



Jurnal Amal Pendidikan

e-ISSN 2721-3668 | p-ISSN 2746-4725

Penerbit: FKIP Universitas Halu Oleo, Jl. HEA Mokodompit Kampus Hijau Tridharma Andounohu Kota
Kendari, website: <https://japend.uho.ac.id/>, email: japend@uho.ac.id



SURAT KETERANGAN PENERIMAAN ARTIKEL

Kepada Yth. Bapak/Ibu. **Sofia, Lukman Ibrahim, Khairina Hasan**

Bersama ini, Pimpinan Redaksi **Jurnal Amal Pendidikan** menyampaikan bahwa Artikel Bapak/Ibu kiriman dinyatakan **DITERIMA** untuk dipublikasikan pada **Jurnal Amal Pendidikan**.

Judul : **Development of Literacy Test Instruments to Improve Mathematical Concept Understanding in High School Students**

Edisi Terbitan : **Volume 6 Nomor 2 : Edisi Agustus 2025**

Kami sampaikan pula bahwa artikel ini telah melalui proses submit, review, revisi daring penuh dan proses review menggunakan sistem *double blind review*. Jurnal Amal Pendidikan telah terakreditasi secara nasional dengan peringkat 4 (**Sinta 4**) sesuai SK Dirjen Dikti Kemdikbudristek Nomor 79/E/ KPT/2023 dengan masa akreditasi mulai terbitan Volume 2 Nomor 1 Tahun 2021 sampai Volume 6 Nomor 3 Tahun 2025. Informasi lainnya terkait editing dan publish artikel jurnal dapat dipantau pada website *Open Jurnal System* Jurnal Amal Pendidikan.

Demikian surat keterangan ini dibuat untuk dipergunakan sebagaimana mestinya. Terima kasih telah memilih Jurnal Amal Pendidikan untuk publikasi artikel anda.

Kendari, 19 Juli 2025

Pimpinan Redaksi



Dr. Mustamin Anggo, M.Si
Chief Editor Jurnal Amal Pendidikan



Pengembangan Instrumen Tes Literasi untuk Meningkatkan Pemahaman Konsep Matematis pada Siswa SMA

Sofia¹⁾ *, Lukman Ibrahim²⁾, Khairina Hasan³⁾

¹Pendidikan Matematika, Universitas Islam Negeri Ar-Raniry. Banda Aceh, Indonesia.

²Pendidikan Matematika, Universitas Islam Negeri Ar-Raniry. Banda Aceh, Indonesia.

³Pendidikan Matematika, Universitas Islam Negeri Ar-Raniry. Banda Aceh, Indonesia.

Received: DD MM YYYY

Revised: DD MM YYYY

Accepted: DD MM YYYY

Abstract

This study aims to develop a valid and reliable mathematical literacy test instrument in measuring the understanding of mathematical concepts of Senior High School (SMA) students. This development is motivated by the need to strengthen students' readiness in facing the Minimum Competency Assessment (AKM), which demands high literacy and numeracy skills. The development model used is four-D with the steps being define, design, develop, and disseminate, which focuses on one of the materials in SMA, namely the Three Variable Linear Equation System. The instrument is designed to measure all indicators of mathematical concept understanding, namely the ability to interpret, apply, and connect concepts in real contexts, as required in AKM questions, as an evaluation of the implementation of the Kurikulum Merdeka. The validation results show that the developed instrument is suitable for use as a measuring tool for mathematical literacy. The product of this development is expected to contribute to reducing the need for assessment instruments as practice materials for students in solving AKM problems and the need for mathematics learning under the Kurikulum Merdeka.

Keywords: *Mathematical Literacy; Concept Understanding; AKM Test Instrument; 4-D Model*

Abstrak

Penelitian ini bertujuan untuk mengembangkan instrumen tes literasi matematika yang valid dan reliabel dalam mengukur pemahaman konsep matematis siswa Sekolah Menengah Atas (SMA). Pengembangan ini dilatarbelakangi oleh perlunya penguatan kesiapan siswa dalam menghadapi Asesmen Kompetensi Minimum (AKM) yang menuntut kemampuan literasi matematis dengan kemampuan numerasi tinggi. Model pengembangan yang digunakan adalah four-D dengan langkah-langkah define, design, develop, disseminate yang difokuskan pada materi Sistem Persamaan Linear Tiga Variabel (SPLTV). Instrumen disusun untuk mengukur semua indikator pemahaman konsep matematis, yaitu kemampuan menginterpretasi, mengaplikasikan, dan menghubungkan konsep dalam konteks nyata, sebagaimana dituntut dalam soal-soal AKM sebagai evaluasi implementasi Kurikulum Merdeka. Hasil validasi menunjukkan bahwa instrumen yang dikembangkan layak digunakan sebagai alat ukur literasi matematis. Produk dari pengembangan ini diharapkan memberikan kontribusi untuk mengurangi kebutuhan instrumen asesmen sebagai bahan berlatih bagi siswa dalam menyelesaikan soal-soal AKM dan kebutuhan pembelajaran matematika sesuai dengan Kurikulum Merdeka.

Kata Kunci: Literasi Matematis; Pemahaman Konsep; Instrumen Tes AKM; Model 4D

PENDAHULUAN

Perkembangan ilmu pengetahuan dan teknologi menuntut siswa tidak hanya menguasai konsep-konsep akademik, tetapi juga memiliki kemampuan berpikir kritis, kreatif, dan kontekstual. Dalam konteks pembelajaran matematika, kemampuan tersebut tercermin dalam literasi matematika, yakni kemampuan untuk merumuskan, menerapkan, dan menafsirkan konsep serta prosedur matematika dalam berbagai situasi kehidupan nyata. Literasi matematika merupakan kompetensi esensial abad ke-21 yang mengaitkan pemahaman konsep dengan

* Author Correspondence. E-mail:

situasi nyata secara kontekstual (Setiawan & Fitriyani, 2021). Literasi matematika tidak hanya mengukur seberapa jauh siswa memahami teori, tetapi juga bagaimana siswa mampu memecahkan masalah kontekstual secara logis dan bernalar. Literasi matematika menuntut kemampuan menyelesaikan masalah terbuka dan tak rutin yang memerlukan nalar matematis (D. Wulandari & Saputra, 2021).

AKM menggeser fokus asesmen dari hafalan semata menjadi pengukuran pada dua domain esensial literasi membaca dan numerasi (Kebudayaan, 2020). Paradigma asesmen pendidikan nasional saat ini menekankan kemampuan reflektif dan bernalar, bukan sekadar reproduksi informasi (Sulastri & Rofiah, 2022). Dalam domain numerasi, siswa diuji kemampuannya dalam menggunakan matematika untuk menyelesaikan masalah kehidupan sehari-hari, yang pada dasarnya merupakan implementasi dari literasi matematika. Hal ini menandakan bahwa peningkatan kualitas pembelajaran matematika di sekolah tidak cukup dilakukan hanya melalui penguatan materi ajar, melainkan juga harus ditunjang dengan pengembangan instrumen asesmen yang relevan dan kontekstual. Literasi matematika menjadi salah satu indikator keberhasilan implementasi Kurikulum Merdeka karena menuntut pemahaman konsep dan koneksi antar topik (M. T. Hidayat & Fsetaraah, 2024).

Literasi yang dikembangkan adalah literasi matematika, karena instrumen dirancang untuk mengukur kemampuan siswa merumuskan, menerapkan, dan menghubungkan konsep matematis dalam konteks nyata sesuai tuntutan AKM dan Kurikulum Merdeka. Meski demikian, literasi numerasi juga tercakup, sebab siswa tetap harus mengolah angka dan melakukan perhitungan. Namun, fokus utama penelitian adalah literasi matematika yang lebih luas, karena menekankan pemahaman konsep dan penerapannya dalam pemecahan masalah, bukan sekadar keterampilan berhitung.

Penelitian oleh Handayani dan Rahayu (2022) (Handayani & Rahayu, 2022) mengembangkan soal numerasi berbasis AKM pada materi bangun datar untuk siswa SMP. Instrumen diuji validitas dan reliabilitasnya dan menunjukkan hasil sangat baik. Persamaannya dengan penelitian ini adalah penggunaan pendekatan AKM dan pengujian validitas-reliabilitas. Perbedaannya terletak pada domain numerasi, jenjang SMP, serta topik yang bersifat geometri, sedangkan penelitian ini berada pada domain literasi matematis SMA dengan materi SPLTV. Penelitian oleh Fitriana dan Sugiyono (2023) (Fitriana & Sugiyono, 2023) mengembangkan instrumen diagnostik berbasis literasi matematika pada materi fungsi di kelas X SMA. Penelitian ini bertujuan mendiagnosis kesulitan awal siswa melalui pendekatan literasi. Persamaannya adalah sama-sama mengembangkan instrumen literasi di jenjang SMA dan melakukan pengujian validitas dan reliabilitas. Perbedaannya terletak pada fungsi asesmen penelitian Fitriana lebih ke diagnosis awal (formatif), sedangkan penelitian ini bertujuan menyediakan asesmen akhir (sumatif) dan pengujian kelayakan soal. Penelitian terbaru oleh Hidayat dan Lestari (2024) (R. Hidayat & Lestari, 2024) mengembangkan instrumen asesmen berbasis AKM multilevel dengan integrasi konteks sosial, sains, dan budaya. Pendekatan mereka menggunakan spiral item development dan analisis menggunakan model Rasch. Persamaannya terletak pada penggunaan pendekatan AKM dan orientasi soal berbasis konteks. Namun, perbedaannya adalah pendekatan yang digunakan (Rasch vs. klasik), serta cakupan soal multilevel, sementara penelitian ini fokus pada satu topik mendalam, yakni SPLTV.

Meskipun soal AKM sudah banyak beredar, kualitasnya belum semuanya terjamin secara ilmiah. Beberapa penelitian menunjukkan bahwa banyak soal yang tersedia masih berfokus pada keterampilan hitung dan prosedur tanpa mengukur indikator pemahaman konsep (Latifah & Ramdhan, 2023). Stimulus yang digunakan juga sering kurang kontekstual dan tidak relevan dengan kehidupan siswa (M. T. Hidayat & Fadilah, 2024). Selain itu, soal AKM yang

beredar jarang diuji kelayakannya melalui analisis validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan kesetaraan antar kelompok, padahal hal ini penting untuk menjamin kualitas asesmen (R. Fauziah & Yulianto, 2023). Laporan AKM Kemendikbudristek (2021) juga mengungkapkan rendahnya capaian literasi matematika siswa Indonesia, khususnya pada soal kontekstual yang memerlukan penalaran logis dan representasi simbolik. Instrumen tes literasi yang dikembangkan memiliki keunggulan dibandingkan soal AKM yang sudah ada. Jika soal yang beredar umumnya masih prosedural dan kurang kontekstual, instrumen ini dirancang untuk benar-benar mengukur pemahaman konsep melalui kemampuan menginterpretasi, mengaplikasikan, dan menghubungkan konsep matematis. Stimulus yang digunakan berbasis masalah nyata dari kehidupan sehari-hari, dengan fokus khusus pada materi SPLTV yang kompleks dan relevan. Selain itu, soal divariasikan dalam bentuk pilihan ganda, pilihan ganda kompleks, benar salah, menjodohkan dan uraian. Keunggulan lainnya, instrumen ini telah teruji validitas, reliabilitas, tingkat kesukaran, daya pembeda, serta terbukti konsisten antar kelompok siswa, sehingga kualitasnya lebih terjamin dan sesuai dengan tuntutan Kurikulum Merdeka.

Indikator yang menunjukkan pemahaman konsep matematika yaitu menyatakan ulang sebuah konsep, mengklasifikasikan objek-objek menurut sifat-sifat tertentu (sesuai dengan konsepnya), memberi contoh dan non-contoh dari konsep, menyajikan konsep dalam berbagai bentuk representasi matematis, mengembangkan syarat perlu dan syarat cukup suatu konsep, menggunakan, memanfaatkan, dan memilih prosedur atau operasi tertentu, mengaplikasikan konsep atau algoritma dalam pemecahan masalah. Dengan demikian, literasi matematika tidak dipahami hanya sebagai kemampuan membaca konteks soal atau menghitung prosedural, tetapi sebagai keterampilan komprehensif yang menghubungkan konsep matematis dengan kehidupan nyata.

Materi Sistem Persamaan Linear Tiga Variabel (SPLTV) dipilih sebagai fokus karena memiliki kompleksitas yang cukup dan keterkaitan dengan persoalan nyata dalam kehidupan sosial dan ekonomi. Materi SPLTV cocok untuk dijadikan konteks literasi karena memiliki representasi simbolik yang kompleks dan bisa dikaitkan dengan masalah ekonomi maupun social (Pratama & Harahap, 2022). Pengembangan instrumen dilakukan untuk memastikan keterukuran, validitas, dan reliabilitas instrumen sebagai alat asesmen literasi.

Dengan adanya instrumen ini, diharapkan proses evaluasi pembelajaran matematika di SMA menjadi lebih bermakna dan sejalan dengan arah kebijakan pendidikan nasional, khususnya dalam memperkuat implementasi AKM dan Kurikulum Merdeka. Hasil penelitian ini juga diharapkan memberikan kontribusi terhadap penyediaan perangkat asesmen yang mampu mendorong siswa untuk berpikir kritis, memahami konsep secara mendalam, dan menghubungkannya dengan konteks kehidupan nyata. Instrumen asesmen yang valid dan kontekstual menjadi penunjang utama implementasi kurikulum merdeka dan asesmen formatif yang bermakna (Rahman & Yuliana, 2023).

METODE

Metode Penelitian ini menggunakan metode pengembangan (*research and development*) yang bertujuan untuk menghasilkan instrumen tes literasi matematika yang valid, reliabel, dan layak digunakan untuk mengukur pemahaman konsep matematis siswa. Model pengembangan yang digunakan adalah 4D (Thiagarajan) yaitu pendefinisian, perancangan, pengembangan, dan penyebaran. Model ini telah digunakan dalam berbagai pengembangan instrumen asesmen pendidikan karena dinilai sistematis dan berorientasi produk (Rahayu & Widodo, 2021).

Tahap pendefinisian bertujuan untuk menentukan dan menetapkan tujuan dari pembelajaran. Pada tahap pendefinisian ini menganalisis terlebih dahulu perangkat tes kemudian menentukan tujuan adanya perangkat tes dan kendalanya. Tahap perancangan bertujuan untuk merancang draf materi pembelajaran berupa perangkat tes. Pada tahap ini digunakan; lembar validasi. Instrumen yang dibutuhkan dalam penelitian ini juga berupa soal. Lembar validasi yang berisi saran, kritikan dan masukan dari validator yang digunakan untuk melihat aspek validitas produk. Tahap pengembangan bertujuan untuk menghasilkan sebuah produk pengembangan melalui validasi dan penilaian ahli, tahap ini bertujuan untuk menghasilkan produk awal menjadi produk yang siap diuji dan direvisi, melalui beberapa proses yaitu pembuatan produk awal, uji validitas ahli, revisi produk berdasarkan validasi ahli, uji coba terbatas, uji lapangan, dan revisi produk berdasarkan uji coba. Teknik penyebaran dalam model pengembangan 4D ini merupakan tahap terakhir yang berfokus pada menyebarkan produk atau hasil proyek kepada pengguna atau target audiens

Subjek penelitian terdiri dari dua kelas X di SMAN 4 Banda Aceh, yaitu kelas A dan kelas B yang masing-masing kelas berjumlah 35 siswa. Kedua kelas tersebut digunakan untuk melihat perbandingan hasil pemahaman konsep matematis yang diperoleh setelah diberi instrumen tes literasi matematika. Instrumen yang dikembangkan berupa soal-soal literasi matematika pada topik SPLTV yang mengukur indikator pemahaman konsep.

Pada tahap pengembangan pada model pengembangan 4D, kelas A berfungsi sebagai subjek uji coba awal terhadap instrumen tes literasi matematika yang telah dikembangkan. Tujuan penggunaan kelas A adalah untuk menguji validitas, reliabilitas, tingkat kesukaran, serta daya pembeda dari butir-butir soal secara komprehensif. Data yang diperoleh dari kelas A memberikan dasar awal untuk menilai kualitas dan konsistensi instrumen sebelum disebarluaskan lebih luas. Sementara itu, kelas B digunakan dalam tahap penyebaran sebagai sarana triangulasi tempat atau pemeriksaan kembali. Fungsi kelas B adalah untuk memastikan bahwa instrumen yang telah diuji dan disempurnakan di kelas A dapat menghasilkan hasil yang konsisten ketika diterapkan pada kelompok lain yang serupa tetapi berbeda secara independen. Hal ini dilakukan untuk menilai kestabilan instrumen antar kelompok serta mendeteksi adanya potensi bias atau ketidaksesuaian dalam penggunaan lintas konteks.

Untuk mendukung proses triangulasi tersebut, digunakan uji T sampel independen. Uji ini bertujuan membandingkan rata-rata nilai pemahaman konsep matematis antara kelas A dan kelas B. Jika nilai signifikansi $\alpha > 0,05$, maka tidak terdapat perbedaan yang signifikan antara kedua kelas, sehingga dapat disimpulkan bahwa instrumen bekerja secara konsisten, terhadap tempat atau kelompok tertentu. Hal ini menjadi bukti bahwa instrumen layak digunakan dalam konteks pembelajaran yang lebih luas. Setelah pengembangan soal, instrumen diuji coba pada kedua kelas. Data hasil tes dianalisis secara kuantitatif menggunakan bantuan perangkat lunak SPSS versi 25. Total soal yang dikembangkan berjumlah 12 butir, yang mewakili seluruh indikator pemahaman konsep dan jenis konteks AKM. Instrumen yang dikembangkan dapat dikategorikan layak, valid, reliabel serta mendukung tujuan asesmen dalam Kurikulum Merdeka dan AKM apabila memenuhi beberapa kriteria, pertama, valid secara isi, yaitu seluruh butir soal memiliki koefisien korelasi positif dan nilai signifikansi $\alpha < 0,05$, kedua, reliabel dengan nilai Cronbach's Alpha $\geq 0,70$ yang menunjukkan konsistensi antar item, ketiga, memiliki tingkat kesukaran sedang, berada di rentang $0,31 < P \leq 0,70$, sehingga sesuai dengan kemampuan mayoritas siswa, keempat, daya beda baik dan sangat baik, dengan Corrected Item-Total Correlation (CITC) $\geq 0,30$, dan terakhir konsisten antar kelompok, dibuktikan dengan hasil uji T, $\alpha > 0,05$ sebagai syarat setara antar responden.

Instrumen yang digunakan dalam penelitian ini merupakan instrumen baru yang dikembangkan sendiri oleh peneliti, bukan merupakan adaptasi dari instrumen yang sudah ada. Pengembangan instrumen dilakukan berdasarkan kebutuhan asesmen literasi matematika dalam konteks implementasi Asesmen Kompetensi Minimum (AKM) dan Kurikulum Merdeka. Instrumen asesmen harus sesuai dengan prinsip diferensiasi dan konteks kehidupan nyata yang menjadi pilar Kurikulum Merdeka (Zahra & Munadi, 2024). Instrumen dirancang khusus untuk mengukur pemahaman konsep matematis siswa Sekolah Menengah Atas (SMA), dengan fokus pada materi Sistem Persamaan Linear Tiga Variabel (SPLTV). Instrumen yang dikembangkan terdiri dari soal-soal berbasis konteks (kontekstual) yang mengacu pada ciri khas AKM, yaitu berbasis stimulus, bentuk soal yang bervariasi, berorientasi pada pemahaman konsep, dan menggunakan table spesifikasi soal.

Berikut disajikan salah satu butir soal yang mewakili karakteristik instrumen akhir, berbasis konteks nyata dan mengukur indikator pemahaman konsep SPLTV.

“Di sebuah toko buah, harga 2 kg apel, 3 kg jeruk, dan 1 kg anggur adalah Rp85.000. Harga 1 kg apel, 1 kg jeruk, dan 2 kg anggur adalah Rp70.000. Sedangkan harga 3 kg apel, 2 kg jeruk, dan 2 kg anggur adalah Rp100.000. Tentukan harga masing-masing jenis buah.”

Tabel 1. Penjelasan Contoh Soal

Aspek	Deskripsi
Indikator	Menganalisis informasi dari soal cerita dan menyusunnya menjadi model SPLTV
Konten	Aljabar – Sistem Persamaan Linear Tiga Variabel (SPLTV)
Konteks	Personal (kehidupan sehari-hari)
Level Kognitif	L3 – C4
Tujuan Soal	Mengukur kemampuan siswa membentuk model matematis dan menyelesaikannya
Alasan Desain	Soal memuat stimulus berbasis konteks nyata sehingga mendukung literasi matematis sesuai AKM

Dalam penelitian pengembangan ini, data diambil dari hasil analisis kebutuhan lembar validasi. Validasi digunakan untuk mengukur kevalidan perangkat tes. Validasi menggunakan instrumen berupa lembar validasi. Validasi diperoleh dari 1 orang Dosen dan 1 orang Guru untuk mengkonfirmasi kevalidan perangkat tes. Dalam penelitian ini, sistem penskoran yang digunakan adalah skor dikotomis, yaitu skor 1 diberikan apabila siswa memberikan jawaban benar, baik secara perhitungan maupun konsep, serta sesuai dengan indikator pemahaman konsep matematis. skor 0 diberikan apabila siswa memberikan jawaban salah, tidak lengkap, atau tidak sesuai dengan prinsip atau pemahaman konsep yang diminta.

Tabel 2. Rubrik penskoran

Aspek Dinilai	Kriteria Skor 1	Kriteria Skor 0
1. Memilih metode yang tepat	Menyebut dan menerapkan metode substitusi atau eliminasi dengan benar	Tidak menyebutkan metode, atau menyebutkan tetapi salah/tidak sesuai
2. Menentukan satu variabel terlebih dulu	Menunjukkan langkah menentukan 1 variabel secara sistematis dan benar	Salah menentukan atau asal mencoba tanpa prosedur jelas

3. Substitusi/eliminasi ke dua persamaan lain	Melakukan substitusi/eliminasi ke dua persamaan lainnya dengan benar	Salah substitusi/eliminasi, atau salah hitung
4. Menentukan nilai semua variabel	Menentukan nilai x, y, z dengan benar	Salah satu atau lebih salah
5. Proses dan penjelasan	Penalaran jelas, langkah-langkah runtut, dan penggunaan operasi aljabar sesuai konsep	Langkah tidak sesuai konsep operasi aljabar atau asal menulis hasil

Analisis dilakukan secara kuantitatif dengan bantuan SPSS melalui melalui beberapa tahap. Pertama, akan dilakukan analisis validitas butir soal. Butir soal dikatakan valid apabila nilai signifikansi $\alpha < 0,05$ dan nilai koefisien korelasi positif. Hasil validitas digunakan untuk menentukan butir soal yang dapat dipertahankan, diperbaiki, atau dibuang.

Rumus uji validitas butir soal:

$$R_{xy} = \frac{(N\sum XY - (\sum X)(\sum Y))}{\sqrt{([N\sum X^2 - (\sum X)^2][N\sum Y^2 - (\sum Y)^2])}}$$

Keterangan:

R_{xy} = koefisien korelasi item

X = skor butir soal

Y = skor total

N = jumlah responden

Tabel 3. Kriteria Validitas

r	Interpretasi
$0,80 \leq r \leq 1,00$	Sangat tinggi
$0,60 \leq r < 0,80$	Tinggi
$0,40 \leq r < 0,60$	Cukup
$0,20 \leq r < 0,40$	Rendah
$0,00 \leq r < 0,20$	Sangat rendah / tidak valid

Sumber: (Arikunto, n.d.)

Pada tahap berikutnya, dilakukan pengujian reliabilitas instrumen. Pengujian dilakukan menggunakan metode *Cronbach's Alpha*. Instrumen dianggap reliabel jika nilai alpha $\geq 0,70$.

Rumus uji reliabilitas:

KR-20

$$r_{11} = \frac{k}{k-1} \left(1 - \frac{\sum pq}{\sigma^2} \right)$$

KR-21

$$r_{11} = \frac{k}{k-1} \left(1 - \frac{M(k-M)}{k\sigma^2} \right)$$

Keterangan:

r_{11} = Koefisien reliabilitas

k = Jumlah butir soal

p = Proporsi jawaban benar

q = Proporsi jawaban salah (= 1 - p)

M = Rata-rata skor

σ^2 = Varians total

Tabel 3. Tingkat Reliabilitas

r	Interpretasi
$0,80 \leq r \leq 1,00$	Sangat tinggi
$0,60 \leq r < 0,80$	Tinggi
$0,40 \leq r < 0,60$	Cukup
$0,20 \leq r < 0,40$	Rendah
$0,00 \leq r < 0,20$	Sangat rendah / tidak reliabel

Sumber: (Arikunto, n.d.)

Pada tahap ketiga, dilakukan analisis tingkat kesukaran soal. Pada analisis tingkat kesukaran, butir soal belum dikatakan baik apabila proporsi kesukaran tiap butir soal masih belum seimbang.

Rumus analisis tingkat kesukaran:

$$P = \frac{B}{JS}$$

Keterangan:

P = indeks kesukaran

B = jumlah siswa yang menjawab benar

JS = jumlah seluruh siswa peserta tes

Tabel 4. Kriteria Taraf Kesukaran

P	Kategori
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

Sumber: (Arikunto, n.d.)

Tahap keempat berfokus pada analisis daya beda. Daya beda yang baik mengindikasikan kemampuan soal dalam mendeteksi variasi pemahaman siswa terhadap konsep (Saputra & Hasan, 2022). Daya beda digunakan untuk mengetahui sejauh mana suatu soal mampu membedakan antara siswa berkemampuan tinggi dan rendah. Daya beda dihitung berdasarkan selisih proporsi siswa kelompok atas dan bawah yang menjawab soal dengan benar.

Rumus analisis daya beda:

$$D = \frac{BA}{JA} - \frac{BB}{JB}$$

Keterangan:

D = Indeks daya beda

BA = Jumlah peserta kelompok atas yang benar

BB = Jumlah peserta kelompok bawah yang benar

JA = Jumlah peserta kelompok atas

JB = Jumlah peserta kelompok bawah

Tabel 5. Kriteria Daya beda

D	Interpretasi
0,70 – 1,00	Sangat baik
0,40 – 0,69	Baik
0,20 – 0,39	Cukup

0,00 – 0,19	Jelek
< 0	Sangat jelek (soal harus dibuang)

Sumber: (Arikunto, n.d.)

Terakhir, dilakukan uji perbedaan menggunakan uji T sampel independen. Untuk mengetahui apakah terdapat perbedaan hasil pemahaman konsep matematis antara kelas A dan kelas B, digunakan uji T sampel independen. Uji ini digunakan untuk membandingkan rata-rata dua kelompok yang tidak saling berhubungan. Instrumen dinyatakan bekerja secara konsisten dan setara apabila hasil uji menunjukkan nilai signifikansi $\alpha > 0,05$, yang berarti tidak terdapat perbedaan hasil yang signifikan antara kedua kelompok.

HASIL DAN PEMBAHASAN

Data hasil tes dianalisis secara kuantitatif menggunakan bantuan perangkat lunak SPSS versi 25. Analisis yang dilakukan meliputi uji validitas, uji reliabilitas, uji Tingkat kesukaran, daya beda, serta uji T sampel independen untuk melihat perbedaan hasil antara dua kelompok siswa.

Pada tahap validasi ahli, proses validasi instrumen dilakukan melalui dua tahap, yaitu validasi oleh dosen ahli sebagai validator pertama dan guru mata pelajaran matematika sebagai validator kedua. Validator pertama memberikan masukan untuk mengganti beberapa stimulus dan model soal karena beberapa stimulus dinilai terlalu abstrak dan tidak langsung mengaitkan konsep dengan situasi nyata. Selanjutnya, validator diminta melakukan penyelarasan antara indikator, stimulus, dan pilihan jawaban. Akhirnya, setelah direvisi dan ditunjukkan kembali, validator menyatakan bahwa instrumen telah valid dan layak untuk diujicobakan.

Hal yang sama, proses validasi dilanjutkan oleh validator kedua. Validator kedua hanya memberikan saran perubahan pada format grafik dan tata letak soal agar lebih mudah dipahami oleh siswa. Akhirnya, setelah direvisi dan ditunjukkan kembali, validator kedua juga menyatakan bahwa instrumen telah valid dan layak untuk diujicobakan.

Setelah merevisi sesuai dengan masukan dari kedua validator diyakini instrument yang dikembangkan valid dari aspek teoretis dan pedagogis, sehingga siap untuk diuji coba. Hasil uji coba instrument ini kepada subjek yang telah ditentukan akan mengkonfirmasi kevalidan instrument yang dikembangkan.

Selanjutnya dilakukan uji validitas butir soal untuk memastikan setiap item instrumen benar-benar mengukur indikator yang telah ditetapkan. Hasil uji validitas pada kedua kelas menunjukkan bahwa seluruh butir soal (Soal 1 hingga Soal 12) memiliki nilai koefisien korelasi (r) yang berada di atas angka 0.30 dengan nilai signifikansi < 0.05 , yang berarti seluruh item dinyatakan valid secara statistik. Bahkan, sebagian besar soal menunjukkan tingkat signifikansi pada level 0.01, yang membuktikan korelasi sangat signifikan. Validitas yang tinggi menunjukkan bahwa item dalam instrumen telah sesuai dengan indikator pemahaman konsep matematis, validitas isi penting dalam konteks AKM karena mendukung pengukuran kompetensi secara autentik (Hasanah & Wulandari, 2023), seperti kemampuan siswa dalam memahami, mengaplikasikan, dan menganalisis konsep dalam konteks masalah nyata. Hasil ini sejalan dengan penelitian oleh Wulandari & Wutsqa (E. S. Wulandari & Wutsqa, 2020), yang menemukan bahwa instrumen literasi matematika yang valid secara isi mampu merepresentasikan kompetensi numerasi dan pemahaman konsep siswa. Selain itu, Pratiwi et al. (Pratiwi et al., 2023) juga menegaskan pentingnya validitas isi dalam mengembangkan instrumen berbasis konteks literasi.

Selanjutnya, dilakukan uji reliabilitas guna melihat konsistensi jawaban peserta tes terhadap instrumen yang dikembangkan. Hasil uji reliabilitas menunjukkan nilai Cronbach's Alpha pada kelas A 0,80 dan pada kelas B sebesar 0,84, yang menunjukkan bahwa instrumen memiliki konsistensi yang digunakan untuk mengukur pemahaman konsep matematis. Reliabilitas tinggi menunjukkan bahwa instrumen menghasilkan hasil yang konsisten ketika diujikan pada kelompok siswa yang berbeda tetapi setara. Nilai $\alpha > 0.70$ menunjukkan konsistensi yang tinggi pada instrumen tes kognitif (Utami & Ramadhan, 2021). Penelitian dari Arifin & Sugiman (Arifin & Sugiman, 2021) menegaskan bahwa reliabilitas tinggi adalah syarat mutlak untuk menjamin stabilitas pengukuran dalam asesmen kemampuan literasi. Hal ini juga dikonfirmasi dalam studi oleh Oktaviani & Maulidian (Oktaviani & Maulidian, 2022) yang mengembangkan instrumen numerasi literatif dengan hasil reliabilitas $\alpha = 0,82$.

Analisis dilanjutkan dengan uji Tingkat kesukaran soal untuk mengetahui sejauh mana tingkat kemudahan atau kesulitan tiap butir soal. Hasil uji Tingkat kesukaran pada kelas A dan kelas B menunjukkan bahwa seluruh butir soal berada dalam kategori sedang, dengan nilai indeks kesukaran berada pada rentang $0,31 < P \leq 0,70$. Hal ini mengindikasikan bahwa tingkat kesulitan setiap soal telah sesuai dengan karakteristik kemampuan siswa secara umum, sehingga instrumen dapat dianggap seimbang dan proporsional dari segi kesukaran.

Pada tahap selanjutnya, dilakukan analisis daya pembeda guna melihat kelayakan setiap soal. Soal-soal dalam instrumen ini telah berhasil menstimulasi kemampuan berpikir tingkat tinggi. Studi oleh Yusnia & Anwar (Yusnia & Anwar, 2020) menyatakan bahwa daya beda yang tinggi sangat penting untuk membedakan kemampuan konseptual siswa dalam konteks literasi.

Tabel 6. Rekapitulasi Daya beda Kelas A dan Kelas B

No	Kode Soal	CITC Kelas A	Kategori A	CITC Kelas B	Kategori B	Keterangan Rekomendasi
1	P1	0.68	Sangat Baik	0.43	Sangat Baik	Layak digunakan
2	P2	0.45	Sangat Baik	0.51	Sangat Baik	Layak digunakan
3	P3	0.39	Baik	0.47	Sangat Baik	Layak digunakan
4	P4	0.38	Baik	0.38	Baik	Layak digunakan
5	P5	0.36	Baik	0.53	Sangat Baik	Layak digunakan
6	P6	0.33	Baik	0.54	Sangat Baik	Layak digunakan
7	P7	0.42	Sangat Baik	0.42	Sangat Baik	Layak digunakan
8	P8	0.47	Sangat Baik	0.47	Sangat Baik	Layak digunakan
9	P9	0.37	Baik	0.67	Sangat Baik	Layak digunakan
10	P10	0.34	Baik	0.57	Sangat Baik	Layak digunakan
11	P11	0.41	Sangat Baik	0.46	Sangat Baik	Layak digunakan

12	P12	0.68	Sangat Baik	0.53	Sangat Baik	Layak digunakan
----	-----	------	-------------	------	-------------	-----------------

Berdasarkan hasil uji reliabilitas instrumen dengan pendekatan *Corrected Item-Total Correlation* (CITC), seluruh butir soal menunjukkan nilai yang berada pada kategori baik dan sangat baik, baik pada kelas A maupun kelas B. Nilai CITC untuk setiap butir soal pada Kelas A berkisar antara 0.33 hingga 0.68, sedangkan pada Kelas B berkisar antara 0.38 hingga 0.67. Tidak terdapat butir soal dengan nilai CITC di bawah 0.30, yang mengindikasikan bahwa seluruh soal memiliki korelasi yang cukup kuat terhadap total skor. Dengan demikian, semua soal dinyatakan layak digunakan karena telah memenuhi syarat sebagai butir yang mampu membedakan kemampuan siswa secara konsisten antar kelas.

Terakhir, dilaksanakan uji T sampel independen. Untuk menguji perbedaan hasil tes antara dua kelompok siswa yang menjadi sampel penelitian. Untuk mengetahui kelayakan instrumen dari segi kemampuan membedakan pemahaman konsep antar kelompok, dilakukan uji T sampel independen terhadap hasil tes literasi siswa dari kelas A dan kelas B. Sebelum dilakukan uji T sampel independen, terlebih dahulu dilakukan uji prasyarat, yaitu uji normalitas dan uji homogenitas varians.

Berdasarkan hasil uji normalitas menggunakan metode *Shapiro-Wilk*, data dari kelas A dan kelas B memiliki nilai signifikansi $\alpha > 0,05$, yang menunjukkan bahwa data berdistribusi normal. Hasil uji homogenitas juga menunjukkan nilai signifikansi $\alpha > 0,05$, yang berarti varians kedua kelompok adalah homogen. Dengan demikian, kedua syarat analisis parametrik telah terpenuhi, sehingga pengujian dapat dilanjutkan dengan menggunakan uji T sampel independen untuk mengetahui apakah terdapat perbedaan pemahaman konsep matematis antara kedua kelas.

Tabel 7. Uji T sampel independen

		<i>Levene's Test for Equality of Variances</i>		<i>t-test for Equality of Means</i>						
		<i>F</i>	<i>Sig.</i>	<i>t</i>	<i>df</i>	<i>Sig. (2-tailed)</i>	<i>Mean Difference</i>	<i>Std. Error Difference</i>	<i>95% Confidence Interval of the Difference</i>	
Nilai	<i>Equal variances assumed</i>	0.729	0.396	-0.417	68	0.678	-0.34286	0.82293	-1.98499	1.29928
	<i>Equal variances not assumed</i>			-0.417	67.549	0.678	-0.34286	0.82293	-1.98519	1.29947

Hasil analisis menunjukkan nilai signifikansi α sebesar $0,68 > 0,05$, yang berarti tidak terdapat perbedaan yang signifikan antara nilai rata-rata kedua kelas. Dalam pengembangan instrumen literasi, uji T digunakan untuk menguji kesetaraan soal pada kelompok yang berbeda

(R. Fauziah & Yulianto, 2023). Hal ini menunjukkan bahwa instrumen tes bekerja secara merata terhadap kedua kelompok siswa. Tidak adanya perbedaan yang signifikan ini dapat ditafsirkan bahwa instrumen sudah layak digunakan secara luas karena mampu memberikan hasil yang konsisten terhadap berbagai kelompok siswa yang berbeda. Hasil ini juga menunjukkan bahwa soal dapat digunakan sebagai alat ukur pemahaman konsep dalam konteks pembelajaran Kurikulum Merdeka dan AKM. Instrumen berbasis AKM yang telah lolos uji T dapat digunakan untuk menghindari bias antar kelompok siswa dalam evaluasi (Febriani & Azzahra, 2024). Ini mengindikasikan bahwa instrumen konsisten terhadap kelompok tertentu, dan mampu mengukur pemahaman konsep secara objektif. Kelayakan ini penting dalam konteks asesmen literasi matematika di berbagai sekolah. Penelitian terbaru oleh Fauziah et al. (I. Fauziah et al., 2024) menekankan pentingnya pengujian antarkelompok untuk menjamin kesetaraan asesmen dalam skala kelas besar.

KESIMPULAN

Penelitian ini berhasil mengembangkan instrumen tes literasi matematika berbasis konteks yang valid, reliabel, dan layak digunakan untuk mengukur pemahaman konsep matematis siswa SMA, khususnya pada materi SPLTV. Instrumen ini disusun untuk mendukung penguatan kompetensi numerasi dalam AKM serta meningkatkan kemampuan literasi matematika yang sesuai dengan tuntutan abad ke-21. Hasil uji validitas menunjukkan bahwa seluruh butir soal dinyatakan valid dan uji reliabilitas menghasilkan nilai Cronbach's Alpha yang mengindikasikan reliabilitas tinggi. Selain itu, tingkat kesukaran soal berada dalam kategori sedang, daya beda tergolong baik dan sangat baik, maka tidak terdapat butir soal yang perlu dibuang. Instrumen ini juga memenuhi uji normalitas dan homogenitas, serta hasil uji T sampel independen menunjukkan nilai signifikansi 0,678 ($> 0,05$), yang berarti tidak terdapat perbedaan signifikan antara pemahaman konsep siswa kelas A dan B. Dengan demikian, instrumen dapat digunakan secara konsisten oleh guru sebagai alat evaluasi pembelajaran maupun sebagai bagian dari asesmen nasional berbasis literasi matematika. Implikasi dari penelitian ini adalah sekolah dapat memanfaatkan instrumen tersebut sebagai alat ukur alternatif dalam menilai kemampuan berpikir siswa secara konseptual dan kontekstual. Selain itu, guru dapat menggunakan hasil tes sebagai dasar dalam menyusun strategi pembelajaran yang lebih bermakna.

REFERENSI

- Arifin, M., & Sugiman. (2021). Pengembangan instrumen penilaian literasi matematika berbasis konteks kehidupan siswa SMP. *Jurnal Riset Pendidikan Matematika*, 8(2), 150–161.
- Arikunto, S. (n.d.). *Dasar-Dasar Evaluasi*.
- Fauziah, I., Fitria, R., & Mukhlisin, M. (2024). Pengembangan instrumen asesmen diagnostik berbasis literasi numerasi pada siswa SMA. *Jurnal Pendidikan Matematika Dan Sains*, 12(1), 22–32.
- Fauziah, R., & Yulianto, D. (2023). Analisis keadilan instrumen berbasis AKM menggunakan uji-t. *Jurnal Evaluasi Pembelajaran*, 9(1), 51–60.
- Febriani, L., & Azzahra, M. (2024). Analisis keadilan instrumen AKM dalam pengukuran literasi matematika siswa SMA. *Jurnal Inovasi Evaluasi Pendidikan*, 6(1), 61–69.
- Fitriana, A., & Sugiyono, M. (2023). Tes Diagnostik Berbasis Literasi pada Materi Fungsi. *Jurnal Ilmiah Pendidikan Matematika*, 9(1), 20–30.
- Handayani, R., & Rahayu, R. S. (2022). Pengembangan Soal Numerasi AKM Berbasis Masalah pada Materi Bangun Datar. *Jurnal Pendidikan Matematika*, 7(2), 77–85.

- Hasanah, U., & Wulandari, T. (2023). Validasi instrumen numerasi berbasis AKM untuk siswa SMA. *Jurnal Inovasi Evaluasi Pendidikan*, 5(1), 23–32.
- Hidayat, A., & Rahmawati, S. (2021). Pengembangan instrumen tes literasi matematika untuk siswa SMA pada materi fungsi. *Jurnal Pendidikan Matematika*, 10(2), 134–144.
- Hidayat, M. T., & Fadilah, N. (2024). Peran literasi matematika dalam penguatan Kurikulum Merdeka. *Jurnal Kurikulum Dan Evaluasi Pendidikan*, 12(1), 25–34.
- Hidayat, R., & Lestari, F. (2024). Pengembangan Instrumen AKM Berbasis Konteks Multilevel. *Jurnal Inovasi Pendidikan*, 10(1), 45–58.
- Kebudayaan, K. P. dan. (2020). *Panduan teknis Asesmen Kompetensi Minimum (AKM)*. Pusat Asesmen dan Pembelajaran, Balitbang.
- Latifah, F., & Ramdhan, D. (2023). Analisis kebutuhan asesmen berbasis kontekstual dalam pembelajaran matematika. *Jurnal Kajian Pendidikan Matematika*, 8(1), 55–63.
- Oktaviani, M., & Maulidian, H. (2022). Pengembangan instrumen numerasi untuk meningkatkan literasi matematika siswa SMP. *Jurnal Numerasi*, 9(1), 45–54.
- Pratama, A. D., & Harahap, R. A. (2022). Konteks SPLTV dalam pengembangan soal literasi matematika. *Jurnal Didaktika Matematika*, 9(1), 51–59.
- Pratiwi, A. E., Maulana, H., & Sundari, L. (2023). Pengembangan soal literasi matematika berbasis konteks PISA pada materi aljabar. *Jurnal Edukasi Matematika*, 4(1), 1–10.
- Rahayu, A., & Widodo, S. A. (2021). Pengembangan instrumen asesmen pembelajaran berbasis model 4D untuk siswa sekolah menengah. *Jurnal Evaluasi Pendidikan*, 12(1), 45–53.
- Rahman, A., & Yuliana, N. (2023). Pengembangan asesmen formatif berbasis konteks dalam mendukung Kurikulum Merdeka. *Jurnal Evaluasi Dan Pembelajaran*, 13(1), 44–52.
- Saputra, H., & Hasan, N. (2022). Daya pembeda dalam pengembangan soal berbasis konteks matematika. *Jurnal Pendidikan Dan Evaluasi*, 14(2), 88–96.
- Setiawan, D., & Fitriyani, R. (2021). Literasi matematika sebagai kompetensi abad 21 dalam pembelajaran berbasis konteks. *Jurnal Numeracy and Education*, 5(2), 98–107.
- Sulastri, R., & Rofiah, I. (2022). Transformasi asesmen nasional: Dari hafalan ke kompetensi. *Jurnal Kebijakan Pendidikan*, 8(1), 14–22.
- Utami, A. P., & Ramadhan, R. (2021). Uji reliabilitas dan konsistensi instrumen tes berbasis HOTS pada pembelajaran matematika. *Jurnal Pengukuran Dan Evaluasi Pendidikan Indonesia*, 11(2), 105–113.
- Wulandari, D., & Saputra, A. (2021). Keterampilan pemecahan masalah dalam literasi matematika berbasis konteks. *Jurnal Pendidikan Matematika Indonesia*, 6(2), 81–89.
- Wulandari, E. S., & Wutsqa, D. U. (2020). Pengembangan instrumen literasi matematika berbasis konteks budaya lokal. *Jurnal Ilmiah Pendidikan Matematika*, 9(1), 77–88.
- Yusnia, M., & Anwar, L. (2020). Analisis karakteristik soal literasi matematika berdasarkan tingkat kesukaran dan daya pembeda. *Delta: Jurnal Ilmiah Pendidikan Matematika*, 8(1), 11–20.
- Zahra, N. L., & Munadi, S. (2024). Pengembangan asesmen berbasis diferensiasi dalam Kurikulum Merdeka. *Jurnal Inovasi Pendidikan Indonesia*, 6(1), 17–26.