullah Wajdi², dan Suryadi Ishak^{3*}

¹Program Studi Penelitian dan Evaluasi Pendidikan, Program Pascasarjana, Universitas Negeri Makassar

²Program Studi S3 Ilmu Pendidikan Bahasa, Universitas Negeri Makassar

Email Corresponding Author : suryadi.ishak@unm.ac.id

Info Artikel

Article history:

Kirim: 25/12/2025

Perbaikan: 31/12/2025

Terima: 03/01/2026

Publikasi: 06/01/2026

Kata-kata kunci:

Analisis Butir Soal;

Asesmen Kompetensi Minimum;

Literasi; Numerasi;

Kurikulum Merdeka.

ABSTRAK

Implementasi Kurikulum Merdeka menempatkan literasi dan numerasi sebagai kompetensi fondasional sehingga diperlukan instrumen berkarakteristik Asesmen Kompetensi Minimum (AKM). Penelitian ini bertujuan menganalisis kualitas butir soal berbasis AKM dalam mengukur kemampuan literasi dan numerasi siswa SMP serta memetakan profil penguasaan berdasarkan proses kognitif. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan populasi siswa SMP di Kota Polman dan sampel 246 siswa melalui purposive sampling. Instrumen berupa tes objektif pilihan ganda 18 butir berbasis AKM dan lembar validasi ahli. Data dianalisis menggunakan koefisien V Aiken untuk validitas isi dan teori tes klasik (CTT) untuk tingkat kesukaran, daya beda, serta keberfungsian pengecoh. Hasil penelitian menunjukkan koefisien V Aiken 0,60–0,85 (rerata 0,76), sehingga seluruh butir valid secara isi. Tingkat kesukaran 0,34–0,79, terdiri atas 15 butir sedang dan 3 butir mudah. Daya beda 0,394–0,817, dengan 13 butir sangat baik, 4 baik, dan 1 cukup, tanpa indeks negatif. Seluruh pengecoh berfungsi dengan proporsi pemilih di atas 5%. Profil numerasi menunjukkan capaian tertinggi pada penerapan (59%), diikuti pemahaman (52%), dan penalaran (50%). Butir yang dikembangkan berkualitas memadai dan layak digunakan sebagai alat ukur literasi dan numerasi SMP serta mendukung asesmen diagnostik dan penguatan pembelajaran berorientasi penalaran dan masalah non-rutin dalam Kurikulum Merdeka bagi guru dalam pelaksanaan pembelajaran dan asesmen.

1. PENDAHULUAN

Transformasi pendidikan nasional melalui kebijakan Kurikulum Merdeka menandai perubahan paradigma pembelajaran dari berorientasi konten menuju penguatan kompetensi esensial peserta didik (Kemendikbudristek, 2022). Pada jenjang Sekolah Menengah Pertama (SMP), penguatan literasi dan numerasi menjadi fokus utama karena kedua kompetensi ini merupakan fondasi bagi penguasaan ilmu pengetahuan, keterampilan berpikir kritis, dan kemampuan adaptif terhadap perkembangan global. Literasi dan numerasi tidak hanya dimaknai sebagai kemampuan membaca teks dan melakukan operasi hitung, tetapi sebagai kemampuan memahami informasi, menganalisis secara logis, serta memecahkan masalah dalam berbagai konteks kehidupan nyata (Abidin, 2018).

Kebijakan ini selaras dengan kerangka evaluasi internasional seperti Programme for International Student Assessment (PISA), yang menekankan kemampuan peserta didik dalam mengaplikasikan pengetahuan pada situasi autentik dan kontekstual (OECD, 2019). Hasil PISA dalam beberapa siklus menunjukkan bahwa kemampuan literasi dan numerasi peserta didik Indonesia masih memerlukan penguatan yang sistematis dan berkelanjutan. Oleh karena itu, reformasi kurikulum dan sistem asesmen menjadi langkah strategis untuk meningkatkan mutu pendidikan secara menyeluruh.

Sebagai bagian dari kebijakan Merdeka Belajar, pemerintah melaksanakan Asesmen Nasional yang mencakup Asesmen Kompetensi Minimum (AKM) sebagai instrumen pengukuran kompetensi literasi membaca dan numerasi (Kemendikbudristek, 2021). AKM dirancang untuk menilai kemampuan memahami, menalar, serta menggunakan informasi dalam penyelesaian masalah, bukan sekadar mengukur penguasaan materi pelajaran. Karakteristik soal berbasis AKM meliputi penggunaan stimulus kontekstual, variasi format soal (pilihan ganda kompleks, isian singkat, uraian), serta tuntutan penalaran tingkat tinggi. Namun demikian, implementasi asesmen berbasis AKM di tingkat sekolah belum sepenuhnya optimal. Dalam praktik penyusunan soal, masih ditemukan butir soal yang berorientasi pada hafalan, prosedur rutin, dan penguasaan fakta semata (Widana, 2017). Kondisi ini menunjukkan adanya kesenjangan antara kebijakan evaluasi nasional dengan praktik evaluasi pembelajaran di SMP. Jika instrumen asesmen tidak mencerminkan karakteristik AKM, maka hasil pengukuran kemampuan literasi dan numerasi siswa menjadi kurang akurat.

Kualitas butir soal memiliki peran krusial dalam menentukan validitas dan reliabilitas hasil asesmen. Instrumen yang baik harus memenuhi kriteria validitas isi dan empiris, memiliki reliabilitas yang memadai, tingkat kesukaran yang proporsional, daya pembeda yang efektif, serta distraktor yang berfungsi optimal (Arikunto, 2018). Analisis butir soal menjadi pendekatan ilmiah yang dapat digunakan untuk mengevaluasi karakteristik setiap item berdasarkan data respons siswa. Pendekatan analisis butir soal berbasis teori tes klasik (Classical Test Theory/CTT) banyak digunakan dalam penelitian evaluasi pendidikan karena mampu memberikan informasi statistik yang komprehensif mengenai kualitas instrumen (Crocker & Algina, 2008). Melalui analisis ini, dapat diidentifikasi butir soal yang layak digunakan, perlu direvisi, atau harus dieliminasi. Dengan demikian, instrumen asesmen yang dihasilkan benar-benar mampu mengukur kompetensi literasi dan numerasi secara tepat.

Sejalan dengan tuntutan implementasi Kurikulum Merdeka, diperlukan kajian empiris mengenai kualitas butir soal berbasis AKM di jenjang SMP. Penelitian ini menjadi penting karena tidak hanya mengevaluasi aspek statistik instrumen, tetapi juga menilai kesesuaiannya dengan karakteristik desain AKM dan tujuan penguatan kompetensi abad ke-21 (Trilling & Fadel, 2009). Hasil penelitian diharapkan memberikan kontribusi nyata dalam pengembangan sistem asesmen yang lebih berkualitas, akurat, dan relevan dengan kebijakan pendidikan nasional. Penelitian mengenai analisis butir soal dan asesmen berbasis HOTS telah banyak dilakukan (Widana, 2017; Brookhart, 2010). Namun, penelitian yang secara khusus menganalisis kualitas butir soal berbasis AKM pada jenjang SMP dalam konteks implementasi

Kurikulum Merdeka masih relatif terbatas. Sebagian penelitian sebelumnya lebih berfokus pada evaluasi umum instrumen atau hasil capaian AKM secara nasional.

Kebaruan penelitian ini terletak pada integrasi antara analisis statistik kualitas butir soal dan evaluasi kesesuaian karakteristik desain AKM. Dengan demikian, penelitian ini tidak hanya menghasilkan data empiris mengenai kualitas instrumen, tetapi juga memberikan arah pengembangan asesmen berbasis kompetensi minimum secara kontekstual.

2. METODE PENELITIAN

Pendekatan Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif dengan metode analisis butir soal (*item analysis*) untuk menganalisis kualitas instrumen asesmen berbasis AKM dalam mengukur kemampuan literasi dan numerasi siswa SMP. Pendekatan kuantitatif dipilih karena penelitian ini berorientasi pada pengukuran empiris karakteristik butir soal berdasarkan data respons siswa. Desain penelitian bersifat evaluatif-analitik, yaitu mengevaluasi instrumen yang telah digunakan dalam implementasi Kurikulum Merdeka untuk mengetahui tingkat kelayakan dan efektivitasnya sebagai alat ukur kompetensi minimum.

Analisis dilakukan berdasarkan teori tes klasik (Classical Test Theory/CTT) yang menekankan hubungan antara skor individu dan karakteristik item (Crocker & Algina, 2008). Parameter yang dianalisis meliputi validitas item, reliabilitas tes, tingkat kesukaran, daya pembeda, dan efektivitas distraktor. Pendekatan ini dipilih karena relevan untuk konteks evaluasi pendidikan di sekolah dan memberikan informasi statistik yang komprehensif mengenai kualitas instrumen (Ebel & Frisbie, 1991). Penelitian dirancang untuk dilaksanakan selama satu tahun anggaran dan disesuaikan dengan Rencana Anggaran Biaya (RAB) yang mencakup tahap persiapan, pengumpulan data, analisis data, penyusunan laporan, serta publikasi ilmiah

Populasi dan Sampel

Populasi penelitian adalah seluruh siswa SMP yang telah menerapkan Kurikulum Merdeka pada wilayah yang menjadi lokasi penelitian. Sampel ditentukan menggunakan teknik purposive sampling dengan mempertimbangkan kesiapan sekolah dalam mengimplementasikan asesmen berbasis AKM. Sampel penelitian terdiri dari siswa kelas VIII (atau kelas yang relevan) pada sekolah mitra yang telah melaksanakan pembelajaran berbasis Kurikulum Merdeka. Pemilihan kelas VIII didasarkan pada pertimbangan bahwa pada tingkat ini siswa telah memiliki pengalaman belajar yang cukup dalam literasi dan numerasi serta telah terpapar karakteristik asesmen berbasis AKM. Jumlah sampel disesuaikan dengan standar analisis butir soal agar memenuhi ketentuan minimal jumlah responden untuk memperoleh parameter statistik yang stabil (Haladyna, 2004).

Instrumen Penelitian

Data penelitian dikumpulkan melalui teknik tes, dokumentasi, dan wawancara terbatas. Teknik tes digunakan untuk memperoleh data kuantitatif berupa skor respons siswa terhadap

butir soal berbasis AKM. Dokumentasi dilakukan untuk mengumpulkan perangkat soal dan dokumen pendukung pembelajaran. Wawancara terbatas dengan guru dilakukan untuk memperoleh informasi kontekstual mengenai proses penyusunan soal dan implementasi asesmen. Penggunaan beberapa teknik pengumpulan data ini bertujuan meningkatkan validitas penelitian melalui triangulasi sumber informasi.

Teknik Analisis Data

Analisis data dilakukan menggunakan pendekatan teori tes klasik. Validitas item dihitung menggunakan korelasi antara skor item dan skor total. Reliabilitas tes dihitung menggunakan koefisien Alpha Cronbach untuk mengetahui konsistensi internal instrumen. Tingkat kesukaran dihitung berdasarkan proporsi siswa yang menjawab benar, sedangkan daya pembeda dianalisis dengan membandingkan kelompok atas dan bawah. Efektivitas distraktor dianalisis melalui distribusi pilihan jawaban siswa. Hasil analisis diinterpretasikan berdasarkan kriteria baku evaluasi pendidikan untuk menentukan kualitas empiris setiap butir soal. Seluruh proses analisis dilakukan menggunakan perangkat lunak statistik guna meningkatkan akurasi perhitungan.

Prosedur Penelitian

Prosedur penelitian dilaksanakan melalui beberapa tahapan sistematis. Tahap pertama adalah studi pendahuluan yang meliputi identifikasi instrumen soal berbasis AKM yang digunakan guru serta koordinasi dengan pihak sekolah. Tahap ini bertujuan memastikan kesesuaian instrumen dengan konteks pembelajaran. Tahap kedua adalah pelaksanaan tes kepada siswa menggunakan instrumen yang telah dikumpulkan. Tes dilaksanakan sesuai prosedur standar untuk menjaga objektivitas dan konsistensi pengukuran. Hasil tes berupa lembar jawaban siswa menjadi data utama penelitian. Tahap ketiga adalah pengolahan dan analisis data butir soal. Setiap butir dianalisis berdasarkan parameter statistik yang telah ditetapkan. Hasil analisis kemudian dikategorikan ke dalam kriteria layak, revisi, atau gugur. Tahap keempat adalah interpretasi hasil dan penyusunan rekomendasi pengembangan instrumen asesmen berbasis AKM. Tahap ini juga mencakup diskusi bersama guru untuk memberikan umpan balik konstruktif. Tahap kelima adalah penyusunan laporan akhir dan publikasi hasil penelitian. Publikasi menjadi bagian penting sebagai bentuk kontribusi ilmiah terhadap pengembangan evaluasi pendidikan.

3. HASIL DAN PEMBAHASAN

Hasil Penelitian

Uji Validitas isi instrumen

Penilaian validitas isi dilakukan oleh lima validator, meliputi empat akademisi dengan kepakaran di bidang literasi, numerasi, dan asesmen, serta satu praktisi guru pada bidang yang sama. Koefisien validitas isi setiap butir dihitung menggunakan formula V Aiken dan hasilnya ditampilkan pada [Tabel 1](#).

Tabel 1. Validitas isi instrumen formula V Aiken

No Item	V Aiken	No Item	V Aiken
1	0,8	11	0,6
2	0,85	12	0,75
3	0,65	13	0,75
4	0,85	14	0,8
5	0,75	15	0,8
6	0,75	16	0,8
7	0,7	17	0,75
8	0,7	18	0,85
9	0,75		
10	0,7		

Hasil analisis pada Tabel 1 menunjukkan bahwa koefisien V Aiken dari 18 butir soal berbasis AKM berada pada rentang 0,60 hingga 0,85 dengan rerata sebesar 0,76. Koefisien tertinggi diperoleh pada butir nomor 2, 4, dan 18 ($V = 0,85$), sedangkan koefisien terendah terdapat pada butir nomor 11 ($V = 0,60$). Berdasarkan kriteria penafsiran V Aiken, sebanyak 7 butir (38,9%) berada pada kategori validitas tinggi ($V \geq 0,80$), yaitu butir nomor 1, 2, 4, 14, 15, 16, dan 18, sementara 11 butir lainnya (61,1%) termasuk kategori sedang dengan koefisien antara 0,60 dan 0,79. Tidak ditemukan butir yang berada pada kategori rendah, sehingga seluruh butir dinyatakan memiliki bukti validitas isi yang memadai dan layak digunakan untuk mengukur kemampuan literasi dan numerasi siswa SMP, dengan catatan bahwa butir-butir berkategori sedang terutama butir nomor 11 dan 3 masih memerlukan perbaikan pada aspek kesesuaian stimulus, kejelasan bahasa, dan ketepatan indikator yang diukur.

Uji Reliabilitas

Pembuktian reliabilitas dilakukan untuk mengetahui tingkat konsistensi internal instrumen dalam mengukur konstruk yang ditetapkan. Berdasarkan data empiris, instrumen literasi numerasi terdiri atas 18 butir soal dengan empat alternatif jawaban yang dikembangkan berdasarkan aspek konten dalam konstruk literasi numerasi. Hasil analisis reliabilitas menggunakan koefisien Cronbach's Alpha menunjukkan nilai sebesar 0,925. Nilai tersebut menunjukkan bahwa instrumen memiliki tingkat konsistensi internal yang sangat baik dan telah melampaui batas minimum reliabilitas sebesar 0,40 sebagaimana dikemukakan oleh Baharuddin et al. (2025). Dengan demikian, instrumen dinilai memiliki reliabilitas yang memadai untuk digunakan dalam mengukur kemampuan literasi numerasi siswa.

Uji Kualitas Butir

Tingkat kesukaran

Tingkat kesukaran merupakan indikator yang menggambarkan proporsi peserta didik yang dapat menjawab suatu butir soal dengan benar dan menjadi salah satu parameter penting dalam menentukan kualitas instrumen tes. Butir yang memiliki tingkat kesukaran terlalu rendah cenderung kurang mampu membedakan kemampuan peserta didik, sedangkan butir yang terlalu sulit dapat meningkatkan kemungkinan peserta didik menjawab secara tidak pasti atau

menebak. Kondisi tersebut dapat menyebabkan skor yang diperoleh kurang merepresentasikan kemampuan peserta didik yang sesungguhnya. Hasil analisis tingkat kesukaran terhadap 18 butir soal disajikan dalam uraian berikut.

Hasil analisis tingkat kesukaran memperlihatkan bahwa indeks kesukaran ke-18 butir soal berbasis AKM tersebar pada rentang 0,34 sampai 0,79 dengan rerata sebesar 0,54. Sebanyak 15 butir (83,3%) tergolong kategori sedang dan 3 butir (16,7%) tergolong mudah, yaitu butir V3 ($p = 0,79$), V10 ($p = 0,72$), dan V11 ($p = 0,72$), sementara tidak ditemukan satu pun butir berkategori sukar. Rerata 0,54 yang mendekati titik tengah rentang indeks kesukaran menunjukkan bahwa perangkat soal secara keseluruhan berada pada taraf yang proporsional, yakni tidak terlalu mudah sehingga gagal membedakan siswa berkemampuan tinggi, dan tidak terlalu sukar sehingga menimbulkan tebakan acak pada siswa berkemampuan rendah.

Daya Beda Butir

Daya pembeda merupakan parameter yang digunakan untuk mengidentifikasi kemampuan suatu butir soal dalam membedakan peserta didik berdasarkan tingkat kemampuannya. Butir dengan daya pembeda yang baik cenderung lebih banyak dijawab dengan benar oleh peserta didik pada kelompok berkemampuan tinggi dibandingkan kelompok berkemampuan rendah, sehingga hasil pengukuran lebih merepresentasikan kemampuan aktual peserta didik. Sebaliknya, butir dengan daya pembeda rendah kurang efektif dalam membedakan kedua kelompok tersebut dan dapat mengurangi ketepatan hasil pengukuran. Kondisi ini dapat disebabkan oleh berbagai faktor, seperti formulasi soal yang kurang jelas, adanya ambiguitas pada pilihan jawaban, atau kemungkinan peserta didik memperoleh jawaban secara kebetulan. Hasil analisis daya pembeda terhadap 18 butir soal disajikan dalam uraian berikut.

Hasil analisis daya beda butir soal berbasis AKM menunjukkan bahwa Indeks daya beda ke-18 butir tersebar pada rentang 0,394 sampai 0,817 dengan rerata sebesar 0,66. Sebanyak 13 butir (72,2%) berada pada kategori sangat baik, yaitu butir 2, 4, 5, 7, 8, 10, 12, 13, 14, 15, 16, 17, dan 18; 4 butir (22,2%) berada pada kategori baik, yaitu 1, 3, 9, dan 11; serta 1 butir (5,6%) berada pada kategori cukup, yaitu 6. Indeks tertinggi diperoleh butir 16 sebesar 0,817, diikuti butir 8 sebesar 0,794, sedangkan indeks terendah terdapat pada butir 6 sebesar 0,394. Tidak terdapat butir dengan indeks daya beda bernilai negatif maupun berada di bawah 0,30, sehingga seluruh butir dinyatakan dapat dipertahankan tanpa penggantian.

Distraktor

Pengecoh atau distraktor merupakan pilihan jawaban yang disediakan selain jawaban yang benar dalam soal pilihan ganda. Suatu distraktor dikatakan berfungsi dengan baik apabila dipilih oleh sekurang-kurangnya 5% peserta didik, yang menunjukkan bahwa pilihan tersebut cukup plausible sehingga mampu menarik peserta didik yang belum menguasai materi. Sebaliknya, distraktor yang tidak berfungsi dapat meningkatkan kemungkinan peserta didik memperoleh jawaban benar secara kebetulan karena pilihan tersebut mudah dieliminasi tanpa melalui proses penalaran yang memadai. Kondisi tersebut dapat menyebabkan tingkat

kesukaran butir menjadi lebih rendah daripada yang seharusnya. Distribusi respons peserta didik terhadap pilihan jawaban pada 18 butir soal disajikan pada [Tabel 2](#).

Tabel 2. Keberfungsian Distraktor

Butir	Pilihan				keberfungsian >5%
	A	B	C	D	
1	32,52%	8,54%	13,82%	45,12%	Berfungsi
2	28,46%	6,10%	54,88%	10,57%	Berfungsi
3	79,27%	8,9%	6,50%	5,28%	Berfungsi
4	34,55%	43,09%	13,41%	8,94%	Berfungsi
5	54,88%	16,26%	13,82%	15,04%	Berfungsi
6	47,15%	12,60%	6,10%	34,15%	Berfungsi
7	23,58%	6,91%	56,50%	13,01%	Berfungsi
8	59,76%	5,7%	28%	6,50%	Berfungsi
9	18,70%	36,59%	7,32%	37,40%	Berfungsi
10	12,60%	9,76%	71,54%	6,10%	Berfungsi
11	9,35%	72,36%	7,72%	10,57%	Berfungsi
12	8,13%	7,32%	28,05%	56,50%	Berfungsi
13	50,00%	8,13%	10,16%	31,71%	Berfungsi
14	32,11%	35,37%	19,51%	13,01%	Berfungsi
15	18,70%	53,66%	12,20%	15,45%	Berfungsi
16	5,28%	12,60%	56,10%	26,02%	Berfungsi
17	42,68%	17,89%	7,32%	32,11%	Berfungsi
18	63,01%	17,48%	9,35%	10,16%	Berfungsi

Analisis keberfungsian pengecoh dilakukan dengan menghitung persentase siswa yang memilih setiap opsi jawaban, dengan kriteria pengecoh dinyatakan berfungsi apabila dipilih oleh lebih dari 5% peserta tes. Hasil selengkapnya disajikan pada [Tabel 2](#). Secara umum, seluruh opsi jawaban pada 18 butir soal memperoleh proporsi pemilih di atas kriteria minimum, sehingga 54 pengecoh yang tersedia dinyatakan berfungsi.

Evaluasi Butir

Kelayakan butir soal tidak ditetapkan melalui satu indikator tunggal, melainkan melalui pembacaan silang atas seluruh hasil analisis. Butir yang tergolong sukar tetap dapat dipertahankan sepanjang daya beda tinggi dan pengecohnya berfungsi, sebaliknya butir dengan tingkat kesukaran ideal masih memerlukan perbaikan apabila validitas isinya belum memadai. Atas dasar itu, hasil validitas isi, tingkat kesukaran, daya beda, dan keberfungsian pengecoh ditelaah secara simultan untuk merumuskan keputusan akhir setiap butir diterima, diterima dengan revisi, atau ditolak sebagaimana direkapitulasi pada [Tabel 3](#).

Tabel 3. Evaluasi Butir

Butir	Kesukaran	Daya beda	Distraktor	Evaluasi awal
1	Sedang	Baik	Berfungsi	Dipertahankan
2	Sedang	Sangat baik	Berfungsi	Dipertahankan
3	Mudah	Baik	Berfungsi	Dipertahankan
4	Sedang	Sangat baik	Berfungsi	Dipertahankan
5	Sedang	Sangat baik	Berfungsi	Dipertahankan
6	Sedang	Cukup/baik	Berfungsi	Dipertahankan
7	Sedang	Sangat baik	Berfungsi	Dipertahankan
8	Sedang	Sangat baik	Berfungsi	Dipertahankan
9	Sedang	Baik	Berfungsi	Dipertahankan
10	Mudah	Sangat baik	Berfungsi	Dipertahankan
11	Mudah	Baik	Berfungsi	Dipertahankan
12	Sedang	Sangat baik	Berfungsi	Dipertahankan
13	Sedang	Sangat baik	Berfungsi	Dipertahankan
14	Sedang	Sangat baik	Berfungsi	Dipertahankan
15	Sedang	Sangat baik	Berfungsi	Dipertahankan
16	Sedang	Sangat baik	Berfungsi	Dipertahankan
17	Sedang	Sangat baik	Berfungsi	Dipertahankan
18	Sedang	Sangat baik	Berfungsi	Dipertahankan

Profil kemampuan siswa

Analisis deskriptif

Statistik deskriptif digunakan untuk memberikan gambaran mengenai kecenderungan pemusatan dan variasi skor hasil pengukuran. Informasi tersebut menjadi landasan dalam menginterpretasikan hasil analisis lebih lanjut, karena nilai rerata yang sama dapat menunjukkan pola distribusi skor yang berbeda sehingga memerlukan strategi tindak lanjut pembelajaran yang berbeda pula. Ringkasan statistik deskriptif kemampuan literasi dan numerasi peserta didik disajikan pada [Tabel 4](#).

Tabel 4. Analisis Deskriptif

Deskriptif	Skor
skor maksimum	18
skor minimum	0
rata-rata	9,65
median	10
Modus	16

Kemampuan siswa terhadap penguasaan

Skor yang diperoleh peserta didik menjadi dasar dalam melakukan interpretasi hasil pengukuran, baik untuk menentukan tingkat pencapaian terhadap kompetensi yang ditargetkan

maupun untuk mengidentifikasi variasi kemampuan antar peserta didik dalam kelas. Distribusi skor peserta didik berdasarkan 18 butir soal yang diujikan disajikan pada [Tabel 5](#).

Tabel 5. Skor Jawaban perolehan siswa

Penguasaan siswa	
0 sampai 5	75
6 sampai 12	79
13 sampai 18	92
Total	246

Profil kemampuan siswa pad aspek Proses Kognitif

Analisis kualitas butir memberikan informasi mengenai karakteristik dan kelayakan instrumen, tetapi belum secara langsung menggambarkan tingkat penguasaan kompetensi peserta didik. Oleh karena itu, hasil pengukuran dianalisis lebih lanjut untuk memperoleh profil kemampuan peserta didik pada masing-masing aspek proses kognitif. Analisis berdasarkan aspek diperlukan karena skor total yang relatif sama belum tentu menunjukkan tingkat penguasaan yang merata pada setiap domain. Identifikasi kesenjangan penguasaan pada masing-masing aspek menjadi penting untuk menentukan prioritas tindak lanjut pembelajaran. Dalam konteks Kurikulum Merdeka, informasi tersebut dapat menjadi landasan bagi guru dalam merancang pembelajaran yang terdiferensiasi sesuai dengan kebutuhan dan karakteristik kemampuan peserta didik. Rata-rata penguasaan peserta didik pada aspek proses kognitif disajikan pada [Tabel 6](#).

Tabel 6. Penguasaan siswa aspek Proses kognitif

Aspek Proses Kognitif	Rerata Penguasaan
Pemahaman Memahami fakta, prosedur serta alat matematika.	52%
Penerapan Mampu menerapkan konsep matematika dalam situasi nyata yang bersifat rutin.	59%
Penalaran Bernalar dengan konsep matematika untuk menyelesaikan masalah yang bersifat non rutin.	50%

Penguasaan siswa dianalisis berdasarkan tiga level proses kognitif numerasi, yaitu pemahaman, penerapan, dan penalaran, sebagaimana disajikan pada [Tabel 6](#). Rerata penguasaan siswa pada ketiga aspek berada pada rentang 50% sampai 59% dengan rerata keseluruhan sebesar 53,67%. Penguasaan tertinggi dicapai pada aspek penerapan, yakni kemampuan menerapkan konsep matematika dalam situasi nyata yang bersifat rutin, sebesar 59%. Aspek pemahaman, yang mencakup penguasaan fakta, prosedur, serta alat matematika, memperoleh rerata 52%, sedangkan penguasaan terendah terdapat pada aspek penalaran, yaitu kemampuan bernalar dengan konsep matematika untuk menyelesaikan masalah non-rutin, sebesar 50%. Selisih antara aspek tertinggi dan terendah sebesar 9%, sementara selisih antara aspek pemahaman dan penalaran hanya 2%.

Pembahasan

Penelitian ini bertujuan menganalisis kualitas butir soal berbasis Asesmen Kompetensi Minimum (AKM) dalam mengukur kemampuan literasi dan numerasi siswa SMP pada implementasi Kurikulum Merdeka. Kebutuhan atas instrumen asesmen yang teruji kualitasnya menjadi semakin mendesak seiring pergeseran orientasi penilaian dari penguasaan konten menuju kemampuan bernalar dan menerapkan pengetahuan dalam konteks nyata, yang menjadi ruh dari Kurikulum Merdeka. Butir soal berbasis AKM menuntut siswa memahami, menggunakan, mengevaluasi, dan merefleksikan informasi dari beragam stimulus, sehingga karakteristiknya berbeda secara mendasar dari butir soal konvensional dan menuntut prosedur pengujian kualitas yang cermat sebelum digunakan sebagai alat ukur. Oleh karena itu, kelayakan butir soal tidak cukup ditinjau dari satu sisi saja, melainkan perlu ditelaah secara menyeluruh melalui validitas isi, reliabilitas, tingkat kesukaran, daya beda, dan keberfungsian distraktor. Pembahasan berikut menguraikan temuan penelitian pada setiap aspek tersebut secara berurutan, membandingkannya dengan hasil penelitian terdahulu, serta menafsirkan implikasinya bagi pengembangan instrumen asesmen literasi dan numerasi di jenjang SMP.

Hasil analisis tingkat kesukaran memperlihatkan bahwa indeks kesukaran ke-18 butir soal berbasis AKM tersebar pada rentang 0,34 sampai 0,79 dengan rerata sebesar 0,54. Sebanyak 15 butir (83,3%) tergolong kategori sedang dan 3 butir (16,7%) tergolong mudah, yaitu butir 3 ($p = 0,79$), 10 ($p = 0,72$), dan 11 ($p = 0,72$), sementara tidak ditemukan satu pun butir berkategori sukar. Rerata 0,54 yang mendekati titik tengah rentang indeks kesukaran menunjukkan bahwa perangkat soal secara keseluruhan berada pada taraf yang proporsional, yakni tidak terlalu mudah sehingga gagal membedakan siswa berkemampuan tinggi, dan tidak terlalu sukar sehingga menimbulkan tebakan acak pada siswa berkemampuan rendah. Dominasi kategori sedang ini mengindikasikan bahwa stimulus kontekstual dan tuntutan penalaran yang menjadi ciri butir AKM masih dapat dijangkau oleh siswa SMP pada tingkat kelas yang diteliti. Meskipun demikian, ketiadaan butir berkategori sukar perlu dicermati, sebab perangkat soal menjadi kurang optimal dalam memetakan siswa pada level kognitif tertinggi, terutama pada tuntutan literasi menginterpretasi mengevaluasi dan numerasi bernalar. Di sisi lain, lima butir dengan indeks terendah ($6 = 0,34$; $14 = 0,35$; $9 = 0,37$; 4 dan 17 masing-masing $0,43$) berada di batas bawah kategori sedang, yang menandakan bahwa butir-butir tersebut menuntut proses kognitif paling kompleks di antara keseluruhan butir sekaligus berpotensi memuat kesulitan pada aspek keterbacaan stimulus.

Temuan ini sejalan dengan penelitian [Purwati, et, al., \(2021\)](#) yang melaporkan bahwa butir asesmen berbasis AKM cenderung terkonsentrasi pada kategori sedang karena karakteristik stimulusnya yang kontekstual dan berbasis situasi nyata, sehingga siswa memiliki pijakan informasi untuk menjawab meskipun tuntutan kognitifnya tinggi. Hasil serupa juga dilaporkan ([Agustiniingsih & Fahrurrozi, 2024](#)) pada instrumen literasi dan numerasi jenjang SMP, dengan proporsi butir sedang yang dominan dan sedikitnya butir sukar. Sebaliknya, temuan ini berbeda dengan [Sari et. al., \(2022\)](#) yang menemukan proporsi butir sukar cukup besar pada soal bertipe AKM; perbedaan tersebut dapat ditelusuri pada karakteristik responden, keterbiasaan siswa mengerjakan soal berbasis stimulus, serta cakupan konten yang diujikan.

Merujuk pada kriteria yang menyatakan bahwa perangkat tes yang baik idealnya memuat komposisi butir mudah, sedang, dan sukar secara berimbang, hasil penelitian ini menunjukkan bahwa perangkat soal berbasis AKM yang dikembangkan telah memenuhi syarat kelayakan dari segi tingkat kesukaran, namun komposisinya belum sepenuhnya seimbang karena tidak terwakilinya kategori sukar. Dengan demikian dapat disimpulkan bahwa ditinjau dari tingkat kesukaran, butir soal berbasis AKM yang dikembangkan layak digunakan untuk mengukur kemampuan literasi dan numerasi siswa SMP pada implementasi Kurikulum Merdeka, dengan catatan perlunya penambahan butir pada level kognitif yang lebih tinggi agar perangkat mampu membedakan capaian siswa secara lebih rentang penuh.

Rerata daya beda sebesar 0,66 dengan 94,4% butir berada pada kategori baik hingga sangat baik menunjukkan bahwa perangkat soal berbasis AKM yang dikembangkan memiliki kemampuan diskriminatif yang tinggi dalam memisahkan siswa berkemampuan literasi dan numerasi tinggi dari siswa berkemampuan rendah. Nilai ini jauh melampaui batas minimum 0,30 yang lazim dijadikan syarat kelayakan butir, bahkan lebih dari dua pertiga butir melewati ambang 0,60. Ketiadaan butir berindeks negatif merupakan indikator penting, sebab indeks negatif menandakan butir yang justru lebih banyak dijawab benar oleh siswa berkemampuan rendah dan biasanya bersumber dari kesalahan kunci jawaban, ambiguitas rumusan, atau stimulus yang menyesatkan. Tingginya daya beda ini dapat ditelusuri pada karakteristik butir AKM itu sendiri, yang menuntut siswa tidak sekadar mengingat informasi melainkan menginterpretasi, menghubungkan, dan mengevaluasi informasi dari stimulus; tuntutan kognitif berjenjang tersebut secara alamiah menghasilkan pemisahan respons yang tajam antarkelompok kemampuan. Butir 16 dengan indeks tertinggi (0,817) merepresentasikan butir yang paling sensitif terhadap perbedaan kemampuan siswa, sementara butir 6 dengan indeks terendah (0,394) menarik dicermati karena butir yang sama juga memiliki tingkat kesukaran terendah ($p = 0,34$); kombinasi keduanya mengindikasikan bahwa butir ini tergolong menantang bagi hampir seluruh siswa termasuk kelompok atas, sehingga rentang perbedaan skor yang dihasilkannya menyempit. Kondisi demikian tidak menggugurkan kelayakan butir, namun mengisyaratkan perlunya penelaahan ulang pada aspek keterbacaan stimulus dan kejelasan tuntutan pertanyaan.

Temuan ini menguatkan hasil penelitian [Agustiningsih & Fahrurrozi, \(2024\)](#) yang melaporkan bahwa instrumen berbasis AKM pada jenjang SMP menghasilkan daya beda dengan proporsi kategori baik dan sangat baik yang dominan, seiring dengan tuntutan penalaran tingkat tinggi pada butir-butirnya. Hasil serupa juga ditemukan [Sari et. al., \(2022\)](#) pada pengembangan soal literasi dan numerasi, yang menyimpulkan bahwa butir berbasis stimulus kontekstual lebih diskriminatif dibandingkan butir berbasis hafalan konten. Sebaliknya, [Zulfikar, et al., \(2023\)](#) menemukan sebagian butir AKM berdaya beda rendah bahkan negatif; perbedaan ini dapat dijelaskan oleh variasi keterbiasaan siswa menghadapi soal bertipe AKM, homogenitas kemampuan responden, serta kualitas pengembangan stimulus dan pengecoh. Apabila hasil daya beda ini dibaca bersama hasil tingkat kesukaran yang didominasi kategori sedang, tampak adanya kesesuaian yang saling menguatkan: butir dengan tingkat kesukaran proporsional memang cenderung menghasilkan daya beda yang optimal, sebagaimana

dinyatakan dalam teori pengukuran klasik. Dengan demikian dapat disimpulkan bahwa ditinjau dari daya beda, butir soal berbasis AKM yang dikembangkan memiliki kualitas yang sangat memadai dan layak digunakan untuk mengukur serta memetakan kemampuan literasi dan numerasi siswa SMP pada implementasi Kurikulum Merdeka, dengan catatan butir 6 sebaiknya direvisi pada aspek redaksional agar kemampuan diskriminatifnya meningkat

Keberfungsian seluruh pengecoh pada 18 butir soal menunjukkan bahwa opsi jawaban yang dikembangkan bukan sekadar pelengkap struktur soal pilihan ganda, melainkan benar-benar merepresentasikan kemungkinan miskonsepsi dan kekeliruan penalaran yang nyata dimiliki siswa. Pengecoh yang berfungsi merupakan prasyarat penting bagi akurasi pengukuran, sebab opsi yang tidak dipilih siapa pun secara efektif mereduksi soal empat pilihan menjadi soal tiga atau dua pilihan, sehingga meningkatkan peluang siswa menjawab benar melalui tebakan dan pada akhirnya mengelembungkan estimasi kemampuan. Temuan ini sekaligus konsisten dengan hasil daya beda yang telah diuraikan sebelumnya: pengecoh yang berfungsi optimal menyerap respons siswa berkemampuan rendah, sehingga pemisahan antara kelompok atas dan kelompok bawah menjadi lebih tajam. Meskipun demikian, terdapat catatan penting pada tiga butir. Pada butir 6, pengecoh A justru dipilih oleh 47,15% siswa, jauh melampaui kunci jawaban yang hanya dipilih 34,15%; pola serupa muncul pada 9 dan 14. Kondisi ini lazim disebut sebagai over-attractive distractor dan dapat bersumber dari dua kemungkinan yang berbeda implikasinya. Kemungkinan pertama, pengecoh tersebut merepresentasikan miskonsepsi yang sangat kuat dan tersebar luas di kalangan siswa, sehingga temuan ini bernilai diagnostik tinggi bagi guru dalam merancang pembelajaran remedial. Kemungkinan kedua, terdapat kelemahan konstruksi pada rumusan pengecoh yang membuatnya tampak sama benarnya dengan kunci, atau stimulus yang memuat informasi ambigu. Mengingat ketiga butir tersebut tetap memiliki daya beda pada kategori cukup hingga sangat baik, kemungkinan pertama lebih dapat dipertanggungjawabkan, sebab butir dengan pengecoh yang keliru secara konstruksi umumnya menghasilkan daya beda rendah atau negatif.

Temuan mengenai berfungsinya seluruh pengecoh sejalan dengan penelitian [Sari et. al., \(2022\)](#) yang melaporkan bahwa butir soal berbasis AKM cenderung memiliki pengecoh yang efektif karena opsi jawabannya dikembangkan dari pola kesalahan autentik siswa dalam menafsirkan stimulus, bukan dibangun secara acak. Hasil serupa dilaporkan [Agustiningsih & Fahrurrozi, \(2024\)](#) pada instrumen literasi dan numerasi jenjang SMP, yang menyimpulkan bahwa penggunaan stimulus kontekstual memperkaya variasi respons siswa sehingga distribusi pemilihan opsi menjadi lebih menyebar. Sebaliknya, [Zulfikar, et al., \(2023\)](#) menemukan sejumlah pengecoh tidak berfungsi pada soal buatan guru, yang umumnya disebabkan oleh opsi yang secara logis mudah dieliminasi atau tidak sejajar secara konten dengan kunci jawaban; perbedaan tersebut menegaskan bahwa kualitas pengecoh sangat ditentukan oleh proses pengembangan butir yang berbasis analisis miskonsepsi. Adapun temuan mengenai pengecoh yang lebih menarik daripada kunci jawaban selaras dengan [Sari et. al., \(2022\)](#), yang memandang fenomena ini bukan semata kelemahan butir melainkan informasi diagnostik yang menunjukkan letak kesulitan konseptual siswa. Dengan demikian dapat disimpulkan bahwa ditinjau dari keberfungsian pengecoh, butir soal berbasis AKM yang dikembangkan telah

memenuhi kriteria kualitas dan layak digunakan untuk mengukur kemampuan literasi dan numerasi siswa SMP, sementara butir V6, V9, dan V14 direkomendasikan untuk ditelaah lebih lanjut guna memastikan apakah pola respons yang muncul mencerminkan miskonsepsi siswa atau menuntut perbaikan redaksional pada opsi jawaban.

Rerata penguasaan sebesar 53,67% menunjukkan bahwa kemampuan numerasi siswa secara umum baru berada sedikit di atas separuh dari capaian maksimal, sehingga tergolong belum optimal untuk memenuhi tuntutan kompetensi minimum yang diharapkan. Pola sebaran antar aspek justru memuat temuan yang lebih penting daripada angka reratanya. Secara teoretis, proses kognitif numerasi tersusun secara hierarkis, sehingga penguasaan diprediksi menurun dari pemahaman, penerapan, hingga penalaran. Data penelitian ini memperlihatkan pola yang menyimpang dari prediksi tersebut, karena aspek penerapan (59%) justru melampaui aspek pemahaman (52%). Penyimpangan ini dapat dimaknai bahwa siswa relatif lebih mampu menyelesaikan soal yang menyajikan konteks situasi nyata dengan prosedur rutin dibandingkan soal yang menuntut penguasaan fakta, prosedur, dan alat matematika secara terpisah dari konteksnya. Dengan kata lain, konteks berfungsi sebagai penopang kognitif yang membantu siswa mengaktifkan pengetahuannya, sementara pengetahuan prosedural yang berdiri sendiri belum dikuasai secara mantap suatu indikasi bahwa penguasaan siswa cenderung bersifat kontekstual dan prosedural, bukan konseptual. Adapun aspek penalaran yang memperoleh capaian terendah (50%) menegaskan bahwa kesulitan terbesar siswa terletak pada masalah non-rutin yang menuntut pemilihan strategi secara mandiri, penghubungan antar konsep, dan justifikasi jawaban, yakni justru kompetensi yang menjadi sasaran utama AKM dan Kurikulum Merdeka.

Temuan bahwa penalaran menjadi aspek terlemah sejalan dengan penelitian [Sari et. al., \(2022\)](#) yang melaporkan bahwa siswa SMP mengalami kesulitan paling besar pada butir numerasi level penalaran, terutama ketika soal tidak menyediakan petunjuk prosedur penyelesaian secara eksplisit. Hasil serupa dilaporkan [Dewi et al., \(2021\)](#) yang menemukan capaian numerasi siswa terkonsentrasi pada level pemahaman dan penerapan dengan penurunan tajam pada level penalaran. Pola capaian penerapan yang melampaui pemahaman dalam penelitian ini berbeda dengan [Estherika & Setianingsih, \(2022\)](#) yang memperoleh urutan capaian sesuai hierarki kognitif; perbedaan tersebut dapat ditelusuri pada karakteristik pembelajaran yang dialami siswa, yang menekankan latihan soal berpola rutin sehingga siswa terbiasa mengenali tipe soal dan menerapkan prosedur yang telah dicontohkan tanpa disertai pemahaman konseptual yang kokoh. Temuan ini juga selaras dengan hasil analisis tingkat kesukaran yang telah diuraikan sebelumnya, yaitu ketiadaan butir berkategori sukar dan rendahnya indeks kesukaran pada butir-butir yang menuntut penalaran, sehingga kedua analisis saling menguatkan dalam menunjukkan letak kesenjangan kemampuan siswa. Dengan demikian dapat disimpulkan bahwa instrumen berbasis AKM yang dikembangkan tidak hanya berfungsi sebagai alat ukur, tetapi juga sebagai sarana diagnostik yang mampu memetakan profil kemampuan numerasi siswa secara berjenjang, dan hasilnya mengisyaratkan perlunya penguatan pembelajaran yang berorientasi pada penalaran dan pemecahan masalah non-rutin dalam implementasi Kurikulum Merdeka.

4. KESIMPULAN

Butir soal berbasis AKM yang dikembangkan terbukti berkualitas memadai untuk mengukur kemampuan literasi dan numerasi siswa SMP pada implementasi Kurikulum Merdeka. Seluruh butir valid secara isi dengan koefisien V Aiken 0,60–0,85, tingkat kesukaran didominasi kategori sedang dengan, daya beda berada pada kategori baik hingga sangat baik tanpa indeks negatif, dan seluruh pengecoh berfungsi. Profil penguasaan siswa tertinggi pada aspek penerapan (59%), diikuti pemahaman (52%), dan terendah pada penalaran (50%), yang menunjukkan penguasaan siswa masih bersifat prosedural dan bergantung pada konteks yang disediakan. Keterbatasan penelitian terletak pada penggunaan teori tes klasik yang hasilnya bergantung sampel, cakupan yang terbatas pada 18 butir dan satu wilayah, serta belum ditelusurinya secara kualitatif penyebab munculnya pengecoh yang lebih menarik daripada kunci jawaban. Penelitian selanjutnya disarankan menerapkan teori respons butir atau pemodelan Rasch untuk memperoleh parameter yang invarian, memperluas responden lintas sekolah dan wilayah, menambah butir pada level penalaran.

Temuan ini berimplikasi secara praktis dan teoretis. Tersedianya perangkat soal AKM yang teruji memberi rujukan konkret bagi guru dalam menyusun asesmen yang selaras dengan Kurikulum Merdeka, sekaligus menegaskan bahwa pengembangan butir harus disertai analisis kualitas secara sistematis. Pola pemilihan pengecoh menunjukkan bahwa analisis butir dapat difungsikan melampaui perannya sebagai prosedur seleksi soal, yaitu sebagai sarana diagnostik untuk memetakan miskonsepsi sebagai dasar perancangan pembelajaran remedial. Rendahnya capaian penalaran menuntut reorientasi pembelajaran dari latihan soal rutin menuju pembiasaan masalah non-rutin yang menuntut pemilihan strategi dan justifikasi jawaban. Secara teoretis, capaian aspek penerapan yang melampaui pemahaman memperkaya kajian hierarki proses kognitif numerasi dan membuka ruang penelitian mengenai peran konteks sebagai penopang kognitif dalam asesmen berbasis stimulus.

REFERENSI

- Abidin, Y. (2018). Pembelajaran literasi. Bumi Aksara.
- Agustiningih, F., & Fahrurrozi. (2024). Pengembangan instrumen asesmen kompetensi minimum (AKM) literasi numerasi untuk siswa kelas IV SD. *Jurnal Pendidikan Matematika*, 2(2), 69–80. DOI: <https://doi.org/10.70115/notasi.v2i2.193>
- Arikunto, S. (2018). Dasar-dasar evaluasi pendidikan. Bumi Aksara.
- Asri, S. A., Syahrul, & Ishak, S. (2025). Analisis kualitas butir soal ulangan akhir semester genap matematika kelas VIII SMPN Sinjai. *Pedagogy: Jurnal Pendidikan Matematika*, 10(1), 120–133. DOI: <https://doi.org/10.30605/pedagogy.v10i1.5672>
- Azwar, S. (2015). Reliabilitas dan validitas. Pustaka Pelajar.
- Baharuddin, H., Riskawati, & Fitrawardani. (2025). Identifikasi kemampuan prosedural dan konseptual siswa dalam memecahkan masalah matematika dengan model DINA. *Jurnal*

DOI: <https://doi.org/10.63976/jimat.v7i3.1568>

-
- Studi Guru dan Pembelajaran, 8(3), 2273–2282. DOI: <https://doi.org/10.30605/jsdp.8.3.2025.8930>
- Borg, W. R., & Gall, M. D. (2003). Educational research: An introduction. Allyn & Bacon.
- Brookhart, S. M. (2010). How to assess higher-order thinking skills in your classroom. ASCD.
- Crocker, L., & Algina, J. (2008). Introduction to classical and modern test theory. Cengage Learning.
- Creswell, J. W. (2014). Research design. Sage Publications.
- Purwati, P. D., Faiz, A., & Widiyatmoko, A. (2021). Asesmen kompetensi minimum (AKM) kelas jenjang sekolah dasar sarana pemacu peningkatan literasi peserta didik. *Sosio Religi: Jurnal Kajian Pendidikan Umum*, 19(1). DOI: <https://doi.org/10.17509/sosioreligi.v19i1.39347>
- Ebel, R. L., & Frisbie, D. A. (1991). Essentials of educational measurement. Prentice Hall.
- Haladyna, T. M. (2004). Developing and validating multiple-choice test items. Lawrence Erlbaum Associates.
- Ishak, S. (2018). Penerapan Metode Eksplorasi Dalam Meningkatkan Hasil Belajar Matematika Pada Peserta Didik Kelas IX SMP Negeri 5 Campalagian. *Pepatudzu: Media Pendidikan dan Sosial Kemasyarakatan*, 13(1), 75-87.
- Ishak, S., Ninsih, N. F., Samad, I., Wibowo, A., & Wajdi, F. (2025). Analisis learning obstacle siswa dalam menyelesaikan soal computational thinking pada materi barisan dan deret. *Jurnal Ilmiah Matematika (JIMAT)*, 6(2), 821–842. DOI: <https://doi.org/10.63976/jimat.v6i2.1191>
- Anggraini, K. E., & Setianingsih, R. (2022). Analisis kemampuan numerasi siswa SMA dalam menyelesaikan soal Asesmen Kompetensi Minimum (AKM). *MATHEdunesa*, 11(3), 837–849. DOI: <https://doi.org/10.26740/mathedunesa.v11n3.p837-849>
- Kemendikbudristek. (2022). Panduan implementasi Kurikulum Merdeka. Kemendikbudristek.
- Kemendikbudristek. (2021). Asesmen Nasional: AKM dan implikasinya. Kemendikbudristek.
- Nitko, A. J., & Brookhart, S. M. (2014). Educational assessment of students. Pearson.
- OECD. (2019). PISA 2018 assessment and analytical framework. OECD Publishing.
- Trilling, B., & Fadel, C. (2009). 21st century skills. Jossey-Bass.
- Sari, V. P., & Sayekti, I. C. (2022). Evaluasi pelaksanaan Asesmen Kompetensi Minimum (AKM) pada kompetensi dasar literasi membaca peserta didik sekolah dasar. *Jurnal Basicedu*, 6(3), 5237–5243. DOI: <https://doi.org/10.31004/basicedu.v6i3.2907>
- Widana, I. W. (2017). Higher order thinking skills assessment. Gramedia.
- Zulfikar, T., Taufiq, & Saputra, H. (2023). Peningkatan kemampuan koneksi matematis siswa SMA dengan menggunakan model pembelajaran Cooperative-Meaningful Instructional Design (MID). *Education Enthusiast: Jurnal Pendidikan dan Keguruan*, 3(4), 24–40.