

Pengembangan Soal Tes Numerasi Berbasis *Hierarchical Content* untuk Fase A Elemen Bilangan dan Aljabar

Hikmah Ramdhani Putri¹, Lalu Hamdian Affandi², Baiq Yuni Wahyuningsih³, Iva Nurmawanti⁴, Vivi Rachmatul Hidayati⁵, Riska Wulandari⁶

^{1,2,3,4,5} Pendidikan Guru Sekolah Dasar, FKIP Universitas Mataram, Mataram, Indonesia

⁶ Mahasiswa Pendidikan Guru Sekolah Dasar, FKIP Universitas Mataram, Mataram, Indonesia

hikmah.ramdhaniputri@unram.ac.id

Diterima:25-11-2025; Direvisi:27-11-2025; Dipublikasi: 30-11-2025

Abstract

Numeracy is an essential skill that students must master. The government administers the Minimum Competency Assessment (AKM) for Numeracy to map students' numeracy abilities in Indonesia. However, in practice, the results of AKM have not fully guided teachers in determining appropriate instructional follow-up, as they only provide information on students' proficiency levels without indicating which specific content has not yet been mastered. This study aims to develop a valid and reliable numeracy test instrument based on hierarchical content for primary school students. The development process followed several stages: defining the test, selecting a scaling method, constructing test items, conducting item trials, revising items, and preparing the final product. The instrument was tested on Grade 3 students at some Elementary School. Analysis of the learning outcomes produced five hierarchical content areas within the Numbers and Algebra element of Phase A, which were developed into ten test items. Rasch Model analysis indicated high reliability, and seven items were found to be valid (items 1, 3, 4, 7, 8, 9, and 10). This instrument is expected to help teachers identify not only students' numeracy levels but also the prerequisite content that has not been mastered, thereby supporting more effective formative assessment.

Keywords: Numeration test; numeration; AKM numerasi; hierarchical content

Abstrak

Numerasi merupakan keterampilan penting yang harus dikuasai siswa. Pemerintah melaksanakan Asesmen Kompetensi Minimum (AKM) Numerasi untuk memetakan kemampuan numerasi siswa di Indonesia. Namun, di lapangan hasil AKM belum sepenuhnya membantu guru menentukan tindak lanjut pembelajaran karena hanya menunjukkan level kemampuan tanpa memberikan informasi mengenai konten apa yang belum dikuasai siswa. Untuk menjawab kebutuhan tersebut, penelitian ini bertujuan mengembangkan instrumen tes numerasi berbasis *hierarchical content* yang valid dan reliabel untuk siswa SD/MI. Instrumen disusun dengan mengikuti tahapan penelitian pengembangan, yaitu mendefinisikan tes, memilih metode penskalaan, menyusun butir tes, menguji butir, merevisi, dan mempersiapkan produk akhir. Uji coba dilakukan pada siswa jenjang Sekolah Dasar. Hasil analisis terhadap capaian pembelajaran menghasilkan lima konten hirarkis pada elemen Bilangan dan Aljabar Fase A yang dikembangkan menjadi sepuluh butir soal. Analisis menggunakan Model Rasch menunjukkan bahwa instrumen yang dikembangkan memiliki reliabilitas tinggi dan tujuh butir soal dinyatakan valid (butir 1, 3, 4, 7, 8, 9, dan 10). Instrumen ini diharapkan membantu guru tidak hanya mengetahui level kemampuan numerasi siswa, tetapi juga mengidentifikasi konten prasyarat yang belum dikuasai sebagai dasar perbaikan pembelajaran.

Kata Kunci: Numerasi; tes numerasi; AKM numerasi; hierarchical content

1. PENDAHULUAN

Numerasi merupakan kemampuan menggunakan angka, data, maupun simbol matematika, serta pengetahuan dan kecakapan dalam penarikan suatu keputusan yang berkaitan dengan masalah nyata di kehidupan sehari-hari (Gerakan Literasi Nasional, 2017). Sejalan dengan itu, Baharuddin mendefinisikan numerasi sebagai pengetahuan dan keterampilan yang terdiri dari (a) penggunaan berbagai jenis angka dan simbol yang terkait dengan matematika dasar untuk memecahkan masalah praktis dalam konteks kehidupan sehari-hari yang berbeda (b) analisis informasi yang disajikan dalam berbagai bentuk (grafik, tabel, peta, dll), dan (c) menggunakan interpretasi untuk memprediksi dan membuat keputusan (Baharuddin et al., 2021; Kemdikbud, 2023a). Berdasarkan definisi-definisi tersebut, jelas bahwa numerasi merupakan keterampilan yang penting yang harus dikuasai oleh siswa.

Hasil numerasi siswa Indonesia dapat tergambar dalam tes PISA. Tahun 2022 skor PISA Indonesia mengalami penurunan skor sebanyak 13 poin dari tahun 2018 meskipun peringkatnya naik. Penurunan skor ini juga terjadi secara internasional akibat adanya *learning loss* (Kemdikbud, 2023b). Meskipun naik peringkat, skor PISA ini belum menunjukkan hasil yang memuaskan karena Indonesia masih menduduki peringkat 10 besar dari bawah. Artinya, kemampuan numerasi siswa kita juga perlu ditingkatkan.

Sebagai bagian dari Merdeka Belajar, Kemendikbud menerapkan program terbaru yang disebut dengan Asesmen Nasional sebagai pengganti Ujian Nasional yang mulai diterapkan pada tahun 2021. Penilaian Asesmen Nasional meliputi tiga aspek, yakni Asesmen Kompetensi Minimum (AKM), Survei karakter, serta Survei lingkungan belajar (Novita et al., 2021). Asesmen Kompetensi Minimum (AKM) adalah penilaian kompetensi mendasar yang dibutuhkan seluruh siswa agar dapat mengembangkan kemampuan diri sendiri serta berperan aktif dalam masyarakat pada kegiatan yang bernilai positif (Mendikbud, 2020). AKM digunakan untuk mengukur kemampuan kognitif siswa dimana aspek yang diukur adalah kemampuan literasi membaca dan literasi numerasi (Novita et al., 2021); (Cahyana, 2020). Asesmen Kompetensi Minimum (AKM) dirancang untuk menghasilkan informasi akurat guna perbaikan kualitas belajar-mengajar yang bertujuan meningkatkan hasil belajar murid. AKM menghasilkan informasi untuk memantau (a) perkembangan mutu dari waktu ke waktu, dan (b) kesenjangan antar bagian di dalam sistem Pendidikan (Kusmaharti et al., 2022). Jika merujuk pada tujuan pelaksanaan AKM, maka mestinya hasil AKM bisa dimanfaatkan untuk keperluan formatif, terutama memberikan arah yang jelas tentang materi dan tujuan pembelajaran yang perlu dipelajari dengan lebih baik untuk menunjang penguasaan siswa terhadap capaian pembelajaran

Berdasarkan studi pendahuluan awal, selama ini untuk meningkatkan capaian kompetensi numerasi siswa, guru memanfaatkan soal-soal numerasi yang ada pada Platform Merdeka Mengajar (PMM). Pemanfaatan soal numerasi di PMM yang bersifat formatif tidak memberikan arah perbaikan pembelajaran yang cukup jelas bagi guru. Hal ini disebabkan karena soal di PMM hanya mampu memberikan gambaran tentang level kemampuan siswa tanpa ada petunjuk tentang materi dan proses berpikir seperti apa

yang sudah dan belum dikuasai siswa. Pemaparan tersebut mengindikasikan bahwa soal numerasi yang disediakan di dalam PMM belum cukup membantu guru untuk secara optimal menerapkan asesmen formatif. Soal-soal tersebut hanya bisa digunakan untuk mendeteksi level kemampuan siswa. Sedangkan untuk melakukan asesmen formatif yang berorientasi pada perbaikan pembelajaran dibutuhkan soal yang mampu menunjukkan materi dan proses berfikir yang perlu dikuasai oleh siswa dalam mencapai tujuan belajarnya. Untuk mengatasi masalah tersebut, dibutuhkan soal yang dikembangkan berdasarkan prinsip *hierarchical content*.

Hierarchical Content dalam konteks evaluasi dan pembelajaran, mengacu pada struktur atau pendekatan yang menempatkan materi atau tugas dalam urutan hierarki berdasarkan tingkat kesulitan atau kompleksitas kognitif. Menurut (Anderson & Krathwohl, 2021), pendekatan ini sangat erat kaitannya dengan Bloom's Taxonomy, yang mengelompokkan keterampilan kognitif menjadi tingkatan mulai dari pemahaman dasar hingga analisis dan evaluasi yang lebih kompleks. Dalam konteks tes, pendekatan hierarkis ini memungkinkan penyusunan soal yang diatur berdasarkan tingkat kesulitan, mulai dari soal yang mengukur pengetahuan dasar hingga soal yang menuntut analisis dan pemecahan masalah tingkat lanjut. *Hierarchical content* yang dimaksud disini adalah konten-konten yang disusun secara hierarkis atau berjenjang. Kita tahu bahwa konten-konten dalam matematika sebagian besar bersifat hierarkis atau terurut. Artinya, untuk menguasai konten tertentu, ada konten-konten lain sebagai prasyarat. Contoh sederhananya, untuk bisa menguasai konsep perkalian siswa harus memahami konsep penjumlahan, dan untuk menguasai konsep penjumlahan, siswa harus memahami konsep membilang.

Hierarchical Content memudahkan pendidik untuk menilai kemampuan siswa secara lebih spesifik. Misalnya, siswa yang mampu menyelesaikan soal dasar tetapi kesulitan dengan soal yang lebih kompleks menunjukkan bahwa mereka mungkin memahami konsep dasar tetapi belum memiliki kemampuan untuk mengaplikasikannya dalam konteks yang lebih rumit. Oleh karena itu, pendekatan ini memungkinkan diagnosis yang lebih tepat terhadap area mana yang memerlukan intervensi atau perbaikan. Tes numerasi yang disusun secara hierarkis berdasarkan konten yang ada pada dokumen kurikulum akan lebih membantu guru dalam mendeteksi konten mana yang perlu dikuasai siswa dalam upayanya mencapai tujuan pembelajaran. Jika konten yang belum dikuasai siswa terdeteksi, maka guru akan lebih mudah menentukan langkah untuk menguatkan konten yang belum dikuasai siswa tersebut. Selain itu, tes numerasi yang disusun hierarkis berdasarkan konten, akan lebih efektif dan efisien dari segi waktu pengerjaannya. Siswa yang sudah tidak bisa atau salah dalam menjawab sebuah soal tidak perlu menjawab soal-soal berikutnya karena untuk bisa mengerjakan soal berikutnya siswa harus bisa menjawab dengan benar soal sebelumnya.

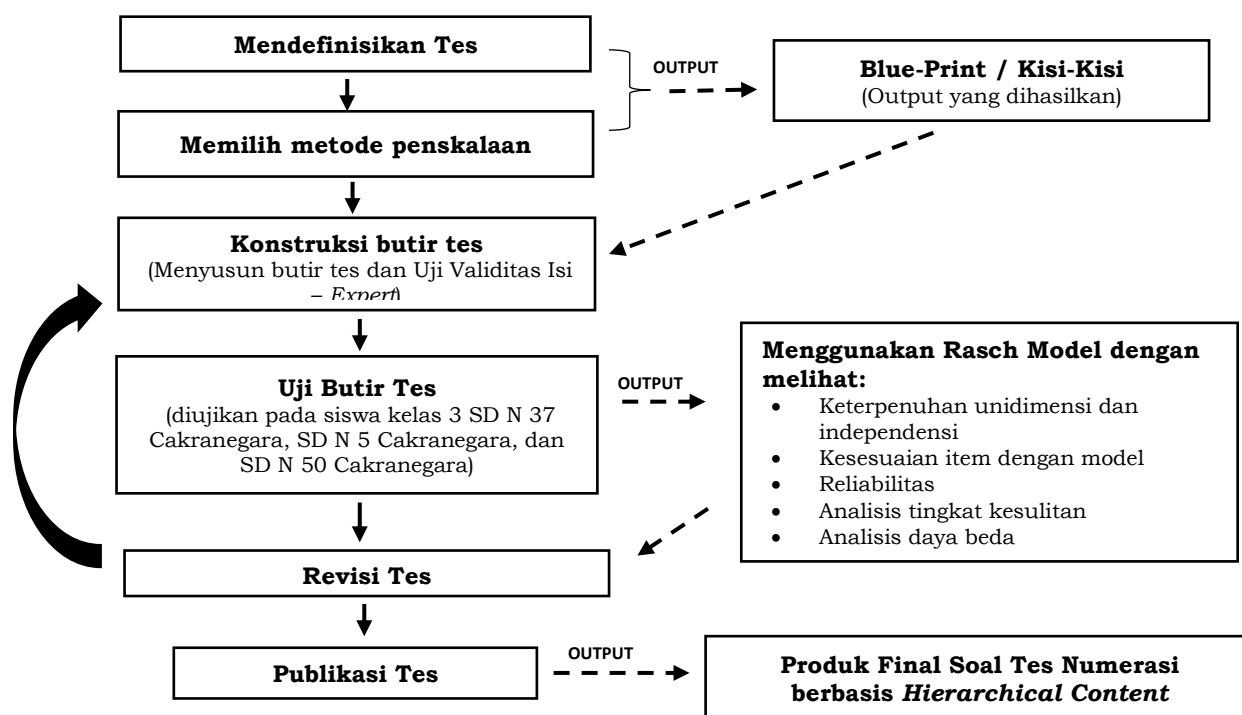
Kebaruan penelitian ini terletak pada pengembangan instrumen tes numerasi yang disusun berdasarkan *hierarchical content* sesuai struktur capaian pembelajaran Fase A. Berbeda dengan instrumen numerasi yang telah tersedia, yang umumnya hanya menunjukkan level kemampuan siswa, instrumen ini dirancang untuk mengidentifikasi

secara spesifik konten prasyarat yang belum dikuasai siswa. Dengan demikian, instrumen yang dikembangkan dapat memberikan informasi diagnostik yang lebih tepat guna untuk mendukung pelaksanaan asesmen formatif dalam Kurikulum Merdeka.

2. METODE PELAKSANAAN

Penelitian ini menggunakan desain penelitian pengembangan (McKenney & Reeves, 2012) mendefinisikan penelitian pengembangan sebagai pendekatan sistematis yang menghasilkan solusi praktis berupa produk, model, atau perangkat pembelajaran, sekaligus menghasilkan pemahaman teoretis tentang proses desain dan implementasinya. Penelitian pengembangan dilakukan secara berulang melalui siklus desain, evaluasi, dan revisi berulang untuk menjamin bahwa produk yang dihasilkan valid, reliabel, dan layak digunakan. Adapun tahapan pengembangan soal tes numerasi yang akan dilakukan dalam penelitian ini menggunakan tahapan konstruksi tes menurut Gregory, yaitu (1) mendefinisikan tes; (2) memilih metode penskalaan; (3) konstruksi butir tes; (4) uji butir tes; (5) revisi tes; (6) publikasi tes (Gregory, 2016).

Model ini sesuai untuk pengembangan instrumen numerasi karena menekankan analisis validitas dan reliabilitas pada setiap tahap, sehingga kualitas butir dapat teruji secara psikometrik dan mendukung pengembangan instrumen yang akurat serta layak digunakan dalam asesmen formatif (Gregory, 2016). Adapun diagram alir dari pengembangan soal tes pada penelitian ini dapat dilihat pada Gambar 1.



Gambar 1 Diagram Alir Pengembangan Soal tes

3. HASIL DAN PEMBAHASAN

3.1 Hasil

Pengembangan instrumen tes numerasi berbasis *hierarchical content* dilakukan melalui enam tahap: (1) mendefinisikan tes, (2) memilih metode penskalaan, (3) konstruksi butir soal, (4) uji butir soal, (5) revisi, dan (6) publikasi. Tahap pendefinisian tes dilakukan dengan menganalisis capaian pembelajaran (CP) elemen Bilangan dan Aljabar Fase A. Pada tahap tersebut dihasilkan 15 tujuan pembelajaran yang disusun secara hierarkis dengan mencakup lima konten utama: bilangan cacah sampai 100, operasi penjumlahan dan pengurangan bilangan sampai 20 menggunakan benda konkret, pecahan sederhana ($\frac{1}{2}$ dan $\frac{1}{4}$), serta pola bukan bilangan. Selanjutnya pada tahap penskalaan, digunakan Model Rasch sebagai dasar analisis butir, dengan penyusunan kisi-kisi yang mencakup sepuluh indikator soal.

Dari 15 tujuan pembelajaran yang diperoleh pada tahap sebelumnya, tidak seluruh tujuan tersebut dikembangkan menjadi indikator tersendiri. Hal ini disebabkan karena beberapa tujuan pembelajaran memiliki keterkaitan konsep yang kuat sehingga dapat diukur melalui satu indikator yang sama. Misalnya, tujuan pembelajaran terkait membandingkan bilangan tidak dijadikan indikator terpisah karena kemampuan tersebut telah terukur ketika siswa diminta mengurutkan bilangan. Dalam proses mengurutkan, siswa secara otomatis melakukan kegiatan membandingkan untuk menentukan posisi bilangan. Contoh ini menunjukkan bahwa satu indikator dapat merepresentasikan lebih dari satu tujuan pembelajaran yang berada pada satu rumpun kemampuan. Pendekatan ini dilakukan untuk menghindari redundansi indikator dan memastikan bahwa instrumen tetap efisien, fokus, serta sesuai dengan struktur *hierarchical content* yang dianalisis. Soal disusun dengan beragam bentuk: pilihan ganda, pilihan ganda kompleks, isian singkat, menjodohkan, dan uraian. Berbagai bentuk soal digunakan karena instrumen ini disesuaikan dengan format soal AKM Numerasi. Penyesuaian ini dilakukan agar instrumen relevan dengan standar asesmen nasional dan mampu mengukur kemampuan numerasi secara lebih komprehensif. Adapun kisi-kisi soal tes dapat dilihat pada Tabel 1.

Tabel 1. Kisi-kisi Soal Tes Berbasis *Hierarchical Content*

Konten	Indikator	Indikator Soal	Bentuk Soal	No. item
Bilangan Cacah sampai 100	Siswa dapat membaca bilangan cacah sampai 100	Konteks soal: antrean Diberikan nomor antrean, siswa menjodohkan angka yang ditunjukkan dengan cara membacanya	Menjodohkan	1
	Siswa dapat menentukan nilai tempat bilangan cacah sampai 100	Konteks soal: tusuk sate diberikan beberapa bilangan, siswa menjodohkan berdasarkan nilai tempatnya dengan benar	Menjodohkan	3
	Siswa bisa mengurutkan bilangan cacah sampai 100	Konteks soal: antrean Diberikan beberapa nomor antrean, siswa menentukan urutan yang akan dipanggil terlebih dahulu	Isian Singkat	2

Konten	Indikator	Indikator Soal	Bentuk Soal	No. item
	Siswa dapat mengurai bilangan cacah sampai 100	Konteks soal: tusuk sate Diberikan sebuah bilangan, siswa menentukan pernyataan yang benar terkait bilangan tersebut	Pilihan Ganda Kompleks	4
Penjumlahan bilangan sampai 20 menggunakan benda konkret	Siswa dapat melakukan operasi penjumlahan dengan bantuan benda konkret yang banyaknya sampai 20	Konteks soal: kelereng Diberikan banyaknya kelereng milik dua anak, siswa menentukan jumlah dari kelereng dua anak tersebut	Pilihan Ganda Biasa	5
	Siswa memahami makna simbol “=” dengan menggunakan gambar	Konteks soal: domino Diberikan sebuah domino, siswa menentukan domino lain yang jumlah titiknya sama dengan domino yang disajikan	Pilihan Ganda Kompleks	7
Pengurangan bilangan sampai 20 menggunakan benda konkret	Siswa dapat menentukan selisih dari banyaknya dua benda yang banyaknya sampai 20	Konteks soal: kelereng Diberikan hasil selisih dari kelereng dua anak, siswa menentukan benar atau salah, dan menjelaskannya	Uraian	6
Pecahan sederhana ($\frac{1}{2}$ dan $\frac{1}{4}$)	Siswa dapat membagi sebuah benda sama besar	Konteks soal: pizza Toni ingin membagikan setengah bagian dari pizzanya, siswa menentukan bagaimana ia membagi pizzanya	Pilihan Ganda Biasa	8
Pola bukan bilangan	Siswa dapat menentukan pola pada urutan tertentu	Konteks soal: bola berwarna Diberikan 15 bola berwarna dengan pola tertentu, siswa menentukan warna bola ke 9	Isian Singkat	9
	Siswa dapat melanjutkan pola	Konteks soal: bola berwarna Urutan pola diubah, siswa menentukan pernyataan yang benar	Pilihan Ganda Kompleks	10

Selanjutnya, tahapan konstruksi butir tes dilakukan dengan berpedoman pada kisi-kisi yang telah disusun. Beberapa butir soal dikelompokkan dengan satu wacana tertentu. Setelah konstruksi soal dilakukan, peneliti melakukan validasi isi/ konten terhadap butir soal yang sudah disusun. Terdapat 4 validator ahli yang merupakan dosen PGSD dengan konsentrasi pendidikan matematika. Hasil validasi isi oleh empat ahli menunjukkan seluruh butir soal dinyatakan valid dengan nilai V Aiken berkisar 0,90–1,00. Revisi dilakukan sesuai masukan validator terkait konteks, kalimat soal, dan kejelasan ilustrasi.

Uji coba soal tes numerasi dilaksanakan di SD N 37 Cakranegara, SD N 5 Cakranegara dan SD N 50 Cakranegara. Tiga sekolah ini dipilih melalui pendekatan *purposive sampling*. Pemilihan ini mempertimbangkan beberapa aspek praktis dan akademis, yaitu ketersediaan siswa kelas 3 sebagai subjek uji coba, kesiapan sekolah dalam menerapkan Kurikulum Merdeka, serta kemudahan akses dan koordinasi yang memungkinkan proses pengambilan data berlangsung dengan lancar. Ketiga sekolah tersebut juga memiliki kondisi pembelajaran yang stabil sehingga sesuai digunakan untuk menguji kelayakan awal instrumen yang dikembangkan. Jumlah seluruh testee yang mengikuti uji coba ini adalah 150 siswa. Analisis dari uji butir tes ini menggunakan Rasch Model dengan melihat: keterpenuhan unidimensi dan independensi, kesesuaian item dengan model, reliabilitas, serta analisis tingkat kesulitan. Analisis ini dilakukan dengan menggunakan aplikasi Winstep. Berdasarkan analisis butir soal menggunakan Rasch model tersebut, diperoleh beberapa kesimpulan, yaitu:

1. Asumsi unidimensi berkaitan dengan jumlah konstruk/ variable/ obyek yang diukur oleh sebuah tes. Tes yang baik hanya mengukur 1 dimensi/ konstruk/ variable. Tes dianggap mengukur 1 dimensi jika *eigens' value of unexplained variance* < 2.0 . Pada tabel *Output Winstep* menunjukkan angka *eigens' value* $1.9 < 2.0$. Dengan demikian dapat disimpulkan bahwa tes mengukur dimensi atau variable tunggal. Artinya, asumsi unidimensi terpenuhi.
2. Asumsi independensi local berkaitan dengan bebasnya respon setiap item dari clue atau jawaban terhadap item lain. Dengan kata lain, jawaban terhadap sebuah item tidak dipengaruhi oleh jawaban pada item yang lain. Asumsi independensi local diuji dengan memeriksa korelasi residu terstandar dari setiap pasangan item. Koefisien korelasi < 0.3 antar item menandakan keterpenuhan asumsi tersebut. Pada tabel *Output Winstep* menunjukkan bahwa semua koefisien korelasi antar item tidak melebihi 0.3. Dengan demikian dapat disimpulkan bahwa semua item yang ada di tes telah memenuhi asumsi independensi local butir tes.
3. Reliabilitas tes menunjukkan konsistensi dan presisi jawaban testee. Konsistensi berkaitan dengan seberapa konsisten jawaban testee ketika diberikan tes yang sama pada waktu yang berbeda. Presisi berkaitan dengan seberapa tepat informasi yang dihasilkan oleh tes. Dengan kata lain, presisi menunjukkan seberapa dekat jawaban terhadap tes dengan kemampuan testee yang sesungguhnya. Hasil analisis menunjukkan nilai reliabilitas sebesar 0.94 (kategori sangat tinggi). Ini menandakan bahwa peluang konsistensi jawaban testee dan kedekatan antara jawaban testee dengan kemampuan yang sebenarnya adalah 94%.
4. Terdapat 3 item (2, 5 dan 6) yang tidak memenuhi kriteria kesesuaian dengan Model Rasch. Ini berarti bahwa ketiga item itu tidak cukup akurat mengukur kemampuan yang ditargetkan. Untuk itu ketiga item tersebut perlu ditelaah lebih jauh untuk menemukan penyebab ketidaksesuaiannya dengan Model rasch.

Sedangkan item yang lain dapat dikatakan valid untuk mengukur kemampuan yang ditargetkan.

5. Item-item menunjukkan tingkat kesulitan yang lebih tinggi daripada kemampuan testee. Pada saat yang sama, item-item di dalam tes juga menunjukkan variasi tingkat kesulitan yang bertingkat, dari yang paling mudah (item 5) hingga yang paling sulit (item 6). Dengan demikian dapat dikatakan bahwa item-item yang ada pada tes cukup mampu memetakan kemampuan yang dimiliki oleh *testee* (peserta tes)

Selanjutnya, tahap revisi tes bertujuan untuk memperbaiki tes berdasarkan hasil uji coba dan melakukan validasi silang (*cross validation*) untuk menentukan: validitas kriteria, validitas prediktif dan validitas konstruk. Revisi dilakukan berdasarkan umpan balik responden uji coba dengan mempertimbangkan (1) perilaku responden, (2) kondisi pemberian tes, (3) kejelasan petunjuk pengerjaan, (4) kenyamanan penggunaan lembar jawaban, (5) kesesuaian tes, (6) kesetaraan budaya, (7) kecukupan waktu, (8) perasaan peserta tes, (9) tingkat tebakan, dan (10) kemungkinan kecurangan. Aspek-aspek tersebut diamati selama proses uji coba dan digunakan untuk mengidentifikasi bagian instrumen yang perlu diperbaiki, seperti kejelasan bahasa, konteks soal, dan tata letak, sehingga instrumen menjadi lebih mudah dipahami dan layak digunakan.

Tahap publikasi merupakan tahap akhir dari penyusunan instrument tes. Adapun yang dilakukan pada tahap ini adalah: (1) Menyiapkan materi tes; (2) Menyusun petunjuk bagi pengguna (penguji); (3) Menyusun petunjuk pengolahan hasil tes (*technical manual*); (4) Mendesain tampilan tes yang *user friendly* dan menarik; dan (5) Mencantumkan tingkat validitas dan reliabilitas tes.

3.2 Pembahasan

Kemampuan asesmen kompetensi minimum, atau yang umum dikenal sebagai AKM, menekankan pentingnya kemampuan numerasi dalam konteks pendidikan. Kemampuan numerasi didefinisikan sebagai kemampuan untuk menggunakan, memahami, dan mengevaluasi informasi matematis dalam kehidupan sehari-hari, seperti berbelanja atau menghitung durasi perjalanan. Sejalan dengan itu, (Zulfayani et al., 2023) mencatat bahwa numerasi adalah kemampuan untuk menerapkan konsep angka dan keterampilan aritmetika dalam kehidupan sehari-hari serta interpretasi informasi kuantitatif di sekitar kita. Sebagian siswa yang memiliki kemampuan matematika tinggi belum tentu memiliki literasi numerasi yang memadai, karena literasi numerasi juga melibatkan kemampuan berpikir logis matematis. Sebagaimana dikemukakan oleh Pratiwi et al., (2024), literasi numerasi meliputi pemahaman kuantitatif sekaligus penalaran logis terhadap informasi matematis, sehingga diperlukan asesmen yang tepat untuk mengevaluasi kompetensi numerasi secara menyeluruh.

Berdasarkan penelitian oleh (Baharuddin et al., 2021), tujuan dari asesmen numerasi adalah untuk mengevaluasi keterampilan kognitif siswa dalam memecahkan masalah praktis menggunakan konsep dan metode matematika. Asesmen numerasi menuntut siswa untuk menerapkan pengetahuan konseptual dan prosedural dalam situasi kehidupan nyata serta menggunakan penalaran matematis untuk menganalisis informasi dan menyelesaikan persoalan kontekstual (Amami et al., 2021; Mariska et al., 2024). Hal ini menciptakan kebutuhan untuk instrumen yang dapat mengukur berbagai aspek pendidikan numerasi, termasuk keterampilan dasar, penerapan, dan pemahaman yang lebih mendalam.

Keterampilan numerasi yang rendah di Indonesia, terutama dalam konteks PISA, menjadi perhatian penting bagi pendidik dan pembuat kebijakan (Aprilie & Kasih, 2023). Penelitian menunjukkan bahwa upaya peningkatan numerasi perlu fokus pada pembiasaan siswa dengan situasi nyata melalui pembelajaran kontekstual, misalnya model *problem-based learning* yang secara langsung mengaitkan matematika dengan kehidupan sehari-hari (Sutrimo et al., 2024). Oleh karena itu, sangat krusial untuk mengembangkan instrumen asesmen numerasi yang tidak hanya mengukur pengetahuan teori, tetapi juga kemampuan aplikasi praktis dari numerasi siswa.

Instrumen asesmen yang dikembangkan harus mencakup berbagai bentuk soal yang menilai tingkat kognitif berbeda, termasuk *knowing*, *applying*, dan *reasoning*. Kerangka kognitif ini dijelaskan secara jelas dalam kerangka penilaian PISA sebagai cara untuk mengukur aspek pemahaman, penerapan, dan penalaran matematis (OECD, 2019). Selain itu, panduan Asesmen Kompetensi Minimum di Indonesia juga menegaskan bahwa soal AKM harus mencerminkan level kognitif yang beragam agar asesmen menjadi lebih representatif terhadap kompetensi numerasi siswa (Kemdikbud, 2021). Dalam konteks ini, penelitian Nurisa Syabaniah et al., (2025) menunjukkan bahwa penggunaan level kognitif PISA dalam konstruksi tes sangat efektif dalam menciptakan butir yang bermakna dan relevan untuk siswa Indonesia. Dalam penelitian ini soal-soal yang disusun mencakup 10 pertanyaan dalam berbagai format, seperti pilihan ganda dan esai, yang dirancang untuk menguji pemahaman siswa di berbagai level kognitif. Contohnya, soal pada level *knowing* akan menguji pengetahuan dasar, sementara soal pada level *reasoning* akan menilai kemampuan analisis dan penalaran siswa dalam menghadapi situasi yang lebih kompleks (Kristianti & Handayani, 2023)

Instrumen tes yang dikembangkan dalam penelitian ini terdiri dari 10 soal yang dirancang dalam berbagai bentuk, yaitu 2 soal pilihan ganda, 3 soal pilihan ganda kompleks, 2 soal isian singkat, 2 soal menjodohkan dan 1 soal *essay*. Soal-soal ini dirancang untuk mencakup berbagai tingkat kognitif, yaitu *knowing*, *applying*, dan *reasoning*. Instrumen ini tidak hanya menguji pengetahuan siswa tentang matematika, tetapi juga kemampuan siswa untuk menerapkan pengetahuan tersebut dalam situasi nyata.

Hasil penelitian menunjukkan bahwa pengembangan soal tes numerasi berbasis *hierarchical content* telah menghasilkan instrumen yang valid dan reliabel untuk mengukur kemampuan numerasi siswa Sekolah Dasar pada Fase A elemen Bilangan dan

Aljabar Kurikulum Merdeka. Temuan ini memperkuat pandangan bahwa asesmen yang disusun berdasarkan urutan konten yang hierarkis dapat memberikan informasi diagnostik yang lebih mendalam tentang kemampuan siswa dibandingkan asesmen konvensional yang hanya berorientasi pada hasil akhir (Black & Wiliam, 2010). Validitas isi yang tinggi pada seluruh butir soal menunjukkan bahwa indikator dan materi yang digunakan telah sesuai dengan capaian pembelajaran elemen Bilangan dan Aljabar. Hal ini sejalan dengan pendapat Gregory, (2016) bahwa validitas isi yang kuat mencerminkan kesesuaian antara item tes dengan konstruk yang diukur. Dalam konteks numerasi, kesesuaian ini penting karena setiap konsep matematika memiliki keterkaitan prasyarat yang membentuk struktur berpikir siswa secara bertahap.

Analisis menggunakan Model Rasch menghasilkan nilai reliabilitas sebesar 0,94, yang mengindikasikan konsistensi instrumen sangat tinggi. Temuan ini menunjukkan bahwa soal-soal yang dikembangkan mampu mengukur kemampuan numerasi siswa secara stabil dan presisi. Menurut Bond et al., (2021) reliabilitas tinggi pada model Rasch menggambarkan kemampuan instrumen untuk membedakan kemampuan siswa secara objektif tanpa dipengaruhi oleh konteks pengujian. Dari hasil analisis unidimensi dan independensi lokal, seluruh butir soal terbukti hanya mengukur satu konstruk, yaitu kemampuan numerasi, dan setiap butir soal bekerja secara independen. Pemenuhan kedua asumsi ini menunjukkan bahwa instrumen telah memenuhi prinsip dasar pengukuran modern dalam Model Rasch (Bond et al., 2021). Dengan demikian, tes yang dikembangkan layak digunakan untuk analisis kemampuan individual maupun pemetaan tingkat penguasaan konsep numerasi (Boone et al., 2014).

Namun demikian, terdapat tiga butir (nomor 2, 5, dan 6) yang menunjukkan ketidaksesuaian terhadap model. Ketidaksesuaian ini dapat disebabkan karena faktor konteks yang kurang familiar bagi peserta tes atau tingkat kesulitan yang melampaui kemampuan rata-rata peserta tes. Sebagaimana ditemukan oleh Aini et al., (2024), soal dengan konteks personal kadang-kadang kurang dipahami siswa bila konteksnya tidak dekat dengan pengalaman mereka sehari-hari. Selain itu, Arofa & Ismail (2022) menunjukkan bahwa interpretasi siswa terhadap soal numerasi sangat tergantung pada konteks sosial maupun saintifik, sehingga variasi konteks perlu diperhatikan dalam pengembangan instrumen asesmen untuk menghindari kesalahan interpretasi dan menjamin validitas pemahaman siswa. Revisi terhadap butir-butir tersebut diperlukan untuk memastikan bahwa setiap soal merepresentasikan tingkat berpikir yang sesuai dengan fase perkembangan kognitif siswa SD. Rentang tingkat kesulitan item (-3,27 hingga 6,28 logit) menunjukkan bahwa soal yang dikembangkan telah mencakup variasi kemampuan dari tingkat rendah hingga tinggi. Artinya, tes ini efektif digunakan untuk memetakan kemampuan numerasi siswa secara berjenjang. Fakta bahwa rata-rata kemampuan siswa masih berada di bawah tingkat kesulitan soal menunjukkan bahwa kemampuan numerasi siswa Fase A masih relatif rendah. Kondisi ini konsisten dengan hasil Asesmen Nasional dan studi PISA yang menunjukkan rendahnya literasi numerasi siswa Indonesia (Kemdikbud, 2023b).

Secara pedagogis, hasil ini menegaskan pentingnya penyusunan soal berdasarkan *hierarchical content*. Dengan mengetahui pada konten mana siswa mengalami kesulitan, guru dapat lebih mudah menentukan strategi pembelajaran remedial dan penguatan. Hal ini mendukung prinsip asesmen formatif dalam Kurikulum Merdeka yang berorientasi pada perbaikan proses belajar, bukan sekadar pengukuran hasil (Kemdikbudristek, 2022)

Dengan demikian, pengembangan tes numerasi berbasis *hierarchical content* tidak hanya menghasilkan instrumen yang valid dan reliabel, tetapi juga memberikan kontribusi praktis bagi guru dalam memfasilitasi pembelajaran diferensiasi dan berbasis kebutuhan siswa. Model pengembangan seperti ini dapat menjadi alternatif asesmen diagnostik yang lebih bermakna bagi pembelajaran numerasi di sekolah dasar.

4. SIMPULAN

Kesimpulan dari pengembangan soal tes numerasi berbasis *hierarchical content* dalam implementasi kurikulum Merdeka pada siswa SD/ MI adalah sebagai berikut:

1. Tahapan pengembangan soal tes numerasi berbasis *hierarchical content* adalah (1) mendefinisikan tes; (2) memilih metode penskalaan; (3) konstruksi butir tes; (4) uji butir tes; (5) revisi tes; (6) publikasi tes.
2. Hasil analisis capaian pembelajaran pada elemen bilangan dan aljabar fase A diperoleh lima konten materi hierarkis yaitu bilangan cacah sampai 100, operasi penjumlahan yang banyaknya sampai dengan 20 menggunakan benda-benda kongkret, operasi pengurangan yang banyaknya sampai dengan 20 menggunakan benda-benda kongkret, pecahan setengah dan seperempat, dan pola bukan bilangan
3. Terdapat 10 soal yang dikembangkan. Validitas isi oleh *expert judgement* ke10 soal dinyatakan valid dengan beberapa revisi. Kesepuluh item tes kemudian diujicobakan dan dianalisis dengan Rasch Model dengan bantuan aplikasi winstep yang menghasilkan (a) Tes dan item yang ada di dalamnya memenuhi asumsi prasyarat yang dibutuhkan dalam analisis Model Rasch (asumsi unidimensi dan asumsi local independensi butir). (b) Tes menunjukkan koefisien reliabilitas yang sangat tinggi sehingga kapan pun digunakan, tes dapat menghasilkan informasi yang konsisten dan presisi. (c) Terdapat 3 item (2, 5 dan 6) yang tidak memenuhi kriteria kesesuaian dengan Model Rasch. Ini berarti bahwa ketiga item itu tidak cukup akurat mengukur kemampuan yang ditargetkan. Untuk itu kedua item tersebut perlu ditelaah lebih jauh untuk menemukan penyebab ketidaksesuaiannya dengan Model Rasch. Sedangkan item yang lain dapat dikatakan valid untuk mengukur kemampuan yang ditargetkan. (d) Item-item menunjukkan tingkat kesulitan yang lebih tinggi daripada kemampuan testee. Pada saat yang sama, item-item di dalam tes juga menunjukkan variasi tingkat kesulitan yang bertingkat, dari yang paling mudah (item 5) hingga yang paling sulit (item 6). Dengan demikian dapat dikatakan bahwa item-item yang ada pada tes cukup mampu memetakan kemampuan yang dimiliki oleh *testee* (peserta tes).

5. UCAPAN TERIMA KASIH

Penelitian ini didanai oleh Dana BLU Universitas Mataram dengan nomor kontrak 2063/UN18.L1/PP/2024. Para penulis menyampaikan terima kasih yang sebesar-besarnya kepada Universitas Mataram atas dukungan pendanaan dan fasilitas yang telah diberikan sehingga penelitian ini dapat diselesaikan dengan baik.

6. REKOMENDASI

Instrumen tes numerasi berbasis *hierarchical content* yang telah dikembangkan terbukti valid dan reliabel, sehingga dapat digunakan guru sebagai alat asesmen formatif dalam pembelajaran numerasi di sekolah dasar. Penelitian selanjutnya disarankan untuk mengembangkan instrumen serupa pada elemen dan fase lainnya agar pemetaan kemampuan numerasi siswa lebih komprehensif. Selain itu, implementasi instrumen ini dalam konteks pembelajaran di kelas perlu diuji lebih lanjut untuk melihat efektivitasnya dalam meningkatkan kemampuan numerasi dan pengambilan keputusan guru dalam merancang tindak lanjut pembelajaran.

7. REFERENSI

- Aini, F. N., Kurniati, D., Oktavianingtyas, E., Murtikusuma, R. P., & Safrida, L. N. (2024). Pengembangan soal numerasi model akm konten aljabar subdomain rasio dan proporsi untuk jenjang smp dengan konteks personal [developing akm model numeracy problems in algebra content for the ratio and proportion subdomain at the junior high school level in a personal context]. *JOHME: Journal of Holistic Mathematics Education*, 8(2), 213–230. <https://doi.org/10.19166/johme.v8i2.8401>
- Amami, C. S., Saefudin, A. A., & Sunanti, T. (2021). Students' numeracy skills in solving minimum competency assessment-type mathematics problems: A case study in junior high school. *Journal of Authentic Research on Mathematics Education*, 7(2), 141–150.
- Anderson, L. W., & Krathwohl, D. R. (2021). *A taxonomy for learning, teaching, and assessing: A revision of Bloom's taxonomy of educational objectives*. Longman.
- Aprilie, T., & Kasih, A. P. (2023, December 6). *Mengapa Kemampuan Literasi dan Numerasi Siswa Indonesia Masih Rendah*. Kompas.Com.
- Arofa, A. N., & Ismail, I. (2022). Kemampuan numerasi siswa MA dalam menyelesaikan soal setara asesmen kompetensi minimum pada konten aljabar. *MATHEdunesa*, 11(3), 779–793. <https://doi.org/10.26740/mathedunesa.v11n3.p779-793>
- Baharuddin, M. R., Sukmawati, S., & Christy, C. (2021). Deskripsi kemampuan numerasi siswa dalam menyelesaikan operasi pecahan. *Pedagogy: Jurnal Pendidikan Matematika*, 6(2), 90–101.
- Black, P., & Wiliam, D. (2010). Inside the Black Box: Raising standards through classroom assessment. *Phi Delta Kappan*, 92(1), 81–90. <https://doi.org/10.1177/003172171009200119>
- Bond, T. G., Yan, Z., & Heene, M. (2021). *Applying the Rasch model: Fundamental measurement in the human sciences* (4th ed.). Routledge.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Springer Netherlands. <https://doi.org/10.1007/978-94-007-6857-4>
- Cahyana, A. (2020). *Prospek AKM dan survei karakter: memperkuat basis praliterasi dan pranuterasi usia dini*. <https://banpaudpnf.kemdikbud.go.id/berita/prospek-akm-dan-survei-karakter-memperkuat-basis-praliterasi-dan-pranuterasi-usia-dini>
- Gregory, R. J. (2016). *Psychological testing: history, principles, and applications* (7th ed). Pearson.
- Kemdikbud. (2023a). *Apa itu literasi dan numerasi?* <https://ditpsd.kemdikbud.go.id/artikel/detail/apa-itu-literasi-dan-numerasi#:~:text=Literasi numerasi adalah kecakapan untuk,macam konteks kehidupan sehari-hari.>

- Kemdikbud. (2023b). *Peringkat Indonesia pada PISA 2022 Naik 5-6 Posisi Dibanding 2018*. <https://www.kemdikbud.go.id/main/blog/2023/12/peringkat-indonesia-pada-pisa-2022-naik-56-posisi-dibanding-2018>
- Kemdikbudristek. (2022). *Panduan pembelajaran dan asesmen kurikulum merdeka*. Kementerian Pendidikan, Kebudayaan, Riset dan Teknologi.
- Kristianti, & Handayani, I. (2023). Analisis literasi numerasi matematis peserta didik kelas xi dalam menyelesaikan soal asesmen kompetensi minimum. *Jurnal Pembelajaran Matematika Inovatif*, 6(4). <https://doi.org/10.22460/jpmi.v6i4.17919>
- Kusmaharti, D., Fiantika, F. R., Rusminati, S. H., Yustitia, V., Studi, P., Guru, P., Dasar, S., & Surabaya, B. (2022). *Penyusunan instrumen asesmen ketuntasan minimum*. *Jurnal Abdimas Bina Bangsa*, 3(1), 196-202.
- Mariska, R., Ratnasari, R., & Nurvicalesi, N. (2024). Kemampuan numerasi siswa smp menggunakan soal tipe asesmen kompetensi minimum. *Differential: Journal on Mathematics Education*, 2(2), 141–155. <https://doi.org/10.32502/differential.v2i2.233>
- McKenney, S. E., & Reeves, T. C. (2012). *Conducting educational design research*. Routledge.
- Mendikbud. (2020). *Asesmen nasional sebagai pengganti ujian nasional*.
- Moh Slamet Sutrimo, Sajdah, S. N., Sinambela, Y. V. F., & Bagas, R. (2024). Peningkatan literasi numerasi melalui model pembelajaran dan hubungannya dengan kemampuan self-efficacy: Systematic literatur review. *JPMI (Jurnal Pembelajaran Matematika Inovatif)*, 7(1), 61–72. <https://doi.org/10.22460/jpmi.v7i1.21650>
- Novita, N., Mellyzar, & Herizal. (2021). Asesmen Nasional (AN): Pengetahuan dan Persepsi Calon Guru. *JISIP (Jurnal Ilmu Sosial Dan Pendidikan)*, 5(1), 72-79.
- Nurisa Syabaniah, T., Ilma Indra Putri, R., & Zulkardi, Z. (2025). Pengembangan soal matematika tipe PISA konten bilangan untuk literasi matematis siswa SMP. *Jurnal Pendidikan Matematika: Judika Education*, 8(4). <https://doi.org/10.31539/ev90kv23>
- Pratiwi, S. A., Nofikusumawati Peni, N. R., & Prabowo, A. (2024). Study on literacy numeracy towards students' logic mathematics: a literature review. *Numeracy*, 11(1), 58–69. <https://doi.org/10.46244/numeracy.v11i1.2601>
- Zulfayani, Z., Ariawan, R., Nufus, H., & Kafrina, K. (2023). Profile of students' numeracy literacy ability and self efficacy in mathematics learning. *Mathematics Research and Education Journal*, 7(2), 25–35. [https://doi.org/10.25299/mrej.2023.vol7\(2\).14381](https://doi.org/10.25299/mrej.2023.vol7(2).14381)