

**PENGEMBANGAN SOAL AKM NUMERASI BERBASIS  
KOMPUTER UNTUK KELAS XI SMA**



**Oleh:**

**IHDA MUTIMMATUL FITRIYAH**

**NIM 21701251006**

**Tesis ini ditulis untuk memenuhi sebagian persyaratan  
untuk mendapatkan gelar Magister Pendidikan**

**PROGRAM STUDI PENELITIAN DAN EVALUASI PENDIDIKAN**

**SEKOLAH PASCASARJANA**

**UNIVERSITAS NEGERI YOGYAKARTA**

**2023**

## ABSTRAK

IHDA MUTIMMATUL FITRIYAH: Pengembangan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA. Tesis. Yogyakarta: Sekolah Pascasarjana, Universitas Negeri Yogyakarta, 2023.

Penelitian ini bertujuan untuk: (1) mendesain konstruk soal AKM numerasi berbasis komputer untuk kelas XI SMA, (2) menganalisis kualitas soal AKM numerasi, (3) mengungkapkan kesetaraan soal AKM numerasi yang telah dikembangkan, (4) mendeskripsikan profil peserta didik dalam menyelesaikan soal yang telah dikembangkan, dan (5) mengungkapkan keefektifan soal yang telah dikembangkan.

Penelitian ini merupakan penelitian pengembangan yang diadopsi dari Retnawati (2016a) dengan sembilan langkah yaitu: (1) menentukan tujuan penyusunan instrumen, (2) menentukan cakupan materi, (3) menyusun indikator dan matriks instrumen soal, (4) menyusun butir soal, (5) validasi isi, (6) revisi berdasarkan masukan validator, (7) uji coba, (8) menentukan validitas konstruk, reliabilitas, model fit, dan karakteristik butir soal, serta (9) merakit soal. Uji coba instrumen melibatkan 1.570 peserta didik dari enam SMA negeri di Kota Yogyakarta. Instrumen pengumpulan data yaitu lembar validasi, AKM numerasi terkomputerisasi LMS Moodle, dan angket. Validasi isi dilakukan oleh ahli dengan menggunakan validasi Aiken's V, dan validitas konstruk dibuktikan melalui Confirmatory Factor Analysis. Karakteristik butir dianalisis menggunakan teori respon butir data campuran dengan model Rasch. Proses penyetaraan dilakukan dengan desain tes dengan butir *anchor* menggunakan metode Rerata & Rerata. Analisis statistik deskriptif digunakan untuk mendeskripsikan profil peserta didik dan keefektifan tes yang dikembangkan.

Hasil penelitian adalah sebagai berikut. (1) Tes yang dikembangkan terdiri atas 54 butir soal, dalam empat paket soal dengan empat butir *anchor* di setiap paket dengan lima bentuk soal (pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian). Konteks yang dilibatkan adalah personal, sosial budaya, dan saintifik dengan level kognitif mencakup pemahaman, penerapan, dan penalaran. (2) Tes yang dikembangkan memiliki kualitas yang baik berdasarkan uji validitas isi, validitas konstruk, reliabilitas, kecocokan butir, dan analisis teori respon butir data campuran model Rasch. Tes tersebut valid, reliabel, dan dapat digunakan untuk mengukur numerasi peserta didik. (3) Empat paket tes tersebut memiliki kesetaraan yang tidak paralel. Setelah paket B, C, dan D disetarakan ke paket A dapat diketahui bahwa paket B dan paket C cenderung setara, namun tidak dengan paket D. (4) Dalam menyelesaikan tes tersebut, siswa berada di bawah kompetensi minimum: 53,7% pada tingkat dasar; 30,9% memerlukan intervensi khusus; 10,8% cakap, dan 4,6% mahir. (5) Keefektifan tes tersebut menunjukkan kepraktisan (72,8%) dan kemanfaatan soal (91,1%).

**Kata Kunci:** AKM, berbasis komputer, LMS moodle, numerasi

## ABSTRACT

IHDA MUTIMMATUL FITRIYAH: Developing Computerized MCA's Numeracy Items for Grade XI. **Thesis. Yogyakarta: Graduate School, Yogyakarta State University, 2023.**

This study aims to: (1) design the constructs of computerized MCA's numeracy items for grade XI, (2) reveal the quality of computerized MCA's numeracy items, (3) reveal the equivalency of the developed MCA's numeracy items, (4) describe student profile in completing the developed test, and (4) reveal the effectiveness of the developed test.

This research is development research with the test development procedure adoption from Retnawati (2016a) which consists of nine steps, namely: (1) determining the purpose of preparing instrument, (2) looking for material coverage, (3) compiling item indicators and matrix, (4) compiling test items, (5) validating content, (6) revisions based on validator input, (7) conducting trials, (8) looking for construct validity, reliability, model fit, item characteristics, and (9) assembling instruments. The trial of the test involved 1,570 students from six public senior high schools in Yogyakarta City. The instruments of data collection are a validation sheet, computerized test on LMS Moodle, and questionnaire. Content validation was carried out by experts using Aiken's V validation, construct validation proven through Confirmatory Factor Analysis. Item characteristics were analyzed using Rasch model from item response theory for mixed item. Equating process was carried out by test design with anchor items using the Average & Average method. Describing the students profile and effectiveness of the developed test was conducted using the descriptive statistical analysis.

The research results are as follows. (1) The developed test consists of 54 questions, in four packages with four anchor items in each package and with five question forms (multiple-choice, complex multiple-choice, matching, short answer, and description). The context involved is personal, socio-cultural, and scientific with a cognitive level including understanding, application, and reasoning. (2) The developed test has good quality based on content validity, construct validity, reliability, item suitability, and Rasch model mixed data item response theory analysis. The test is valid, reliable, and can be used to measure students' numeracy. (3) The four test packages have non-parallel equivalence. After packages B, C, and D are equated to package A, it can be seen that package B and package C tend to be equivalent, but not package D. (4) In completing the test, students are below the minimum competency: 53.7% at the basic level; 30.9% require special intervention; 10.8% skillful, and 4.6% proficient. (5) The effectiveness of the test shows its practicality (72.8%) and the usefulness of the questions (91.1%).

**Keywords:** AKM, computer-based, Moodle LMS, numeracy

## PERNYATAAN KEASLIAN KARYA

Saya yang bertanda tangan di bawah ini:

Nama : Ihda Mutimmatul Fitriyah

NIM : 21701251006

Program Studi : Penelitian dan Evaluasi Pendidikan

Dengan ini menyatakan bahwa tesis ini merupakan hasil karya saya sendiri dan belum pernah diajukan untuk memperoleh gelar magister di suatu perguruan tinggi, dan sepanjang pengetahuan saya dalam tesis ini tidak terdapat karya atau pendapat yang pernah ditulis atau diterbitkan oleh orang lain kecuali yang secara tertulis diacu dalam naskah ini dan disebutkan dalam daftar pustaka.

Yogyakarta, 14 Agustus 2023

Yang membuat pernyataan



Ihda Mutimmatul Fitriyah

NIM 21701251006

**LEMBAR PERSETUJUAN**

**PENGEMBANGAN SOAL AKM NUMERASI BERBASIS KOMPUTER  
UNTUK KELAS XI SMA**

**IHDA MUTIMMATUL FTIRIYAH  
NIM 21701251006**

Tesis ini ditulis untuk memenuhi sebagai persyaratan  
mendapatkan gelar Magister Pendidikan  
Program Studi Penelitian dan Evaluasi Pendidikan

Menyetujui untuk diajukan pada ujian tesis  
Pembimbing,



Prof. Dr. Heri Retnawati, M.Pd.

Mengetahui:

Sekolah Pascasarjana

Universitas Negeri Yogyakarta

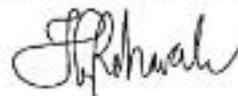
Direktur,



Chufon, M.Pd.

NIP 19621111988031001

Koordinator Program Studi,



Prof. Dr. Heri Retnawati, M.Pd.

NIP 19730103 200003 2 001

LEMBAR PENGESAHAN

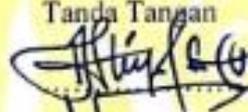
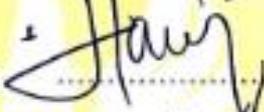
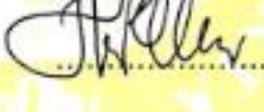
PENGEMBANGAN SOAL AKM NUMERASI BERBASIS  
KOMPUTER UNTUK KELAS XI SMA

TESIS

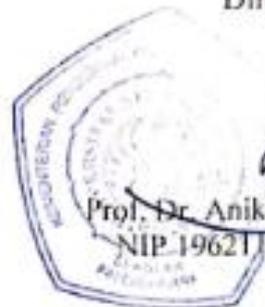
IHDA MUTIMMATUL FITRIYAH  
NIM 21701251006

Telah dipertahankan di depan Dewan Penguji Tesis  
Sekolah Pascasarjana Universitas Negeri Yogyakarta  
Tanggal: 14 September 2023

DEWAN PENGUJI

Nama/Jabatan	Tanda Tangan	Tanggal
Prof. Dr. Edi Istiyono, M.Si. (Ketua/Penguji)		13/10 2023
Dr. Syukrul Hamdi, S.Pd., M.Pd. (Sekertaris/Penguji)		11/10 2023
Prof. Dr. Sugiman, M.Si. (Penguji I)		11/10 2023
Prof. Dr. Heri Retnawati, M.Pd. (Penguji II/Pembimbing)		13/10 - 2023

Yogyakarta, 13 Oktober 2023  
Sekolah Pascasarjana  
Universitas Negeri Yogyakarta  
Direktur,



  
Prof. Dr. Anik Ghufron, M.Pd.  
NIP.1962111988031001

## **KATA PENGANTAR**

Puji syukur penulis panjatkan kepada Allah SWT yang telah melimpahkan berkah rahmat-Nya sehingga tesis dengan judul “Pengembangan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA” dapat selesai di waktu yang tepat dengan baik. Penulisan tesis ini tak lepas dari doa, bantuan, serta dukungan dari berbagai pihak. Oleh karena itu, pada kesempatan ini penulis sampaikan terima kasih kepada:

1. Prof. Dr. Sumaryanto, M. Kes., selaku Rektor Universitas Negeri Yogyakarta;
2. Prof. Dr. Anik Ghufron, M.Pd., selaku Direktur Sekolah Pascasarjana Universitas Negeri Yogyakarta;
3. Prof. Dr. Heri Retnawati, M.Pd., selaku dosen pembimbing dan Koordinator Program Studi Magister Penelitian dan Evaluasi Pendidikan yang telah membimbing dengan sangat baik, begitupun dengan segala arahan dan motivasinya untuk penulis segera menyelesaikan tesis ini dengan baik;
4. Dr. Syukrul Hamdi, M.Pd., Dr. Kana Hidayati, M.Pd., Dr. Siti Lailiyah, M.Si., Prof. Ir. Moh. Khairudin, M.T., Ph.D., IPU., ASEAN. Eng., Dr. Risky Setiawan, M.Pd., Risa Utaminingsih, M.Pd., Itsna Lailatul Mas’udah, M.Pd., Durrotun Nabilah, M.Pd. selaku validator ahli instrumen dan media yang saya kembangkan dalam penelitian;
5. Bapak Ibu Dosen Program Studi Magister Penelitian dan Evaluasi Pendidikan Universitas Negeri Yogyakarta yang dengan tulus dan ikhlas mendampingi penulis untuk bertumbuh pada rumpun ilmu yang sedang penulis fokus;
6. Seluruh Staff Akademik dan Tata Usaha Sekolah Pascasarjana Universitas Negeri Yogyakarta yang telah membantu penulis dari awal menempuh pendidikan hingga dalam menyelesaikan tesis sebagai tugas akhir;
7. Bapak Ibu Kepala SMA Negeri di Kota Yogyakarta tempat penulis melakukan penelitian yang telah memberikan izin dan dukungannya sehingga proses pengambilan data terlaksana dengan sangat baik;

8. Bapak Ibu Guru Matematika dan Peserta Didik kelas XI yang telah menjadi subjek penelitian, terima kasih atas bantuan dan kerjasamanya sehingga penulis dapat memperoleh data dengan sangat baik;
9. Kedua orang tua, saudara kandung beserta keluarga besar, dan sahabat-sahabat penulis yang selalu memberikan dukungan dan motivasi yang tak terhingga;
10. Teman-teman Program Studi Penelitian dan Evaluasi Pendidikan baik seangkatan, kakak tingkat, serta adik tingkat yang telah saling memberi dukungan dan menyediakan ruang diskusi yang baik dan nyaman;
11. Pihak-pihak lain yang tidak bisa penulis sebutkan satu persatu, terima kasih atas segala bantuan, dukungan dan motivasi yang diberikan dengan ikhlas.

Teriring doa semoga Allah SWT memberi balasan kebaikan kepada semua pihak. Penulis juga menyadari akan keterbatasan pengetahuan, kemampuan, dan pengalaman yang penulis miliki. Oleh karena itu saran dan masukan yang membangun sangat penulis harapkan untuk perbaikan karya tulis ini. Terakhir, besar harapan penulis agar tesis ini dapat bermanfaat dengan baik.

Yogyakarta, 14 Agustus 2023

Penulis,

Ihda Mutimmatul Fitriyah

## DAFTAR ISI

<b>ABSTRAK</b> .....	<b>i</b>
<b>ABSTRACT</b> .....	<b>ii</b>
<b>PERNYATAAN KEASLIAN KARYA</b> .....	<b>iii</b>
<b>LEMBAR PERSETUJUAN</b> .....	<b>iv</b>
<b>LEMBAR PENGESAHAN</b> .....	<b>v</b>
<b>KATA PENGANTAR</b> .....	<b>vi</b>
<b>DAFTAR ISI</b> .....	<b>viii</b>
<b>DAFTAR TABEL</b> .....	<b>x</b>
<b>DAFTAR GAMBAR</b> .....	<b>xii</b>
<b>DAFTAR LAMPIRAN</b> .....	<b>xiv</b>
<b>BAB I PENDAHULUAN</b>	
A. Latar Belakang Masalah .....	1
B. Identifikasi Masalah .....	8
C. Pembatasan Masalah .....	9
D. Rumusan Masalah .....	9
E. Tujuan Pengembangan .....	10
F. Spesifikasi Produk yang Dikembangkan.....	10
G. Manfaat Penelitian.....	11
H. Asumsi Pengembangan .....	12
<b>BAB II KAJIAN PUSTAKA</b>	
A. Kajian Teori.....	13
1. Asesmen .....	13
2. Asesmen Kompetensi Minimum.....	15
3. Numerasi .....	37
4. Penilaian Berbasis Komputer (CBT).....	52
B. Kajian Penelitian yang Relevan .....	54
C. Kerangka Pikir.....	60
D. Pertanyaan Penelitian .....	64

### **BAB III METODE PENELITIAN**

A. Model Pengembangan .....	66
B. Prosedur Pengembangan .....	66
C. Desain Uji Coba Produk.....	70
1. Desain Uji Coba .....	70
2. Subjek Uji Coba .....	70
3. Teknik dan Instrumen Pengumpulan Data .....	72
4. Teknik Analisis Data .....	73

### **BAB IV HASIL PENELITIAN DAN PEMBAHASAN**

A. Hasil Pengembangan Produk Awal.....	83
B. Hasil Uji Coba Produk .....	104
C. Revisi dan Produk Akhir .....	144
D. Kajian Produk Akhir .....	151
E. Keterbatasan Penelitian.....	160

### **BAB V SIMPULAN DAN SARAN**

A. Simpulan tentang Produk.....	161
B. Saran Pemanfaatan Produk .....	162
C. Diseminasi dan Pengembangan Produk Lebih Lanjut .....	162

<b>DAFTAR PUSTAKA .....</b>	<b>164</b>
-----------------------------	------------

<b>LAMPIRAN.....</b>	<b>178</b>
----------------------	------------

## DAFTAR TABEL

<b>Tabel 1.</b> Kisi-kisi Numerasi Jenjang SMA/SMK/Sederajat .....	11
<b>Tabel 2.</b> Konten Pada Literasi Membaca dan Numerasi .....	19
<b>Tabel 3.</b> Level Kognitif Pada Literasi Membaca dan Numerasi .....	20
<b>Tabel 4.</b> Jumlah Pernyataan Soal Pilihan Ganda Kompleks .....	24
<b>Tabel 5.</b> Distribusi Domain Geometri dan Pengukuran .....	47
<b>Tabel 6.</b> Distribusi Domain Aljabar .....	48
<b>Tabel 7.</b> Distribusi Domain Data dan Ketidakpastian .....	49
<b>Tabel 8.</b> Distribusi Konteks Numerasi .....	49
<b>Tabel 9.</b> Distribusi Level Kognitif Numerasi .....	51
<b>Tabel 10.</b> Kriteria <i>Goodness of Fit Test</i> .....	74
<b>Tabel 11.</b> Kategori Koefisien Reliabilitas .....	77
<b>Tabel 12.</b> Tingkat Kompetensi yang Dicapai Peserta Didik .....	80
<b>Tabel 13.</b> Kategori Keefektifan .....	82
<b>Tabel 14.</b> Kisi-kisi Numerasi SMA .....	85
<b>Tabel 15.</b> Nama-nama Validator Isi Soal AKM Numerasi .....	100
<b>Tabel 16.</b> Perhitungan Hasil <i>Aiken's V</i> .....	101
<b>Tabel 17.</b> Nama-nama Validator Media CBT .....	102
<b>Tabel 18.</b> Subjek Uji Coba Penelitian .....	104
<b>Tabel 19.</b> Uji Kecocokan Butir Paket A .....	118
<b>Tabel 20.</b> Uji Kecocokan Butir Paket B .....	119
<b>Tabel 21.</b> Uji Kecocokan Butir Paket C .....	120
<b>Tabel 22.</b> Uji Kecocokan Butir Paket D .....	121
<b>Tabel 23.</b> Nilai Parameter Butir Paket A .....	124
<b>Tabel 24.</b> Nilai Parameter Butir Paket B .....	126
<b>Tabel 25.</b> Nilai Parameter Butir Paket C .....	128
<b>Tabel 26.</b> Nilai Parameter Butir Paket D .....	130
<b>Tabel 27.</b> Indeks Numerasi Peserta Didik SMA Kelas XI .....	136
<b>Tabel 28.</b> Hasil Uji Pasang (Uji <i>Tukey</i> ) .....	137
<b>Tabel 29.</b> Hasil Uji Kesetaraan Keempat Paket Soal .....	138
<b>Tabel 30.</b> Sebaran <i>Anchor</i> dan Jumlah Responden .....	138

<b>Tabel 31.</b> Statistik Deskriptif Tingkat Kesukaran Empat Paket Soal.....	139
<b>Tabel 32.</b> Hasil Konstanta Penyetaraan Metode Rerata & Rerata.....	140
<b>Tabel 33.</b> Persamaan Kesetaraan Skor Metode Rerata & Rerata .....	141
<b>Tabel 34.</b> Hasil Skor Penyetaraan untuk Beberapa <i>Ability</i> .....	141

## DAFTAR GAMBAR

<b>Gambar 1.</b> Tampilan Domain <i>Hosting</i> .....	87
<b>Gambar 2.</b> Cuplikan Kisi-kisi Instrumen Soal Tahap Awal.....	88
<b>Gambar 3.</b> <i>Storyboard</i> Tampilan Halaman Utama.....	89
<b>Gambar 4.</b> <i>Storyboard</i> Tampilan Halaman Administrator .....	90
<b>Gambar 5.</b> <i>Storyboard</i> Tampilan Halaman Pendidik .....	91
<b>Gambar 6.</b> <i>Storyboard</i> Tampilan Halaman Peserta Didik.....	92
<b>Gambar 7.</b> Sampel Kartu Soal AKM Numerasi .....	93
<b>Gambar 8.</b> Tampilan Halaman Utama.....	94
<b>Gambar 9.</b> Tampilan Menu <i>Dashboard</i> .....	95
<b>Gambar 10.</b> Tampilan Menu <i>Site Home</i> .....	95
<b>Gambar 11.</b> Tampilan Menu <i>Site Home</i> Pendidik.....	96
<b>Gambar 12.</b> Tampilan <i>Edit Quiz</i> .....	96
<b>Gambar 13.</b> Tampilan <i>Edit Soal</i> .....	97
<b>Gambar 14.</b> Tampilan Aktivitas dalam <i>Course</i> .....	97
<b>Gambar 15.</b> Tampilan Soal Numerasi dalam Aktivitas <i>Quiz</i> .....	98
<b>Gambar 16.</b> Tampilan dalam Aktivitas Kuesioner .....	98
<b>Gambar 17.</b> <i>Flowchart</i> LMS Moodle AKM Numerasi .....	99
<b>Gambar 18.</b> Hasil Validasi Media CBT.....	102
<b>Gambar 19.</b> Nilai GOF Instrumen Soal AKM Numerasi .....	105
<b>Gambar 20.</b> Diagram Path CFA Soal Numerasi Paket A Tahap I.....	107
<b>Gambar 21.</b> Diagram Path CFA Soal Numerasi Paket A Tahap II .....	108
<b>Gambar 22.</b> Diagram Path CFA Soal Numerasi Paket B .....	108
<b>Gambar 23.</b> Diagram Path CFA Soal Numerasi Paket C .....	109
<b>Gambar 24.</b> Diagram Path CFA Soal Numerasi Paket D Tahap I.....	110
<b>Gambar 25.</b> Diagram Path CFA Soal Numerasi Paket D Tahap II .....	111
<b>Gambar 26.</b> Hasil Uji Realibilitas .....	111
<b>Gambar 27.</b> Hasil <i>Total Variance and Eigenvalues Explained</i> .....	113
<b>Gambar 28.</b> Hasil <i>Scree Plot Uji Unidimensi</i> .....	114
<b>Gambar 29.</b> Hasil Invariansi Parameter Butir .....	115
<b>Gambar 30.</b> Hasil Invariansi Parameter Kemampuan .....	116

<b>Gambar 31.</b> <i>Item Characteristics Curve</i> Paket A .....	124
<b>Gambar 32.</b> <i>Item Characteristics Curve</i> Paket B .....	126
<b>Gambar 33.</b> <i>Item Characteristics Curve</i> Paket C .....	128
<b>Gambar 34.</b> <i>Item Characteristics Curve</i> Paket D .....	130
<b>Gambar 35.</b> Hasil Fungsi Informasi dan Kesalahan Pengukuran .....	132
<b>Gambar 36.</b> Grafik Distribusi Kemampuan Peserta Didik .....	134
<b>Gambar 37.</b> TCC Keempat Paket Soal Sebelum Penyetaraan.....	140
<b>Gambar 38.</b> TCC Setelah Penyetaraan .....	142
<b>Gambar 39.</b> Respon Peserta Didik Terhadap Kepraktisan Soal .....	143
<b>Gambar 40.</b> Respon Pendidik Terhadap Kemanfaatan Soal.....	144

## DAFTAR LAMPIRAN

<b>Lampiran 1.</b> Perangkat Produk .....	182
A. <i>Blue Print</i> Instrumen AKM Numerasi .....	182
B. Kisi-kisi Instrumen AKM Numerasi .....	184
C. Instrumen AKM Numerasi .....	202
D. Pedoman Penskoran dan Kunci Jawaban .....	249
E. Panduan Media CBT .....	262
F. Instrumen Keefektifan Produk .....	309
<b>Lampiran 2.</b> Data Hasil Penelitian .....	313
A. Hasil Validasi Isi Instrumen AKM Numerasi .....	313
B. Hasil Penilaian Ahli Media CBT .....	339
C. Hasil Analisis Uji Coba Instrumen AKM Numerasi .....	345
D. Hasil Pengukuran .....	363
<b>Lampiran 3.</b> Persuratan .....	365
A. Surat Keterangan Validitas Instrumen AKM Numerasi .....	365
B. Surat Keterangan Vaiditas Media CBT .....	371
C. Surat Keterangan Telah Penelitian .....	373
<b>Lampiran 4.</b> Dokumentasi .....	379

# **BAB I**

## **PENDAHULUAN**

### **A. Latar Belakang Masalah**

Pendidikan merupakan pilar penting dalam mengembangkan kecerdasan intelektual setiap manusia. Pendidikan memiliki peran dalam menumbuhkan potensi diri setiap manusia melalui kecerdasan yang dimilikinya. Pendidikan pada ruang lingkup sekolah memiliki andil dalam mengembangkan potensi diri setiap peserta didik. Hal tersebut sejalan dengan Peraturan Pemerintah Republik Indonesia Nomor 57 tahun 2021 tentang Standar Nasional Pendidikan, yang menyatakan bahwa pendidikan merupakan usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik mampu secara aktif mengembangkan potensi diri untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa, maupun negara. Berpijak pada hal tersebut, maka sudah semestinya pelaksanaan pendidikan di sekolah dilaksanakan dengan sebaik-baiknya. Dalam mendukung keberhasilan pelaksanaan pendidikan di sekolah, terdapat faktor-faktor pendukung yang memiliki keterkaitan yang sangat erat, salah satu faktor tersebut yakni penilaian pendidikan.

Penilaian menjadi salah satu faktor penting dalam mendukung keberhasilan pelaksanaan pendidikan di sekolah karena informasi yang diperoleh dari kegiatan tersebut menjadi pedoman dalam memutuskan ketercapaian tujuan pendidikan (Kusainun, 2020, p. 1). Menurut Tosuncuoglu (2018, p. 1) penilaian tidak hanya memberikan informasi terkait keputusan intruksional yang dibuat setiap hari, dan membantu pendidik dalam mendiagnosis kekuatan dan kelemahan peserta didik dalam kegiatan belajar, namun penilain juga memberikan umpan balik kepada peserta didik dalam mendukung pembelajaran. Pusmenjar (2021b, p. 34) juga menegaskan bahwa umpan balik pelaksanaan penilaian tidak hanya untuk pendidik dan peserta didik, namun juga untuk orang tua peserta didik. Dengan adanya umpan balik tersebut, maka hasil penilaian diharapkan dapat dijadikan panduan dalam memperbaiki pelaksanaan pembelajaran, memperbaiki strategi belajar, hingga mampu meningkatkan prestasi belajar peserta didik.

Penilaian pendidikan pada abad ke-21 tentu tidak lepas dari kecakapan abad ke-21 itu sendiri, dimana antara penilaian dan tujuan pengembangan kecakapan haruslah selaras. Hal tersebut bertujuan agar penilaian secara eksplisit mengkomunikasikan sifat pembelajaran yang diharapkan (Griffin et al., 2012, p. 24). Pernyataan tersebut diperkuat dengan hasil penelitian Andrian & Rusman (2019, p. 21) dimana 57,9% penilaian pembelajaran abad 21 dalam pelaksanaan kurikulum 2013 sudah berjalan dengan baik. Kemudian dengan adanya *Covid-19* di tahun 2020, maka pelaksanaan pendidikan mengalami perubahan dan penyesuaian ulang, seperti halnya penerapan kurikulum 2013 yang disederhanakan atau disebut dengan kurikulum darurat dengan pengurangan beban belajar dan pemangkasan waktu belajar peserta didik. Hal tersebut berdampak pada proses penilaian pendidikan.

Salah satu penelitian yang melaporkan upaya pendidik dalam melakukan penilaian selama pembelajaran daring atau masa *Covid-19* adalah penelitian Fakhrunisa & Prabawanto (2020), dimana pendidik memberikan penilaian (1) berupa masukan pada tugas kinerja; (2) berupa poin lebih tinggi untuk peserta didik tercepat dalam mengumpulkan jawaban; dan (3) berdasarkan aktivitas peserta didik selama diskusi *online*. Hasil penelitian tersebut menegaskan bahwa banyak langkah yang telah diupayakan pendidik agar proses penilaian pembelajaran selama *Covid-19* berjalan dengan baik, seperti halnya pelaksanaan penilaian pembelajaran sebelum *Covid-19*. Sejak memasuki abad ke-21 hingga sekarang ini, sudah banyak transformasi sistem pendidikan yang telah dilakukan terutama dalam hal penilaian, namun hal yang tidak pernah berubah hingga sekarang ini yakni berkaitan dengan salah satu kompetensi dasar yang harus dimiliki peserta didik, yaitu numerasi.

Numerasi merupakan salah satu dari enam literasi dasar yang harus selalu ditingkatkan peserta didik agar mampu menghadapi segala tantangan pada abad ke-21. *United Nations Educational, Scientific, and Cultural Organization* (UNESCO) di tahun 2006 menegaskan bahwa salah satu kemampuan yang dapat menentukan kemajuan dari suatu bangsa adalah numerasi (Hendrowati & Faelasofi, 2021, p. 366; Kemdikbud, 2017, p. 2). Menurut Nortvedt & Wiese (2020, p. 529) numerasi merupakan kemampuan dalam mengakses, menggunakan, menafsirkan, serta

mengkomunikasikan informasi dan ide-ide matematika dalam memecahkan masalah sehari-hari. Ide-ide matematika yang dimaksud berupa adanya keterlibatan angka, data, dan simbol matematika (Anderha & Maskar, 2021, p. 2; Barham et al., 2019, p. 176; Kustantina et al., 2022, p. 11) yang tersajikan dalam grafik, bagan, tabel, dll (Geiger et al., 2015, p. 533). Sebagai suatu kemampuan yang dibutuhkan dalam mendukung aktivitas sehari-hari, kemampuan numerasi yang baik sangatlah diperlukan, karena Hendrowati & Faelasofi (2021, p. 365) menjelaskan bahwa tanpa sadar kemampuan numerasi tertanam dalam aktivitas sehari-hari seperti pada pendidikan, kegiatan perekonomian, kesehatan, dll. Hal tersebut menegaskan bahwa kemampuan numerasi sangat penting dan perlu perhatian lebih terutama dalam dunia pendidikan.

Pentingnya numerasi pada pendidikan terletak bagaimana OECD (*Organization for Economic Cooperation and Development*) pada tahun 1997 menggagas suatu penilaian skala internasional yang dilaksanakan setiap tiga tahun sekali, dan putaran pertama terlaksana pada tahun 2000. Penilaian tersebut dikenal dengan sebutan PISA (*Programme for International Students Assessment*). Geiger et al. (2015, p. 533) menyatakan bahwa adanya PISA memberikan definisi tersendiri terhadap numerasi atau literasi matematika yang merupakan bagian relevan dari penilaian, serta kerangka analitis yang digunakan untuk mengembangkan instrumen penilaian numerasi itu sendiri. Numerasi yang dimaksud dalam PISA mencakup logika matematika serta penggunaan konsep, prosedur, fakta, dan perangkat matematika yang digunakan untuk penggambaran, penguraian, serta perkiraan suatu masalah dalam berbagai konteks sehari-hari (Kemdikbud, 2019, p. 20). Hal tersebut menunjukkan bahwa harapan dari adanya numerasi pada pelaksanaan PISA yakni makin mampunya peserta didik dalam memperkuat dan memperluas kemampuan analisisnya berdasarkan pengetahuan yang dimiliki serta mengimplementasikannya pada konteks masalah yang tidak biasa.

Data OECD menunjukkan bahwa rata-rata numerasi PISA untuk seluruh peserta dari berbagai negara bergerak fluktuatif. Dari putaran pertama tahun 2000 hingga putaran ketujuh tahun 2018, rata-rata numerasi tertinggi ada pada tahun

2003 yakni sebesar 499, selisih 10 poin dari tahun 2018 yang sebesar 489 (OECD, 2018). Kefluktuatifan tersebut merupakan dampak dari setiap negara peserta yang memiliki prioritas penilaian tersendiri dalam hal ini numerasi, sehingga keragaman kemampuan numerasi antar peserta menjadi hal yang lumrah. Disamping itu, pemerintah Indonesia melalui Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi (Kemdikbudristek) juga memperhatikan hasil partisipasi Indonesia dalam pelaksanaan PISA dari tahun ke tahun untuk bidang numerasi. Pada tahun 2000 Indonesia berada pada peringkat 39 dari 41 negara dengan rata-rata 367, tahun 2003 peringkat 38 dari 40 negara dengan rata-rata 360, tahun 2006 peringkat 50 dari 56 negara dengan rata-rata 391, tahun 2009 peringkat 61 dari 65 negara dengan rata-rata 371, tahun 2012 peringkat 64 dari 65 negara dengan rata-rata 375, tahun 2015 peringkat 62 dari 70 negara dengan rata-rata 386, dan tahun 2018 peringkat 72 dari 77 negara dengan rata-rata 379 (Hewi & Shaleh, 2020; Kemdikbud, 2019, p. 42). Data tersebut memberi pertanda bahwa perlu ada terobosan baru terkait penilaian numerasi sehingga ada upaya dalam melatih hingga meningkatkan kemampuan numerasi secara signifikan.

Pada tahun 2021 Kemdikbudristek menerapkan terobosan baru terkait pelaksanaan penilaian skala nasional, dimana mengganti UN dengan penilaian baru yang disebut dengan Asesmen Nasional (AN) yang salah satu aspek ujinya yakni Asesmen Kompetensi Minimum (AKM), dengan salah satu kompetensi dasar yang diukur adalah numerasi. Rapor pendidikan untuk AKM numerasi tahun 2021 menunjukkan bahwa secara keseluruhan baik jenjang SD/Sederajat, SMP/Sederajat, dan SMA/SMK/Sederajat kemampuan numerasi peserta didik di bawah kompetensi minimum atau kurang dari 50% peserta didik telah mencapai kompetensi minimum untuk numerasi. Sedangkan hasil AKM numerasi masing-masing Provinsi untuk ketiga jenjang yakni (1) untuk jenjang SD/Sederajat, satu dari 34 Provinsi yakni Maluku Utara hasilnya jauh di bawah kompetensi minimum atau sebagian besar peserta didik belum mencapai batas kompetensi minimum untuk numerasi, sedangkan 33 Provinsi lain hasilnya di bawah kompetensi minimum; (2) untuk jenjang SMP/Sederajat, dua dari 34 Provinsi yakni DI Yogyakarta dan DKI Jakarta hasilnya mencapai kompetensi minimum atau

sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi, namun perlu upaya mendorong lebih banyak peserta didik menjadi mahir, sedangkan 32 Provinsi lain hasilnya di bawah kompetensi minimum; dan (3) untuk jenjang SMA/SMK/Sederajat, 4 dari 34 Provinsi yakni DI Yogyakarta, DKI Jakarta, Jawa Tengah, dan Kepulauan Riau hasilnya mencapai kompetensi minimum, sedangkan 30 Provinsi lain hasilnya di bawah kompetensi minimum. Data tersebut menunjukkan bahwa numerasi di Indonesia melalui pelaksanaan AKM masih sangat rendah bahkan belum cukup merata dari ke-34 Provinsi yang ada.

Rendah dan belum cukup meratanya numerasi peserta didik se-Indonesia melalui pelaksanaan AKM mendapat perhatian dari beberapa peneliti. Kustantina et al. (2022, p. 11) menyatakan bahwa salah satu penyebab rendahnya numerasi adalah kurangnya motivasi belajar peserta didik selama pembelajaran matematika. Peserta didik dengan asumsi bahwa matematika identik dengan pelajaran sulit juga berimbas pada numerasi, yang mana keduanya adalah dua hal yang saling berkaitan. Hal tersebut sejalan dengan pernyataan Anam et al. (2019, p. 1) bahwa numerasi merupakan subjek yang sulit dipahami tanpa adanya fungsi dan kegunaan. Sehingga peserta didik mengalami ketidakpedulian, kebingungan, serta kelelahan setiap diberi pelajaran numerasi. Pada dasarnya, harapan pemerintah pada pelaksanaan AKM yakni menghasilkan informasi yang memicu adanya peningkatan kualitas pembelajaran dan kemampuan peserta didik, sehingga dapat meningkatkan hasil belajar (Purnomo et al., 2022, p. 636), namun tidak dapat dijadikan dasar seleksi ke jenjang selanjutnya (Herman et al., 2022, p. 1). Untuk mewujudkan harapan tersebut, melalui AKM pemerintah mencoba mengetahui sejauh mana kemampuan peserta didik melalui proses kognitif yang diukur yakni pemahaman, penerapan, serta penalaran dengan berbagai variasi bentuk soal.

Sebagai penilaian skala nasional, AKM merupakan penilaian pertama yang dalam pelaksanaannya terdapat berbagai variasi bentuk soal lebih dari dua. Karena sebelumnya pada pelaksanaan UN hanya terdapat dua variasi soal yakni pilihan ganda dan isian singkat (Meriana et al., 2021, p. 111; Widarti et al., 2022, p. 66). Variasi bentuk soal yang dimaksudkan pada AKM yakni pilihan ganda, pilihan

ganda kompleks, menjodohkan, isian singkat, serta uraian (Aufa & Manoy, 2022, p. 220). Adanya variasi tersebut memacu peserta didik agar makin terampil dalam memahami dan menganalisa soal, bahkan hasil penelitian Adelia & Deta (2022, p. 7) mengungkapkan bahwa AKM mampu mengukur kemampuan tingkat tinggi peserta didik. Berkenaan dengan ketersediaan variasi bentuk soal untuk persiapan pelaksanaan AKM 2021 lalu, Purnomo et al. (2022, p. 637) melaporkan bahwa pendidik berupaya untuk mengembangkan instrumen AKM numerasi sendiri namun kenyataannya instrumen tersebut tidak dapat dipergunakan untuk mengukur numerasi sesuai dengan kriteria yang ada pada AKM. Selain itu hasil penelitian Widarti et al. (2022, p. 66) juga menegaskan bahwa ketersediaan variasi bentuk soal AKM dari berbagai sumber yang dipergunakan pendidik untuk persiapan AKM persentasenya masing-masing di bawah 50%. Hal tersebut menunjukkan bahwa belum cukup tersedianya variasi bentuk soal AKM dalam hal ini numerasi yang bisa dijadikan bahan referensi untuk mempersiapkan pelaksanaan AKM.

Permasalahan di atas tidak bisa didiamkan begitu saja karena tentunya akan menghambat persiapan peserta didik ataupun pendidik dalam pelaksanaan AKM. Oleh karena itu ada upaya yang perlu dilakukan untuk mengatasi masalah tersebut yakni dengan menyediakan soal AKM numerasi dengan melakukan pengembangan soal sesuai dengan kerangka kerja atau standar AKM. Pengembangan soal ini merupakan usaha yang tepat dalam membantu pendidik dalam menyediakan kebutuhan soal AKM numerasi, terlebih untuk melatih peserta didik dalam mempersiapkan pelaksanaan AKM. Bellini et al. (2019, p. 9) menegaskan bahwa pengembangan soal numerasi yang memadai merupakan aset sistem penilaian di Indonesia karena AKM menjadi rangkaian asesmen nasional yang berjalan hingga beberapa tahun kedepan. Berbagai upaya telah dilakukan beberapa peneliti dalam mengembangkan soal AKM numerasi yakni Machromah et al. (2021), Nuzulia & Gafur (2022), Purnomo et al. (2022), serta Sulistyani & Kusumawardana (2022). Penelitian-penelitian tersebut ditujukan untuk jenjang SD/Sederajat dan SMP/Sederajat, namun belum pernah ditemui pengembangan soal AKM numerasi untuk jenjang SMA/SMK/Sederajat, sehingga untuk penelitian ini akan difokuskan kepada jenjang tersebut terlebih untuk kelas XI sebagai target pelaksanaan AKM.

Semakin berkembangnya zaman yang diiringi dengan teknologi yang semakin berkembang pesat sebagai akibat dari gelombang revolusi industri 4.0 tentu berdampak pada pemanfaatan komputer dalam pelaksanaan pendidikan, terlebih dalam penilaian. Hal tersebut sejalan dengan Khoshsima et al. (2019, p. 86) yang menyatakan bahwa teknologi komputer telah banyak dipergunakan dalam tes bahasa untuk menganalisis skor dan hasil dengan cepat. Selain itu sejak adanya *Covid-19* pelaksanaan penilaian pendidikan serba dilaksanakan secara *online*, hal tersebut tentu sudah menjadikan hal yang lumrah terkait adanya alat ukur sebagai bentuk penilaian berbasis komputer. Penilaian berbasis komputer yang kemudian dikenal dengan *Computerized Based Test* (CBT) dikembangkan untuk memberikan kemudahan dalam sistem penilaian pendidikan. Hal tersebut mempertegas bahwa teknologi sudah mampu berdampingan dengan pelaksanaan pendidikan.

Media CBT dalam penerapan penilaian pendidikan dapat dilakukan menggunakan *Learning Management System* (LMS). LMS merupakan perangkat lunak berbasis internet yang dipergunakan untuk pengelolaan *e-learning* sehingga memungkinkan adanya administrasi, penyampaian kursus (konten pembelajaran), pengujian, dan lainnya (Cavus, 2015, p. 873; Xin & Singh, 2021, p. 838). Adanya penggunaan LMS membuat peserta didik dapat mengakses informasi terkait pembelajarannya kapanpun dan dimanapun (Alkhateeb & Abdalla, 2021, p. 138). Salah satu *platform* LMS yang efisien dan populer dalam dunia pendidikan adalah Moodle.

Moodle merupakan LMS yang bersifat *open source* dengan bebas penggunaan tanpa ada tarif pembayaran. Moodle dikembangkan menggunakan bahasa pemrograman PHP dan 2002 merupakan tahun pertama perilisan. Pada tahun 2021 pengguna pendaftar Moodle mencapai lebih dari 278 juta pengguna (Xin & Singh, 2021, p. 838). Hal tersebut dapat terjadi karena Moodle memiliki fleksibilitas dan skalabilitas yang sangat tinggi disesuaikan untuk mendukung kebutuhan kelas kecil maupun organisasi besar dalam pelaksanaan pendidikan (Al-Hamad, 2022, pp. 2–3). Alkhateeb & Abdalla (2021, p. 139) menjelaskan bahwa *platform* Moodle memungkinkan adanya pembuatan tes dan kuesioner serta pengelolaannya. Berkenaan dengan itu pengembangan soal yang akan dilakukan

mengarah pada bentuk *Computer Based Test* (CBT) berbantuan LMS Moodle. Fungsi CBT dalam pengembangan soal ini sebagai alat menorehkan jawaban dengan lima variasi bentuk soal yang akan dikembangkan. Dipilihnya LMS Moodle sebagai media CBT karena pada Moodle terdapat fitur simbol matematika, *equation*, serta fitur unggah gambar untuk melampirkan penyelesaian, terlebih dalam soal yang dikembangkan terdapat bentuk soal uraian. Berdasarkan uraian di atas penulis termotivasi untuk melakukan penelitian dengan judul “Pengembangan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA”.

## **B. Identifikasi Masalah**

Berdasarkan latar belakang masalah di atas, diperoleh beberapa identifikasi masalah yang muncul yakni sebagai berikut.

1. Penilaian pendidikan merupakan aspek penting dalam mendukung keberhasilan serta mengukur ketercapaian tujuan pendidikan. Maka dari itu banyaknya transformasi sistem pendidikan yang telah diterapkan juga berdampak pada transformasi penilaian. Salah satu halnya yakni transformasi penilaian yang dilakukan semenjak adanya *Covid-19*, dimana bentuk penilaian pembelajaran harus disesuaikan ulang dan diefisiensikan dengan pelaksanaan pembelajaran;
2. Numerasi merupakan kemampuan penting yang harus dimiliki peserta didik sebagai bekal menghadapi segala tantangan pada abad ke-21, namun realita menunjukkan bahwa kemampuan numerasi peserta didik di Indonesia tergolong sangat rendah jika dibandingkan dengan negara-negara peserta PISA dari setiap periode pelaksanaannya;
3. AKM merupakan terobosan baru dalam penilaian di Indonesia dimana hasil yang diperoleh dijadikan acuan dalam meningkatkan kualitas pembelajaran dan kemampuan peserta didik, namun tidak dipergunakan sebagai dasar seleksi jenjang sekolah. Realita menunjukkan bahwa hasil AKM dapat memetakan sekolah serta daerah-daerah berdasarkan kompetensi minimum yang dihasilkan;
4. Hasil pemetaan numerasi untuk setiap Provinsi disetiap jenjang sekolah menunjukkan bahwa mayoritas kemampuan numerasi peserta didik di bawah kompetensi minimum, akan tetapi beberapa Provinsi menunjukkan hasil yang

cukup bagus dimana kemampuan numerasi sudah mencapai kompetensi minimum;

5. Pendidik maupun peserta didik memerlukan instrumen AKM numerasi untuk mempersiapkan pelaksanaan AKM disemua jenjang pendidikan, akan tetapi dari berbagai sumber, instrumen yang dibutuhkan masih terbatas terlebih untuk jenjang SMA/Sederajat. Selain itu pada persiapan pelaksanaan AKM 2021 pendidik melakukan upaya untuk mengembangkan instrumen AKM numerasi, akan tetapi instrumen tidak dapat dipergunakan untuk mengukur numerasi peserta didik karena tidak sesuai kriteria yang ada pada AKM;

### **C. Pembatasan Masalah**

Berdasarkan identifikasi masalah di atas, peneliti membatasi permasalahan agar penelitian yang dilakukan lebih spesifik dan terarah, sehingga pembatasan masalahnya ada pada poin 5, yakni dilakukan penelitian dengan mengembangkan soal AKM numerasi untuk kelas XI SMA yang disesuaikan dengan kerangka kerja AKM serta uji coba dilaksanakan berbasis komputer (CBT) berbantuan LMS Moodle.

### **D. Rumusan Masalah**

Berdasarkan identifikasi masalah dan pembatasan masalah di atas, maka rumusan masalah dalam penelitian ini antara lain:

1. Bagaimana konstruk soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
2. Bagaimana kualitas soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
3. Bagaimana kesetaraan soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
4. Bagaimana profil peserta didik kelas XI SMA dalam menyelesaikan soal AKM numerasi berbasis komputer yang dikembangkan?
5. Bagaimana keefektifan produk soal AKM numerasi berbasis komputer yang telah dikembangkan?

### **E. Tujuan Pengembangan**

Berdasarkan rumusan masalah di atas, tujuan dari penelitian pengembangan ini antara lain:

1. Mendesain konstruk soal AKM numerasi berbasis komputer untuk kelas XI SMA;
2. Menganalisis kualitas soal AKM numerasi berbasis komputer untuk kelas XI SMA;
3. Mengetahui kesetaraan soal AKM numerasi berbasis komputer untuk kelas XI SMA yang telah dikembangkan;
4. Mendeskripsikan profil peserta didik kelas XI SMA dalam menyelesaikan soal AKM numerasi berbasis komputer;
5. Mengetahui keefektifan produk soal AKM numerasi berbasis komputer yang telah dikembangkan.

### **F. Spesifikasi Produk yang Dikembangkan**

Spesifikasi produk yang dikembangkan pada penelitian ini adalah perangkat soal AKM numerasi untuk kelas XI SMA yang meliputi kisi-kisi soal, item soal, alternatif jawaban, serta rubrik penskoran. Item soal yang dikembangkan tersusun dalam 4 paket soal dengan masing-masing paket terdapat 16-17 soal dengan 4 item *anchor* disetiap paket. Variasi bentuk soal meliputi pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, serta uraian. Kemudian kompetensi yang diujikan disesuaikan dengan kisi-kisi AKM numerasi jenjang SMA/SMK/Sederajat seperti Tabel 1. Produk yang dikembangkan diuji cobakan secara CBT berbasis *web* yakni menggunakan *Learning Management System (LMS) Moodle*.

**Tabel 1.** *Kisi-kisi Numerasi Jenjang SMA/SMK/Sederajat (Pusmenjar, 2020, p. 95)*

Domain	Subdomain	Kompetensi
Aljabar	Persamaan dan pertaksamaan	Menyelesaikan persamaan dan pertaksamaan kuadrat
		Menyelesaikan sistem persamaan linear dua atau tiga variabel
	Relasi dan fungsi (termasuk pola bilangan)	Memahami barisan aritmatika dan geometri Memahami fungsi kuadrat dan grafiknya, serta sifat-sifatnya
Geometri dan Pengukuran	Bangun geometri	Menggunakan perbandingan trigonometri Menghitung volume dan luas permukaan limas segi- $n$ , kerucut, dan bola
Data dan ketidakpastian	Data dan representasinya	Menginterpretasikan penyajian data Menentukan ukuran penyebaran data (jangkauan, simpangan, dan variansi)
	Ketidakpastian dan peluang	Menggunakan sifat-sifat peluang suatu kejadian Menggunakan sifat-sifat peluang kejadian majemuk

### G. Manfaat Penelitian

Adanya penelitian ini diharapkan dapat memberikan manfaat terhadap pelaksanaan asesmen pembelajaran terutama pada pembelajaran matematika yang masih memiliki keterkaitan dengan numerasi. Adapun manfaat secara teoritis, metodologis, dan praktis dari penelitian ini antara lain:

1. Secara teoritis, penelitian ini diharapkan menjadi salah satu bahan dalam pengembangan khazanah ilmu pengetahuan, terkhusus dalam pengembangan soal AKM numerasi yang dapat menstimulasi dengan baik kemampuan tingkat tinggi peserta didik, terutama jika dikaitkan dengan konteks AKM numerasi yang orientasinya pada permasalahan di lingkungan sekitar;

2. Secara metodologis, penelitian ini diharapkan menjadi bahan rujukan untuk mengembangkan soal AKM dengan konteks pembelajaran atau tingkat pendidikan yang berbeda;
3. Secara praktis, penelitian ini diharapkan memberi manfaat kepada pendidik dan peserta didik.
  - a. Bagi pendidik, penelitian ini diharapkan menjadi alat identifikasi sejauh mana pemahaman peserta didik dalam menyelesaikan soal AKM numerasi. Hasil dari penggunaan soal ini juga diharapkan mampu meningkatkan proses pembelajaran, serta menjadi bahan rujukan dalam mempersiapkan peserta didik untuk pelaksanaan AKM;
  - b. Bagi peserta didik, penelitian ini diharapkan menjadi informasi terkait sejauh mana kemampuan peserta didik dalam menyelesaikan soal AKM numerasi, serta memberi pengalaman kepada peserta didik dalam mengerjakan soal tersebut berbasis komputer;

#### **H. Asumsi Pengembangan**

Asumsi-asumsi pada penelitian pengembangan soal AKM numerasi untuk kelas XI SMA berbasis komputer antara lain:

1. Proses pengembangan soal AKM numerasi disusun dengan langkah yang sistematis;
2. Soal AKM numerasi yang dikembangkan sesuai dengan kerangka kerja AKM numerasi yang telah ditetapkan pemerintah;
3. Instrumen divalidasi oleh validator yang memiliki pengetahuan mengenai kriteria soal AKM numerasi dan instrumen yang baik.

## **BAB II**

### **KAJIAN PUSTAKA**

#### **A. Kajian Teori**

##### **1. Asesmen**

Asesmen atau penilaian merupakan hal yang cukup identik dalam dunia pendidikan, karena dari pelaksanaannya akan diperoleh informasi-informasi penting yang dapat dijadikan acuan dalam melihat sejauh mana kemampuan peserta didik. Amnie et al. (2021, p. 1) mendefinisikan asesmen sebagai salah satu cara untuk meningkatkan kualitas hasil pendidikan. Hal tersebut dapat terjadi karena penilaian menjadi sebuah alat ukur untuk mengetahui pencapaian pemahaman dan tingkat kemampuan peserta didik. Sejalan dengan hal tersebut, Looney et al. (2018, p. 442) berpendapat bahwa asesmen sebagai proses untuk meninjau apa yang telah diketahui dan kemampuan apa yang sudah tercapai oleh peserta didik. Sehingga kemudian dapat dinyatakan bahwa salah satu komponen utama pada proses pembelajaran hingga evaluasi adalah asesmen.

Asesmen berdasarkan Keputusan Menteri Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia Nomor 56/M/2022 tentang Pedoman Penerapan Kurikulum dalam Rangka Pemulihan Pembelajaran merupakan sebuah proses pengumpulan dan pengolahan informasi untuk mengukur pencapaian hasil belajar peserta didik. Selain itu, Mardapi (2016, pp. 10–11) juga mengemukakan bahwa asesmen mencakup prosedur dalam pengumpulan data yang dilakukan terhadap peserta didik. Dua definisi di atas kemudian diperjelas oleh Kusainun (2020, p. 1) bahwa asesmen merupakan proses terencana dan sistematis yang dilakukan pendidik dalam mengambil keputusan tentang keberhasilan atau ketercapaian tujuan pendidikan dalam hal ini pada proses pembelajaran. Proses terencana dan sistematis yang dimaksud meliputi proses mengumpulkan data atau informasi dari peserta didik, menganalisis, serta menginterpretasi.

Tujuan dari adanya pelaksanaan asesmen menurut Fong & Kaur (2015, pp. 22–23) yaitu: (1) sumber informasi bagi pendidik dalam berinovasi dan meningkatkan pembelajaran di kelas, sehingga peserta didik dapat mempelajari

pengetahuan dengan lebih baik; (2) bahan pertimbangan dalam pengambilan kebijakan pendidik pada suatu negara; dan (3) sumber informasi bagi orang tua dan peserta didik sendiri mengenai hasil kemajuan belajar peserta didik. Sedangkan prinsip-prinsip asesmen pada penerapan kurikulum merdeka sebagai kurikulum baru yakni sebagai berikut (Anggraena et al., 2022, pp. 8–9).

- a. Asesmen merupakan bagian terpadu pada proses pembelajaran, fasilitasi pembelajaran, serta penyediaan informasi yang holistik, sebagai umpan balik untuk pendidik, peserta didik, dan orang tua agar dapat memandu peserta didik dalam menentukan strategi pembelajaran selanjutnya. Contoh implementasinya yakni pendidik merancang asesmen untuk mendorong peserta didik untuk terus meningkatkan kompetensinya melalui asesmen dengan tingkat kesulitan yang tepat dan umpan balik yang membangun;
- b. Asesmen dirancang dan dilakukan sesuai dengan fungsi asesmen, dengan keleluasaan pendidik dalam menentukan waktu dan teknik pelaksanaan asesmen agar efektif mencapai tujuan pembelajaran. Contoh implementasinya yakni pendidik memikirkan tujuan pembelajaran pada saat merencanakan asesmen dan memberikan kejelasan pada peserta didik mengenai tujuan asesmen di awal pembelajaran;
- c. Asesmen didesain secara adil, proporsional, valid, serta dapat dipercaya untuk menjelaskan kemajuan belajar, menentukan keputusan tentang langkah dan sebagai dasar untuk menyusun program pembelajaran selanjutnya yang sesuai. Contoh implementasinya yakni pendidik menyediakan waktu dan durasi yang cukup agar asesmen menjadi sebuah proses pembelajaran dan bukan hanya untuk kepentingan menguji;
- d. Laporan kemajuan belajar dan pencapaian peserta didik bersifat sederhana dan informatif, memberikan informasi yang bermanfaat tentang karakter dan kompetensi yang dicapai, serta strategi tindak lanjut. Contoh implementasinya yakni pendidik menyusun laporan kemajuan belajar

secara ringkas, mengutamakan informasi yang paling penting untuk dipahami oleh peserta didik dan orang tua;

- e. Hasil asesmen digunakan oleh peserta didik, pendidik, tenaga kependidikan, serta orang tua sebagai bahan refleksi untuk meningkatkan mutu pembelajaran. Contoh implementasinya yakni pendidik memberikan umpan balik secara berkala kepada peserta didik dan mendiskusikan tindak lanjutnya bersama orang tua.

Berdasarkan beberapa definisi yang dinyatakan oleh beberapa ahli di atas, selanjutnya dapat disimpulkan bahwa asesmen merupakan serangkaian proses terpadu dalam pembelajaran yang meliputi proses mengumpulkan informasi, menganalisis, serta menginterpretasikan informasi sehingga diperoleh hasil berupa keputusan terkait ketercapaian belajar peserta didik. Tidak berhenti pada proses tersebut, keputusan yang didapat dari asesmen dijadikan sebagai umpan balik dan bahan refleksi baik untuk peserta didik, pendidik, bahkan orang tua agar dapat dicapai mutu pembelajaran yang lambat laun semakin meningkat.

## **2. Asesmen Kompetensi Minimum**

### **a. Konsep Dasar AKM**

Kemdikbudristek tahun 2021 melaksanakan asesmen baru sebagai pengganti Ujian Nasional (UN) dengan istilah Asesmen Nasional (AN). Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021 menyatakan bahwa AN merupakan evaluasi pemerintah dalam memetakan mutu sistem pendidikan pada tingkat satuan pendidikan dasar dan menengah. Pelaksanaan AN mencakup tiga kegiatan penting yakni survei karakter, survei lingkungan belajar, dan asesmen kompetensi minimum (Herman et al., 2022, p. 4; Widiyanti & Hidayati, 2021, p. 28). Pada bagian ini akan diperinci penjelasan mengenai AKM yang menjadi fokus utama pada penelitian.

Asesmen kompetensi minimum yang selanjutnya dikenal dengan sebutan AKM merupakan salah satu dari cakupan pembelajaran paradigma baru yang orientasinya mengarah kepada praktik pembelajaran yang berpusat pada peserta didik. Dimana pendidik memiliki ruang yang lebih leluasa dalam merancang asesmen yang sesuai dengan kebutuhan dan karakteristik peserta didik (Pusmenjar, 2021b). AKM didefinisikan sebagai penilaian kemampuan mendasar yang diperlukan peserta didik agar mampu mengembangkan kapasitas diri serta berpartisipasi positif dalam masyarakat (Herman et al., 2022, p. 4; Meriana et al., 2021, p. 113; Widarti et al., 2022, p. 63). Selain itu, Andiani et al. (2020, p. 82) mendefinisikan AKM sebagai penilaian dengan tujuan yang diharapkan adalah mewujudkan kecakapan hidup abad ke-21 yang seharusnya dapat dimiliki semua peserta didik. Hal tersebut sejalan dengan pernyataan Aufa & Manoy (2022, p. 220) bahwa dengan adanya AKM peserta didik dapat berpikir kritis dalam kehidupan sehari-hari.

Kemdikbudristek merancang AKM sebagai tanggapan atas kebutuhan global saat ini hingga masa depan. Dimana peserta didik diharapkan mampu beradaptasi secara cepat terhadap perubahan yang ada di masyarakat serta bisa berpartisipasi aktif didalamnya (Pusmenjar, 2021a, pp. 1–2). Sehingga dengan sadar, peserta didik dituntut menjadi pembelajar sepanjang hayat. Disamping itu, adanya AKM memiliki tujuan yang sangat baik untuk kemajuan negara, yakni dapat meningkatkan kualitas pendidikan di Indonesia (Herman et al., 2022, p. 4). Berkaitan dengan pernyataan tersebut, Meriana et al. (2021, pp. 112–113) mempertegas bahwa hasil AKM dapat memetakan sekolah-sekolah dan daerah-daerah berdasarkan kompetensi minimum yang dihasilkan. Dengan demikian, setiap sekolah ataupun daerah belum tentu memiliki hasil yang sama, sehingga akan berdampak pada adanya keragaman perbaikan kualitas sistem pendidikan.

Berdasarkan beberapa definisi yang dinyatakan oleh beberapa ahli di atas, selanjutnya dapat disimpulkan bahwa AKM merupakan

serangkaian penilaian kompetensi mendasar peserta didik yang dilaksanakan untuk melatih dan mengembangkan kapasitas diri sesuai dengan kebutuhan global serta mampu berkontribusi baik dalam masyarakat.

#### **b. Komponen AKM**

Sebagai suatu rangkaian penilaian, AKM memiliki ruang lingkup yang diujikan kepada peserta didik. Kemdikbudristek menetapkan bahwa AKM ditujukan untuk mengukur hasil belajar kognitif peserta didik dalam literasi membaca dan numerasi. Berkaitan dengan hal tersebut Herman et al. (2022, p. 1) menjelaskan bahwa ruang lingkup pada AKM diadaptasi dari tes PISA yang telah ada sejak tahun 2000. Berikut adalah pemaparan mengenai literasi membaca dan numerasi yang ada pada AKM.

##### 1) Literasi membaca

*National Institute for Literacy* mendefinisikan literasi sebagai suatu kemampuan dalam membaca, menulis, berbicara, menghitung, serta memecahkan masalah pada level kemampuan yang dibutuhkan dalam pekerjaan, keluarga, dan lingkungan sosial (Pinastiti et al., 2022, p. 204). Berawal dari definisi dasar tersebut, kemudian PISA memberikan definisi untuk literasi membaca sebagai kemampuan peserta didik dalam memahami, menggunakan, merefleksikan, serta melakukan interaksi dengan teks tulis yang tersajikan agar dapat meraih tujuan pribadi, pengembangan pengetahuan dan potensi, dengan hasil peserta didik mampu berpartisipasi positif sebagai masyarakat dalam suatu negara (Pusmenjar, 2021a, p. 9). Sedangkan Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021 menyatakan bahwa literasi membaca dalam AKM merupakan kemampuan peserta didik dalam memahami, menggunakan, mengevaluasi, merefleksikan berbagai jenis teks untuk memecahkan masalah serta mengembangkan

kapasitas diri dan berkontribusi secara produktif dalam bermasyarakat secara nasional maupun global.

## 2) Numerasi

Secara terminologi numerasi berkaitan dengan kemampuan dalam menggunakan matematika pada konteks non-matematika baik dalam kehidupan sehari-hari secara pribadi ataupun dalam masyarakat umum (Queensland, 2015, p. 2). Kemudian PISA memberikan definisi terhadap numerasi sebagai kemampuan dalam menggunakan, menerapkan, menginterpretasi, serta mengkomunikasikan ide dan informasi matematika dalam kehidupan sehari-hari (Geiger et al., 2015, p. 534). Sedangkan Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021 menyatakan bahwa numerasi dalam AKM merupakan kemampuan berpikir dalam menggunakan konsep, prosedur, fakta, serta alat matematika yang digunakan dalam memecahkan masalah sehari-hari pada berbagai konteks yang relevan untuk individu sebagai warga Indonesia maupun global. Mengenai konsep dasar numerasi akan dipaparkan lebih rinci di sub-bab berikutnya mengingat numerasi merupakan fokus penelitian ini.

Selain dua ruang lingkup di atas, pada AKM juga terdapat komponen-komponen penting yang mendukung tujuan dari pelaksanaan AKM yakni konten, konteks, serta level kognitif (Rohim, 2021, pp. 58–59). Berikut adalah pemaparan mengenai ketiga komponen AKM tersebut.

### 1) Konten

Konten merupakan pengetahuan yang mencakup sebuah fakta, konsep, teori, serta prinsip pada suatu hal yang menjadi topik pembahasan. Karena pada pelaksanaan AKM terbagi menjadi dua ruang lingkup, maka secara garis besar komponen konten juga terbagi menjadi dua domain yakni konten domain pada literasi membaca dan konten domain pada numerasi. Tabel 2 adalah beberapa konten domain yang diujikan pada literasi membaca dan numerasi berdasarkan Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021.

**Tabel 2.** *Konten Pada Literasi Membaca dan Numerasi*

<b>Ruang Lingkup AKM</b>	<b>Konten Domain yang Diujikan</b>
Literasi membaca	Teks sastra/fiksi; Teks informasi
Numerasi	Bilangan; Aljabar; Geometri dan pengukuran; Data dan ketidakpastian

### 2) Konteks

Meskipun dalam pelaksanaan AKM terbagi menjadi dua ruang lingkup, namun untuk komponen konteks antara literasi membaca dan numerasi adalah sama yakni mencakup konteks personal, sosial budaya, dan saintifik. Hal tersebut telah ditetapkan dalam Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021.

### 3) Level kognitif

Level kognitif merupakan level berpikir yang mencakup kegiatan mengingat, memahami, menerapkan, menganalisis, mengevaluasi, serta menciptakan (Widiana et al., 2018, p. 260). Level kognitif yang dimaksudkan dalam AKM adalah level proses berpikir peserta didik

dalam memecahkan masalah-masalah sehari-hari yang mencakup kegiatan-kegiatan yang dimaksudkan di atas. Kemudian karena pada pelaksanaan AKM terbagi menjadi dua ruang lingkup, maka secara garis besar komponen level kognitif juga terbagi menjadi dua, karena level kognitif berkaitan erat dengan konten domain yang diujikan. Level kognitif yang dimaksud yakni level kognitif pada literasi membaca dan level kognitif pada numerasi. Tabel 3 adalah beberapa level kognitif yang diujikan pada literasi membaca dan numerasi berdasarkan Peraturan Kepala Badan Standar, Kurikulum, dan Asesmen Pendidikan Kemdikbudristek Nomor 030/H/PG.00/2021 tentang Prosedur Operasional Standar Penyelenggaraan Asesmen Nasional tahun 2021.

**Tabel 3.** *Level Kognitif Pada Literasi Membaca dan Numerasi*

Ruang Lingkup AKM	Level Kognitif yang Diujikan
Literasi membaca	<ol style="list-style-type: none"> <li>1. Menemukan informasi (<i>access and retrieve</i>)</li> <li>2. Menafsirkan dan mengintegrasikan (<i>interpret and integrate</i>)</li> <li>3. Mengevaluasi dan merefleksi (<i>evaluate and reflect</i>)</li> </ol>
Numerasi	<ol style="list-style-type: none"> <li>1. Pemahaman (<i>knowing</i>)</li> <li>2. Penerapan (<i>applying</i>)</li> <li>3. Penalaran (<i>reasoning</i>)</li> </ol>

**c. Bentuk Soal AKM**

Sebagai suatu penilaian, selain memiliki konten domain, konteks, dan level kognitif yang beragam, AKM juga memiliki beragam bentuk soal yang diujikan. Aufa & Manoy (2022, p. 220) menyatakan bahwa bentuk soal yang dimaksud terdiri atas soal pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian. Kelima bentuk soal AKM di atas

akan dipergunakan pada penelitian ini, sehingga berikut adalah penjabaran dari masing-masing bentuk soal AKM.

#### 1) Soal pilihan ganda

Soal pilihan ganda merupakan bentuk soal yang populer pada penilaian pendidikan, baik skala nasional maupun internasional. Menurut Gierl et al. (2017, p. 1082), soal pilihan ganda merupakan bentuk penilaian yang paling efektif, dapat bertahan lama, serta bentuknya yang ekonomis. Selain itu, soal pilihan ganda mudah dalam penskoran, menawarkan peningkatan akurasi, handal, serta objektif pada proses penilaiannya (Romm et al., 2019, p. 1; Tangianu et al., 2018, p. 88). Berdasarkan konstruksinya, soal pilihan ganda berupa pokok soal (*stem*) dan beberapa opsi jawaban yang hanya ada satu jawaban benar (Butler, 2018, p. 1; Haladyna et al., 2019, p. 350). Sedangkan opsi yang bukan jawaban benar disebut dengan pengecoh (*distractor*).

Keberadaan pengecoh dalam soal pilihan ganda merupakan hal yang penting, karena Papenberg & Musch (2017) meyakini soal pilihan ganda berkualitas tinggi jika pengecohnya masuk akal dan berfungsi dengan baik. Sajjad et al. (2020, p. 982) juga menjelaskan bahwa pengecoh yang baik setidaknya 5% dipilih oleh peserta tes. Pada implementasinya di sekolah, jumlah opsi jawaban pada soal pilihan ganda disesuaikan dengan jenjang sekolah peserta didik. Hal tersebut didasari atas perbedaan kemampuan peserta disetiap jenjangnya. Jenjang SD/Sederajat kelas I hingga kelas III terdapat 3 opsi jawaban yakni A, B, dan C. Jenjang SD/Sederajat kelas IV hingga kelas VI dan SMP/Sederajat kelas VII hingga kelas IX terdapat 4 opsi jawaban yakni A, B, C, dan D. Jenjang SMA/SMK/Sederajat kelas X hingga kelas XII terdapat 5 opsi jawaban yakni A, B, C, D, dan E. Karena pada penelitian ini ditujukan untuk peserta didik kelas XI, maka jumlah opsi jawaban yang dikembangkan pada soal AKM numerasi bentuk pilihan ganda sebanyak 5 opsi.

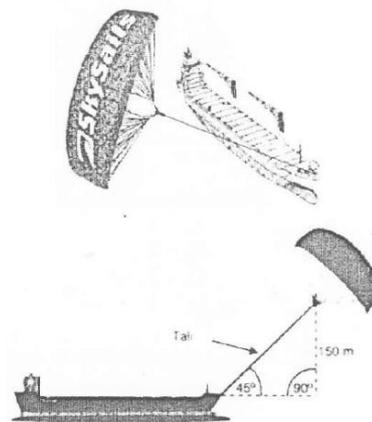
Berikut adalah salah satu contoh soal pilihan ganda yang dipergunakan pada pelaksanaan UN matematika, sebelum adanya AKM (Retnawati et al., 2017, p. 264).

Perhatikan gambar kapal layar berikut!

Sekitar 95% komoditas perdagangan dunia melalui transportasi laut dengan menggunakan sekitar 50.000 kapal tanker, kapal-kapal pengirim, dan pengangkut barang raksasa. Sebagian besar kapal-kapal ini menggunakan bahan bakar solar.

Para insinyur berencana membangun tenaga pendukung menggunakan angin untuk kapal-kapal tersebut. Ide mereka adalah dengan memasang layar berupa layang-layang ke kapal dan menggunakan tenaga angin untuk mengurangi pemakaian solar serta dampak solar terhadap lingkungan.

Berdasarkan hal di atas, berapa panjang tali layar dari layang-layang agar layar tersebut menarik kapal pada sudut  $45^\circ$  dan berada pada ketinggian vertikal  $150\text{ m}$  seperti pada gambar?

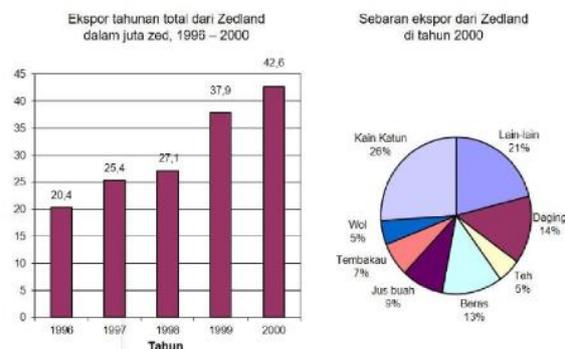


Sedangkan berikut adalah beberapa contoh bentuk soal pilihan ganda pada AKM.

**Pertama** (Pusmenjar, 2021a, p. 98)

### EKSPOR

Grafik di bawah ini memberikan informasi tentang ekspor dari Zedland, sebuah negeri yang menggunakan satuan mata uang zed.



Pertanyaan:

Berapa harga jus buah yang diekspor dari Zedland di tahun 2000?

- a. 1,8 juta zed
- b. 2,3 juta zed
- c. 3,4 juta zed
- d. 3,8 juta zed

**Kedua** (Pusmenjar, 2021a, p. 102)

Hasil dari suatu perlombaan lompat jauh di suatu sekolah dilaporkan sebagai berikut:

Rata-rata Panjang Lompatan	
Tim A	3,6 m
Tim B	4,8 m

Banyak anggota dalam setiap tim adalah sama.

Pertanyaan:

Manakah pernyataan berikut yang pasti benar tentang hasil perlombaan tersebut?

- a. Setiap peserta didik di tim B melompat lebih jauh daripada peserta didik mana pun di tim A
- b. Setelah setiap peserta didik di tim A melompat, ada salah seorang peserta didik di tim B yang melompat lebih jauh
- c. Sebagai sebuah grup, tim B melompat lebih jauh daripada tim A
- d. Beberapa peserta didik di tim A melompat lebih jauh dari beberapa peserta didik di tim B

## 2) Soal pilihan ganda kompleks

Soal pilihan ganda kompleks merupakan pengembangan dari soal pilihan ganda biasa. Butler (2018, p. 2) menegaskan bahwa penggunaan bentuk soal ini telah meningkat penggunaannya karena membutuhkan kemampuan berpikir tingkat tinggi dalam pembuatan soal dan pengerjaannya. Konstruksi soal pilihan ganda kompleks terdiri atas pokok

soal (*stem*), satu set tanggapan utama atas pokok soal, serta satu set pilihan jawaban dengan lebih dari satu jawaban yang benar (Haberkorn et al., 2016, p. 524). Sebelum adanya AKM, pelaksanaan penilaian di Indonesia yang menggunakan bentuk soal ini yakni pada tes seleksi masuk perguruan tinggi. Kemudian dengan adanya AKM bentuk soal pilihan ganda kompleks ditetapkan sebagai salah satu variasi bentuk soal yang harus ada dari jenjang SD/Sederajat hingga jenjang SMA/SMK/Sederajat. Kemudian Pusmenjar (2020, p. 5) menegaskan bahwa pada bentuk soal ini terdapat beberapa pernyataan yang harus dipilih peserta didik menggunakan tanda centang pada bagian yang telah disediakan. Selain itu Pusmenjar (2020, p. 5) membatasi variasi soal pilihan ganda kompleks yang digunakan untuk AKM yakni (1) berupa pilihan ganda dengan lebih dari satu jawaban benar; dan (2) berupa pilihan ganda kompleks benar/salah, ya/tidak, atau kategori lainnya.

Jumlah pernyataan yang digunakan dalam bentuk soal pilihan ganda kompleks dapat dilihat pada tabel 4 (Pusmenjar, 2021a, p. 106).

**Tabel 4.** *Jumlah Pernyataan Soal Pilihan Ganda Kompleks*

<b>Pilihan Ganda Lebih dari Satu Jawaban Benar</b>		<b>Pilihan Ganda Kompleks Benar/Salah, Ya/Tidak, atau Lainnya</b>	
<b>Kelas</b>	<b>Jumlah pernyataan</b>	<b>Kelas</b>	<b>Jumlah pernyataan</b>
I – VI	3 pilihan	I – VI	2 pilihan
VII – X	4 pilihan	VII – X	3 pilihan
XI – XII	5 pilihan	XI – XII	4 pilihan

Berikut adalah salah satu contoh bentuk soal pilihan ganda kompleks yang dipergunakan pada seleksi masuk perguruan tinggi negeri di Indonesia (Eduka, 2020, p. 48).

Diketahui  $S = \{6, 7, 7, 7, 11, 13, 13, 24\}$ . Mana dari pernyataan di bawah ini yang nilainya juga merupakan anggota dari  $S$ ?

- (1) Median
  - (2) Rata-rata
  - (3) Jangkauan
  - (4) Modus
- a. (1), (2), dan (3) SAJA yang benar
  - b. (1) dan (3) SAJA yang benar
  - c. (2) dan (4) SAJA yang benar
  - d. HANYA (4) yang benar
  - e. SEMUA pilihan benar

Sedangkan berikut adalah beberapa contoh soal pilihan ganda kompleks pada pelaksanaan AKM.

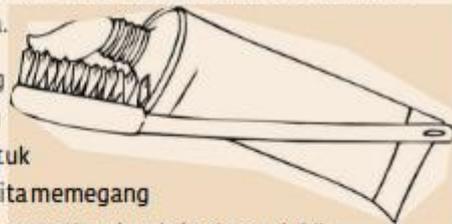
**Pertama** (Pusmenjar, 2021a, p. 47)

### Menggosok Gigi

Apakah dengan menggosok gigi semakin lama dan semakin keras gigi kita akan semakin bersih?

Peneliti dari Inggris menjawab tidak. Mereka sudah mencoba berbagai alternatif, dan akhirnya menemukan cara yang sempurna untuk menggosok gigi. Cukup menggosok gigi selama 2 menit, tanpa harus menggosok dengan keras, akan memberikan hasil terbaik. Menggosok terlalu keras akan membahayakan email gigi dan gusi kita tanpa melepaskan sisa makanan dan plak yang menempel di gigi kita.

Bente Hansen, seorang pakar di bidang gigi, mengatakan bahwa cara yang paling baik untuk memegang sikat gigi adalah seperti kita memegang pulpen. "Dimulai dari satu sudut dan gosok seluruh barisan gigi," Jangan lupa menggosok lidah! Pada lidah biasanya terkandung banyak bakteri yang dapat menyebabkan bau mulut.



"Menggosok Gigi" adalah artikel yang diambil dari majalah Norwegia.

Pertanyaan:

Menurut Bente Hansen, mengapa kita harus menggosok lidah kita?

Pilihlah jawaban yang benar berikut dengan memberi tanda centang (✓).

Jawaban benar lebih dari satu.

- Untuk menghilangkan bakteri
- Dapat menghilangkan bau mulut kita
- Melindungi email gigi kita
- Menjaga bakteri baik di lidah kita

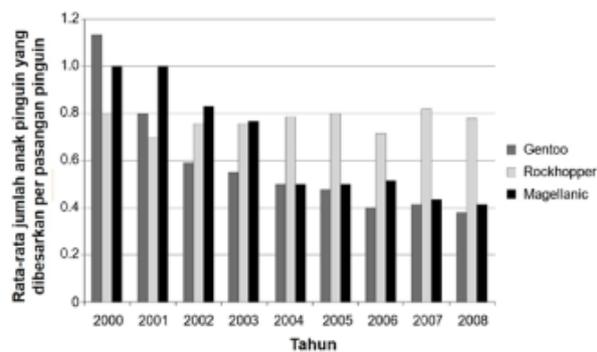
**Kedua** (Pusmenjar, 2021a, p. 100)

### PENGUIN

Setelah pulang dari perjalanan wisata, Jean mencari informasi di internet tentang berapa rata-rata anak penguin yang dibesarkan pasangan penguin.

Dia menemukan diagram batang berikut untuk tiga jenis penguin yaitu *Gentoo*, *Rockhopper* dan *Magellanic*

Jumlah Anak Penguin yang Dibesarkan per Pasangan Penguin Setiap Tahun



Pertanyaan:

Berdasarkan diagram di atas, apakah pernyataan berikut benar atau salah?

Beri tanda centang (✓) pada kolom Benar atau Salah untuk setiap pernyataan!

Pernyataan	Benar	Salah
Pada tahun 2000, rata-rata jumlah anal penguin yang dibesarkan setiap pasangan penguin lebih besar dari 0,6		
Pada tahun 2006, secara rata-rata kurang dari 80% pasangan penguin membesarkan seorang anak		
Rata-rata jumlah anak penguin <i>Magellanic</i> yang dibesarkan per pasangan penguin menurun antara tahun 2001 dan 2004		

### 3) Soal menjodohkan

Soal menjodohkan bermakna mencocokkan antara dua hal yang sedang dibandingkan (Lee, 2013, pp. 47–48). Bentuk soal ini umum dipergunakan dalam penilaian serta tidak memiliki batasan jenjang untuk penggunaannya. Pusmenjar (2020, p. 5) menjelaskan bahwa pada soal menjodohkan terdapat proses menyesuaikan dan menghubungkan antar pernyataan. Sehingga pada bentuk soal ini terdapat dua lajur yakni (1) lajur kiri berupa pokok soal atau premis; dan (2) lajur kanan berupa jawaban atau respon. Pada pelaksanaan AKM, Pusmenjar (2021a, p. 106) menetapkan bahwa pokok soal pada jalur kiri harus lebih banyak daripada jumlah jawaban pada jalur kanan. Untuk ketentuan jumlah pernyataan, jenjang SD/Sederajat kelas I hingga kelas VI sebanyak 3 pilihan pernyataan, jenjang SMP/Sederajat kelas VII hingga kelas IX dan jenjang SMA/SMK/Sederajat kelas X sebanyak 4 pilihan pernyataan, serta kelas XI hingga kelas XII sebanyak 5 pilihan pernyataan.

Berikut adalah salah satu contoh bentuk soal menjodohkan yang diujikan kepada mahasiswa (Draaijer et al., 2018, p. 216).

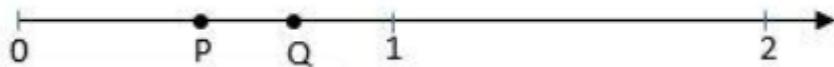
Cocokkan jenis-jenis pertanyaan pada kuis di sebelah kanan dengan deskripsi yang benar di sebelah kiri!

Deskripsi	
Peserta didik harus membuat asosiasi antar item soal pada dua daftar	1
Peserta didik menilai kebenaran secara deklaratif proposisi	2
Peserta didik memilih satu jawaban benar dari satu set pilihan jawaban	3

Jenis-jenis pertanyaan	
a	Esai
b	Menjodohkan
c	Pilihan majemuk
d	Benar – salah

Sedangkan berikut adalah beberapa contoh soal menjodohkan pada pelaksanaan AKM.

**Pertama** (Pusmenjar, 2020, p. 102)



P dan Q merepresentasikan bilangan pecahan pada garis bilangan di atas.

Pertanyaan:

Pasangkan P dan Q dengan bilai pecahan yang mungkin.

P	$\frac{5}{3}$
Q	$\frac{1}{2}$
	$\frac{3}{4}$

**Kedua** (Pusmenjar, 2021a, p. 56)

Level 5

### Hijau Kampungku di Tengah Kota: Aku dan Belimbing Wuluh

"Aku tadi menawarkan bibit pohon belimbing kepada beberapa tetangga. Kalau mereka ikut menanam pohon, jalan kampung ini akan menjadi lebih teduh. Tidak seperti sekarang, ada bagian yang teduh, ada bagian yang panas," keluhku. "Ada yang beralasan, air sedang susah didapat. Ada yang berkata, 'Nanti, ya, tunggu musim hujan datang lagi.' Aku kecewa, Yah."

"Hmm... Mereka tidak ingin menggunakan air terlalu banyak. Saat ini, memang sebaiknya kita hemat air," kata Ayah. Tentu saja, aku semakin merengut. Kalau air tetap sukar didapat, tidak ada orang yang mau menanam pohon belimbing wuluh itu.

"Tapi, jangan khawatir. Masalah air untuk menyiram tanaman sebentar lagi akan terbantu oleh proyek Pak RT," kata Ayah. "Kampung kita akan punya Pandora L."

"Pandora L? Apa itu, Yah?"

Maket Pandora L (Pendaur Ulang Limbah Air)

"Ini maket dari pengolahan limbah yang baru saja selesai dibangun di kampung kita. Bangunan ini ada di dalam tanah dan berguna untuk mengolah limbah rumah tangga saja, seperti air cucian," kata Ayah.

Air hasil pengolahan dari Pandora L digunakan untuk kegiatan menyiram tanaman dan mencuci kendaraan. Pemasangan Pandora L terletak di Kampung Genteng Candirejo di tengah kota Surabaya, Ibu kota Provinsi Jawa Timur.

(Diadaptasi dari *Hijau Kampungku di Tengah Kota: Aku dan Belimbing Wuluh* karya Tyas KW)

Pertanyaan:

Pasangkanlah sesuatu yang disebutkan oleh tokoh dalam cerita tersebut dengan benda yang dimaksud!

Pernyataan	
Hal yang ditawarkan oleh tokoh Aku kepada tetangganya	1
Hal yang akan segera dimiliki oleh warga kampung	2
Hal yang disarankan oleh tokoh Ayah untuk dihemat	3

Benda	
a	Pandora L
b	Air
c	Pohon belimbing wuluh
d	Bibit pohon belimbing

#### 4) Soal isian singkat atau jawaban singkat

Soal isian singkat atau jawaban singkat merupakan bentuk soal yang pokok soalnya terdapat elemen yang sengaja dikosongkan dan tidak ada pilihan jawaban yang diberikan (Cardona-Zea & Rojas-Galeano, 2012, p. 1). Sehingga peserta didik dituntut menemukan jawaban yang singkat baik berupa kata, frasa, bilangan, atau simbol. Kemudian letak perbedaan soal isian singkat dengan jawaban singkat yakni pada soal isian singkat, pokok soal berupa kalimat yang tidak lengkap sedangkan soal jawaban singkat, pokok soal berupa kalimat tanya (Pusmenjar, 2021a, p. 107). Goshima et al. (2019, p. 908) menjelaskan bahwa peserta tes memiliki peluang yang sama dalam menjawab soal isian singkat atau jawaban singkat berdasarkan pemahaman masing-masing peserta tes. Sama halnya dengan bentuk soal menjodohkan, bentuk soal ini juga umum dipergunakan pada berbagai penilaian.

Berikut adalah salah satu contoh bentuk soal isian singkat yang diujikan kepada peserta didik SMP (Charmila et al., 2016, p. 205).

## Gunung Kerinci

Gunung Kerinci merupakan gunung api tertinggi di Indonesia dengan ketinggian 3.805 mdpl (meter di atas permukaan laut). Gunung ini termasuk ke dalam kawasan Taman Nasional Kerinci Seblat dan pegunungan Bukit Barisan yang ada pada Kabupaten Kerinci, Jambi.

Menjadi salah satu gunung yang digemari para pendaki, tentu membuat gunung Kerinci memiliki beragam jalur pendakian, bahkan pada salah satu jalur terdapat tanjakan dengan kemiringan  $60^\circ$ . Jalur umum yang sering dilewati para pendaki untuk sampai ke puncak bernama jalur Kersik Tuo.

Berikut merupakan jarak dan rata-rata waktu pendakian satu tim pendaki gunung Kerinci melalui jalur Kersik Tuo.



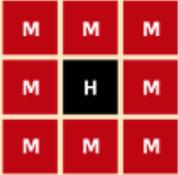
Perkiraan kecepatan rata-rata para pendaki mulai dari pos 3 hingga puncak Indrapura adalah ... km/jam.

Sedangkan berikut adalah beberapa contoh soal isian singkat atau jawaban singkat pada pelaksanaan AKM.

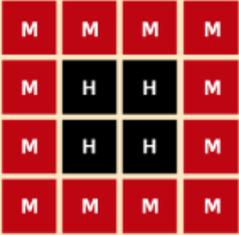
**Pertama** (Pusmenjar, 2021a, p. 94)

Pat memiliki kepingan berbentuk persegi yang berwarna merah dan hitam. Pat menggunakan kepingan-kepingan tersebut untuk membentuk persegi yang lebih besar.

Persegi ukuran 3 x 3 memiliki 1 keping hitam dan 8 keping merah.



Persegi ukuran 4 x 4 memiliki 4 keping hitam dan 12 keping merah.



 Merah  
 Hitam

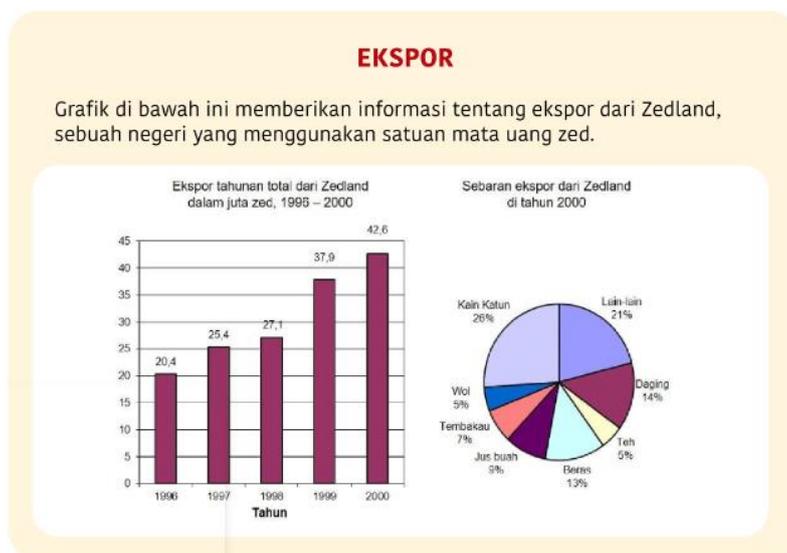
Pat melanjutkan untuk membentuk persegi ukuran  $4 \times 4$ ,  $5 \times 5$ ,  $6 \times 6$ , dan seterusnya menggunakan kepingan-kepingan tersebut dengan pola yang sama seperti pola pada kedua persegi di atas.

Pertanyaan:

Persegi yang berhasil dibentuk oleh Pat memiliki 49 kepingan hitam. Berapa banyak kepingan merah yang digunakan Pat untuk membentuk persegi tersebut?

... kepingan merah

**Kedua** (Pusmenjar, 2021a, pp. 98–99)



Pertanyaan:

Berapakah nilai ekspor total dari Zedland (dalam juta zed) di tahun 1998?

Jawaban: .....

5) Soal uraian

Soal uraian merupakan bentuk soal yang menuntut peserta didik berpikir logis, bernalar kritis, dan berpikir kreatif dalam mengorganisasikan gagasan dalam bentuk uraian tertulis (Uto & Okano, 2021, p. 763). Sedangkan Winarni et al. (2022, p. 85) menjelaskan bahwa konstruk soal uraian terdiri atas pertanyaan terstruktur. Peserta tes memiliki kebebasan dalam menjawab sesuai dengan pemahaman dan kebahasaan masing-masing. Maka dari itu akan diperoleh beragam tanggapan atau jawaban dari peserta tes yang menjadi ciri identik bentuk soal ini. Bentuk soal uraian tergolong umum dipergunakan dalam segala jenis penilaian, terutama pada penilaian harian.

Ditetapkan sebagai salah satu bentuk soal AKM, soal uraian memiliki pedoman penskoran yang bergantung pada kompleksitas jawaban peserta tes. Pusmenjar (2021a, p. 108) menetapkan penskorannya berupa data politomus, yakni skor 2 untuk jawaban yang memenuhi semua kriteria kunci jawaban, skor 1 untuk jawaban yang kurang memenuhi kriteria kunci jawaban, dan skor 0 untuk jawaban salah. Berikut adalah salah satu contoh bentuk soal uraian yang diujikan kepada peserta didik SMA (Apino & Retnawati, 2017, p. 104).

Seorang pengusaha mebel membutuhkan biaya untuk memproduksi beberapa unit lemari selama satu minggu sebesar Rp37.500.000,00. Hasil produksi lemari selama seminggu tersebut dipasarkan dan berhasil terjual dengan sisa 3 unit saja. Adapun total hasil penjualan lemari-lemari tersebut Rp36.000.000,00 dengan keuntungan tiap lemari Rp500.000,00. Berdasarkan ilustrasi tersebut, buatlah persamaan untuk menentukan jumlah lemari yang diproduksi selama seminggu.

Sedangkan berikut adalah beberapa contoh soal isian singkat atau jawaban singkat pada pelaksanaan AKM.

**Pertama** (Pusmenjar, 2021a, p. 38)

**Level 2**

### KUE TRADISIONAL INDONESIA

Kelapa, gula nira, dan aneka beras sering digunakan untuk membuat kue tradisional. Oleh karena itu, banyak kue Indonesia yang memiliki rasa sama meskipun namanya berbeda. Contohnya adalah kue klepon yang berbentuk bola, terbuat dari tepung beras ketan dan gula merah. Di Jawa, kue tersebut disebut klepon, sementara di Sumatra dan Sulawesi disebut onde-onde. Kerupuk yang terbuat dari beras kering dan dilumuri gula nira cair disebut intip di Solo, sementara di Minang disebut batiah.

Walaupun terbuat dari bahan yang sama, cara masak yang berbeda akan menghasilkan rasa yang berbeda pula. Contoh kue-kue itu adalah serabi dan putu. Kedua kue terbuat dari campuran tepung beras, santan kelapa, dan gula merah. Bedanya, serabi dibuat dengan cara dipanggang, sementara putu dikukus. Bentuk yang dihasilkan kedua kue tersebut juga berbeda.

MEMASAK KUE PUTU

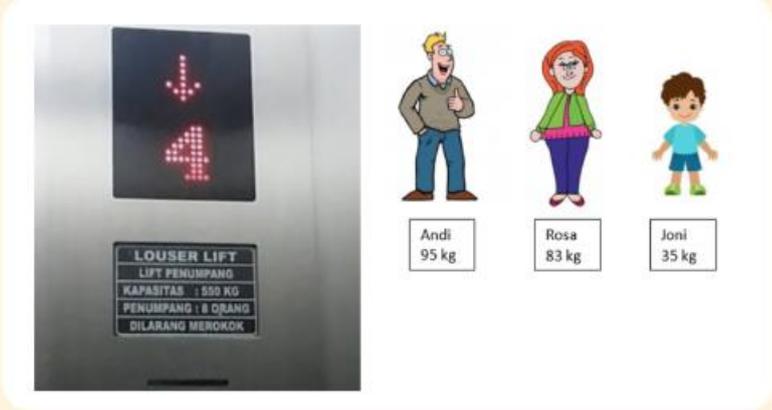
MEMASAK SERABI

Pertanyaan:

Serabi dan putu dibuat dari bahan yang sama. Mengapa rasanya berbeda? Uraikan pendapatmu!

**Kedua** (Pusmenjar, 2021a, p. 93)

**Perhatikan keterangan pada gambar lift di bawah ini!**



Di luar lift, Andi, Rosa, dan Joni akan masuk ke dalam lift yang sama.

Lift sudah berisi 5 orang dengan jumlah berat badan 380 *kg*. Di luar lift Andi, Rosa, dan Joni akan masuk ke dalam lift yang sama.

Pertanyaan:

Jika Rosa dan Joni masuk bersamaan, kemudian bergabung dengan kelima orang di dalam lift, berapa kilogram barang paling berat yang dapat mereka bawa? Jelaskan alasanmu secara matematis!

**d. Capaian Hasil AKM**

Pada pelaksanaannya AKM tidak menggunakan standar capaian hasil dengan skala 0 – 100 namun dilaporkan dengan rentang nilai 1,00 – 3,00. Sistem penilaian nasional ini memiliki karakteristik tersendiri dalam menilai dan mendeskripsikan hasil AKM. Hal tersebut sesuai dengan tujuan pelaksanaan AKM dimana hasil AKM dirancang untuk memberikan informasi mengenai tingkat kompetensi peserta didik yang mana akan dimanfaatkan pendidik untuk menyusun strategi pembelajaran yang efektif dan berkualitas. Karena pembelajaran yang dirancang dengan memperhatikan tingkat capaian peserta didik dapat memudahkan peserta didik dalam menguasai konten atau kompetensi yang diharapkan dalam suatu mata pelajaran. Berikut adalah pelaporan hasil AKM sesuai standar pemerintah, yang kemudian dalam penelitian ini digunakan dalam

mendeskripsikan profil peserta didik dalam menyelesaikan soal AKM numerasi. Capaian hasil AKM terbagi menjadi dua yakni atribut sekolah dan atribut peserta didik. Atribut sekolah menunjukkan capaian hasil sekolah secara keseluruhan sedangkan atribut peserta didik menunjukkan capaian hasil peserta didik secara individu (Kemendikbudristek, 2022, p. 5).

1) Atribut sekolah

Atribut sekolah mencakup empat tingkat penilaian, yakni:

a) Di atas kompetensi minimum

Peserta didik di sekolah menunjukkan tingkat numerasi yang cakap dan cukup banyak peserta didik berada pada level mahir. Rentang nilai yang dicapai sebesar 2,10 s.d. 3,00.

b) Mencapai kompetensi minimum

Sebagian besar peserta didik telah mencapai batas kompetensi minimum untuk numerasi namun perlu upaya mendorong lebih banyak peserta didik menjadi mahir. Rentang nilai yang dicapai sebesar 1,80 s.d. 2,09.

c) Di bawah kompetensi minimum

Kurang dari 50% peserta didik telah mencapai kompetensi minimum untuk numerasi. Rentang nilai yang dicapai sebesar 1,40 s.d. 1,79.

d) Jauh di bawah kompetensi minimum

Sebagian besar peserta didik belum mencapai batas kompetensi minimum untuk numerasi. Rentang nilai yang dicapai sebesar 1,00 s.d. 1,39.

2) Atribut peserta didik

Atribut peserta didik mencakup empat tingkat penilaian, yakni:

a) Tingkat mahir

Peserta didik mampu bernalar untuk menyelesaikan masalah kompleks serta nonrutin berdasarkan konsep matematika yang dimilikinya.

b) Tingkat cakap

Peserta didik mampu mengaplikasikan pengetahuan matematika yang dimiliki dalam konteks yang lebih beragam.

c) Tingkat dasar

Peserta didik memiliki keterampilan dasar matematika: komputasi dasar dalam bentuk persamaan langsung, konsep dasar terkait geometri dan statistika, serta menyelesaikan masalah matematika sederhana yang rutin.

d) Tingkat perlu intervensi khusus

Peserta didik hanya memiliki pengetahuan matematika yang terbatas (penguasaan konsep yang parsial dan keterampilan komputasi yang terbatas).

### 3. Numerasi

#### a. Konsep Dasar Numerasi

Numerasi merupakan hal yang populer di dunia pendidikan baik secara nasional maupun global. Hal tersebut terjadi karena numerasi masih berkaitan dengan matematika, namun keduanya bukan hal yang sama. Westwood (2008, pp. 1–2) menyatakan bahwa istilah numerasi secara resmi digagas kemudian dilaporkan oleh Dewan Penasehat Pusat untuk Pendidikan di Inggris tahun 1959 sebagai kemampuan yang berkaitan dengan perhitungan kuantitatif dalam kehidupan sehari-hari. Disamping itu Balt et al. (2019, p. 200) juga mendefinisikan numerasi sebagai kemampuan dalam menerapkan konsep dasar numerik. Kedua definisi tersebut menegaskan bahwa objek numerasi sebagai kemampuan adalah pengetahuan kuantitatif atau numerik atau bisa disebut dengan matematika.

Secara global numerasi mendapat perhatian lebih, terutama dalam dunia pendidikan setelah OECD menggagas pelaksanaan PISA pada tahun 1997 dengan salah satu bidang asesmennya adalah numerasi atau lebih dikenal dengan literasi matematika. Sejak periode pertama pelaksanaannya (tahun 2000) PISA memberikan definisi tersendiri

terhadap literasi matematika di setiap periodenya. Tahun 2000 PISA mendefinisikan literasi matematika sebagai kemampuan individu untuk mengidentifikasi, memahami, serta terlibat peran dalam matematika pada kehidupan sehari-hari baik kehidupan pribadi, lingkungan pekerjaan, dan kehidupan sosial (OECD, 2000, p. 5). Selanjutnya untuk periode tahun 2003, 2006, dan 2009 PISA mendefinisikan literasi matematika sebagai kemampuan individu untuk mengidentifikasi dan memahami peran matematika yang ada pada dunia, serta melibatkan matematika dalam memenuhi kebutuhan hidup (OECD, 2009, p. 84).

Selanjutnya periode tahun 2012, PISA mendefinisikan literasi matematika sebagai kemampuan individu untuk memformulasikan, menggunakan, serta menginterpretasikan matematika kedalam berbagai konteks. Termasuk juga penalaran matematis dengan menggunakan konsep, prosedur, fakta, dan alat matematika dalam mendeskripsikan, menjelaskan, dan memprediksi fenomena. Hal tersebut membantu individu dalam mengenali peran matematika pada dunia (OECD, 2013, p. 7). Ketiga definisi di atas menegaskan bahwa terdapat dua karakter kunci literasi matematika berdasarkan PISA yakni (1) kemampuan dalam menggunakan matematika dalam kehidupan sehari-hari; dan (2) kemampuan dalam memahami dan mengapresiasi informasi yang tersajikan dalam istilah matematika.

Melihat hasil partisipasi Indonesia dalam pelaksanaan PISA di setiap tahunnya yang tergolong sangat rendah, pemerintah di bawah naungan Kemendikbud pada tahun 2017 menggagas sebuah gerakan literasi nasional yang selanjutnya dikenal dengan sebutan GLN. Pada GLN numerasi didefinisikan sebagai pengetahuan serta kecakapan untuk (1) memperoleh, menginterpretasi, menggunakan, dan mengkomunikasikan angka-angka dan simbol matematika dalam menyelesaikan masalah praktis dengan berbagai konteks kehidupan sehari-hari; serta (2) menganalisis informasi yang ditampilkan ke dalam grafik, tabel, bagan, dll guna mengambil sebuah keputusan (Atmazaki et al., 2017, p. 7). Definisi

tersebut sejalan dengan PIAAC Numeracy Expert Group (2009, p. 6) yang mendefinisikan numerasi sebagai kemampuan untuk mengakses, menggunakan, menginterpretasi, dan mengkomunikasikan ide dan informasi matematika untuk dilibatkan dalam pengelolaan matematika kedalam berbagai konteks kehidupan sehari-hari.

Lingkungan sekolah merupakan salah satu fokus implementasi GLN, sehingga tentu seluruh sekolah berupaya dengan baik dalam melatih dimensi-dimensi literasi pada GLN, dengan salah satunya adalah numerasi. Getenet (2022, p. 664) menjelaskan bahwa peserta didik bernumerasi ketika mereka mengembangkan pengetahuan serta kemampuannya untuk menggunakan matematika dengan percaya diri di seluruh mata pelajaran (selain matematika) sekolah serta dalam kehidupan mereka secara lebih luas. Kemudian Knowles et al. (2021, p. 243) menyatakan bahwa memiliki pengetahuan, kemampuan, dan kepercayaan diri dalam menggunakan alat matematika dalam berbagai disiplin ilmu adalah definisi dari numerasi. Definisi dari kedua ahli tersebut semakin mempertajam definisi numerasi, bahwa seseorang bisa bernumerasi dari segala disiplin ilmu dengan berbagai konteks.

Kemudian tahun 2021 Kemdikbudristek melaksanakan AKM sebagai terobosan baru yang salah satu ruang lingkup yang diujikan adalah numerasi. Numerasi dalam pelaksanaan AKM didefinisikan sebagai kemampuan berpikir dalam menggunakan konsep, prosedur, fakta, serta alat matematika yang digunakan dalam memecahkan masalah sehari-hari pada berbagai konteks yang relevan untuk individu sebagai warga Indonesia maupun global (Pusmenjar, 2021a, pp. 65–66). Menurut Nortvedt & Wiese (2020, p. 529), hanya mampu menerapkan pengetahuan dan keterampilan matematika saja tidak cukup, namun dengan mewajibkan belajar dapat mengindikasikan adanya kontribusi besar dalam numerasi peserta didik dengan fokus pada matematika terapan hingga matematika dalam kehidupan sehari-hari.

Berikut adalah beberapa contoh soal numerasi yang diujikan pada berbagai penilaian.

**Pertama**, contoh numerasi berdasarkan LANTITE atau *Literacy and Numeracy Test for Initial Teacher Education Students* di Australia (Hall & Zmood, 2019, p. 3).

Numeracy Sample Question 10

GEOGRAPHICAL DISTRIBUTION OF AUSTRALIANS

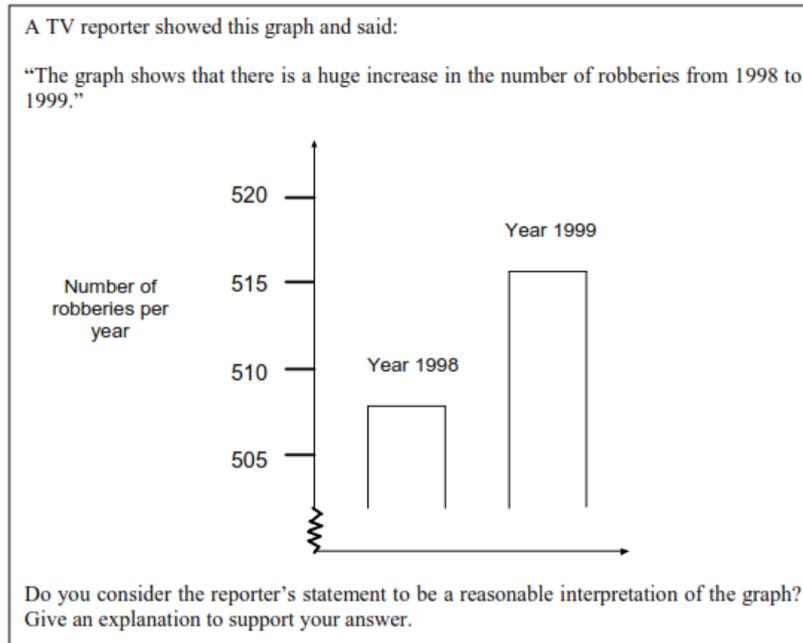
The Australian Bureau of Statistics conducts a census every five years.  
In 2011, the population of Australia was 22 million.  
About 2% of these people lived in remote or very remote areas.

About how many people lived in remote or very remote areas in Australia in 2011?

**A** 11 000  
**B** 44 000  
**C** 110 000  
**D** 440 000

Tanpa menggunakan bantuan alat hitung, soal di atas meminta peserta didik menghitung banyaknya manusia yang tinggal di daerah pedalaman atau sangat pedalaman di Australia tahun 2011 dengan 2% dari 22 juta populasi. Jawaban yang benar adalah 440.000 (opsi D), untuk opsi B 44.000 sebagai pengecoh dengan asumsi peserta didik mengalami kesalahan hitung atau kurang teliti, sedangkan opsi A dan C sebagai pengecoh dengan asumsi persen yang dimaksudkan adalah  $\frac{1}{2}\%$  bukan 2%. Dalam menyelesaikan soal tersebut peserta didik harus menerjemahkan 22 juta populasi menjadi angka 22.000.000, dan selanjutnya menghitung 2% dari 22.000.000 atau  $\frac{2}{100} \times 22.000.000 = 440.000$ .

**Kedua**, contoh numerasi atau literasi matematika pada PISA (Gal & Tout, 2014, p. 40).



Soal numerasi di atas menuntut peserta didik memberikan argumen terkait kelogisan reporter dalam menyatakan bahwa dari tahun 1998 hingga 1999 terdapat peningkatan yang cukup besar dalam perampokan yang didukung dengan grafik. Karena ini sebuah argumen tentu antar peserta didik akan memiliki narasi yang berbeda-beda. Namun jika peserta didik mampu memahami grafik dengan baik maka keragaman narasi tersebut akan memiliki inti argumen yang sama, dimana peningkatan perampokan dari tahun 1998 hingga 1999 tidak cukup besar, karena di tahun 1998 sebanyak 508 kasus, 1999 sebanyak 515 kasus, peningkatan hanya sebesar 7 kasus.

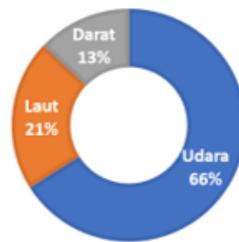
**Ketiga**, contoh numerasi yang diharapkan pada pelaksanaan AKM (Komariah et al., 2020, p. 156)

#### Potensi Sektor Wisata di Indonesia

Pariwisata merupakan salah satu program unggulan pemerintah Indonesia untuk mendapatkan sumber pendapatan devisa negara. Sektor pariwisata ini dipercaya terus mencatatkan pertumbuhan yang positif. Pada tahun 2018 pertumbuhan sektor pariwisata Indonesia, yaitu 12,58% lebih tinggi dibandingkan rata-rata pertumbuhan sektor pariwisata dunia yang hanya mencapai 5,6% serta

ASEAN sebesar 7,4%. Pertumbuhan sektor pariwisata ini tentu tak lepas dari kenaikan jumlah wisatawan asing yang berkunjung ke Indonesia dari tahun ke tahun.

Sarana transportasi kunjungan wisatawan mancanegara ke Indonesia masih didominasi lewat udara. Hal ini karena letak Indonesia yang berada di kepulauan, akses mudah dan cepat, dan beberapa tempat wisata yang menjadi destinasi favorit berada di banyak daerah. Melimpahnya potensi wisata alam yang dimiliki Indonesia, membuka peluang untuk terus meningkatkan jumlah kunjungan wisatawan. Salah satu upaya yang dilakukan pemerintah melalui Kementerian Pariwisata adalah menetapkan 10 destinasi wisata prioritas. Kesepuluh destinasi ini adalah Danau Toba di Sumatera Utara, Tanjung Kelayang di Kepulauan Bangka Belitung, Tanjung Lesung di Banten, Kepulauan Seribu di Jakarta, Borobudur di Jawa Tengah, Bromo Tengger Semeru di Jawa Timur, Mandalika di Nusa Tenggara Barat, Labuan Bajo di Nusa Tenggara Timur, Wakatobi di Sulawesi Tenggara, dan Morotai di Maluku Utara. Data jumlah wisatawan bulan Juli 2019 ke Indonesia, yaitu ada 1.547.231 orang. Jika dibedakan berdasarkan cara masuk, hasilnya sebagai berikut.



Jodohkanlah antara pernyataan dan jawaban yang tersedia.

1. Jumlah wisatawan dari Bandara Ngurah Rai	<input type="checkbox"/>	A. 40.847 orang
2. Jumlah wisatawan dari Bandara Soekarno Hatta	<input type="checkbox"/>	B. 83.551 orang
3. Selisih banyaknya wisatawan dari Bandara Soekarno Hatta dan Sultan Hasanuddin	<input type="checkbox"/>	C. 102.117 orang
		D. 367.622 orang
		E. 408.469 orang

Soal numerasi di atas menuntut kejelian dan pemahaman peserta didik dalam membaca informasi terutama data yang telah disediakan, karena untuk menyelesaikan bentuk soal menjodohkan tersebut data memiliki peran penting. Telah diketahui bahwa jumlah wisatawan bulan Juli 2019 sebanyak 1.547.231 orang. Kemudian wisatawan yang datang melalui jalur udara sebanyak  $\frac{66}{100} \times 1.547.231 = 1.021.172,46$  yang

dilakukan pembulatan ke bawah karena ini berkaitan dengan kuantitas manusia sehingga sebanyak 1.021.172 orang.

Pertanyaan pertama meminta peserta didik untuk menghitung jumlah wisatawan melalui Bandara Ngurah Rai dimana  $\frac{10}{100} \times 1.021.172 = 102.117,2 \approx 102.117$  orang. Pertanyaan kedua menghitung jumlah wisatawan melalui Bandara Soekarno Hatta dimana  $\frac{40}{100} \times 1.021.172 = 408.468,8 \approx 408.469$  orang. Pertanyaan ketiga menghitung selisih wisatawan melalui Bandara Soekarno Hatta dan Sultan Hasanuddin, namun sebelum itu harus dihitung terlebih dahulu jumlah wisatawan melalui Bandara Sultan Hasanuddin dimana  $\frac{4}{100} \times 1.021.172 = 40.846,88 \approx 40.847$  orang. Sehingga selisih wisatawan Bandara Soekarno Hatta dan Sultan Hasanuddin sebanyak  $408.469 - 40.847 = 367.622$  orang. Sehingga opsi yang benar untuk pertanyaan pertama adalah C, pertanyaan kedua adalah E, dan pertanyaan ketiga adalah D.

Berdasarkan beberapa definisi yang dinyatakan oleh beberapa ahli di atas, selanjutnya dapat disimpulkan bahwa numerasi merupakan kemampuan berpikir yang didalamnya terdapat proses memformulasikan, menggunakan, menginterpretasi, serta mengkomunikasikan berbagai konsep, prosedur, fakta, dan alat matematika dalam menyelesaikan masalah dalam berbagai konteks sehari-hari.

#### **b. Komponen Konten Pada Numerasi**

Numerasi sebagai suatu kemampuan yang diujikan tentu memiliki domain ukur atau konten yang perlu diketahui berbagai pihak, dalam hal ini adalah pendidik dan peserta didik. Selain itu, numerasi yang menjadi subjek pengujian baik skala nasional atau global tentu memiliki beberapa perspektif terkait konten yang diukur. Berikut adalah konten domain numerasi dari perspektif LANTITE, PISA, hingga AKM.

LANTITE atau *Literacy and Numeracy Test for Initial Teacher Education Students* merupakan pengujian literasi dan numerasi di Australia yang ditujukan kepada para lulusan sarjana atau pascasarjana pendidikan guru sebagai syarat mengajar. Konten domain numerasi yang diujikan pada LANTITE ini sebanyak 3 cakupan yakni (1) bilangan dan aljabar; (2) pengukuran dan geometri; serta (3) statistika dan peluang (Thai et al., 2021, p. 74).

1) Bilangan dan aljabar

Proporsi konten domain bilangan dan aljabar dalam pelaksanaan LANTITE sebanyak 40 – 50% dari total soal, dengan topik yang diuji berupa (1) aritmatika dasar; (2) pembagian; (3) desimal; (4) persentase; (5) rasio; (6) aljabar; dan (7) matematika keuangan (Hall & Zmood, 2019, p. 2; Thai et al., 2021, p. 79).

2) Pengukuran dan geometri

Proporsi konten domain pengukuran dan geometri dalam pelaksanaan LANTITE sebanyak 20 – 30% dari total soal, dengan topik yang diuji berupa (1) memperkirakan, mengkonversi, dan membaca; (2) waktu dan penjadwalan; (3) jarak dan keliling; (4) luas; (5) volume; (6) sudut; serta (7) ruang, bentuk dan kesimetrisan (Hall & Zmood, 2019, p. 2; Thai et al., 2021, p. 79).

3) Statistika dan peluang

Proporsi konten domain statistika dan peluang dalam pelaksanaan LANTITE sebanyak 25 – 35% dari total soal, dengan topik yang diuji berupa (1) interpretasi data; (2) statistika; (3) kombinasi; dan (4) peluang (Hall & Zmood, 2019, p. 2; Thai et al., 2021, p. 79).

PISA atau *Programme for International Student Assessment* merupakan penilaian pengetahuan dan kemampuan skala internasional untuk peserta didik berumur 15 tahun. Tiga kompetensi yang dinilai pada pelaksanaan PISA yakni literasi membaca, literasi matematika, dan literasi sains (Rahmawati et al., 2021, p. 1). Konten domain literasi matematika

yang diujikan sebanyak 4 cakupan yakni (1) kuantitas; (2) ruang dan bentuk; (3) perubahan dan hubungan; serta (4) data dan ketidakpastian (Ozkale & Ozdemir Erdogan, 2020, p. 3).

1) Kuantitas

Kuantitas merupakan aspek matematis yang paling meresap dan esensial untuk terlibat dan berfungsi di dunia sehingga membutuhkan konsep pemahaman dalam pengukuran, hitungan, besaran, hingga pola bilangan. Kuantitas berkaitan erat dengan seluruh konten PISA yang lainnya karena memungkinkan adanya pemodelan antara masing-masing konten PISA tersebut (OECD, 2019, p. 85). Dengan demikian konten domain kuantitas menerapkan segala pengetahuan bilangan, serta operasi bilangan dalam berbagai pengaturan.

2) Ruang dan bentuk

Ruang dan bentuk mencakup berbagai fenomena visual yang melibatkan pola, sifat objek, posisi dan orientasi, representasi objek, kode informasi visual, navigasi, serta interaksi dinamis mengenai bentuk benda. Konten ruang dan bentuk pada literasi matematika ini dapat diterapkan pada kegiatan memahami perspektif, mengkreasikan dan membaca peta, menafsirkan bangun tiga dimensi, serta merepresentasikan bentuk benda (OECD, 2019, p. 85).

3) Perubahan dan hubungan

Konten domain perubahan dan hubungan melibatkan pemahaman serta model matematika untuk menggambarkan serta memprediksikan perubahan dan hubungan sesuai dengan fungsi dan persamaan. Sehingga pada konten ini mencakup ekspresi aljabar, persamaan dan pertidaksamaan, serta representasi tabel dan grafik. Penerapan konten ini ada pada beberapa konteks masalah sehari-hari seperti pertumbuhan organisme, musik, siklus musiman, pola cuaca, tingkat pekerjaan, serta kondisi ekonomi (OECD, 2019, p. 84).

#### 4) Data dan ketidakpastian

Konten domain data dan ketidakpastian merupakan konten dalam PISA yang meliputi ketidakpastian dan kesalahan dalam pengukuran, serta berkaitan dengan kemungkinan dan peluang. Pada penerapannya, konten ini banyak menerapkan teori statistika dan peluang, sehingga sering ditemukan bentuk soal yang berupa tabel atau berbagai penyajian data statistik (OECD, 2019, pp. 85–86).

Ketika konten domain dari kedua pengujian tersebut dibandingkan, dapat diketahui bahwa inti konten dari keduanya adalah sama. Poin umum ada pada LANTITE sedangkan untuk perincinya lebih kepada PISA. Sedangkan untuk konten domain pada AKM numerasi berjumlah empat, yang merupakan hasil adaptasi dari PISA. Hal tersebut dipertegas oleh Herman et al. (2022, p. 1) yang menjelaskan bahwa ruang lingkup pada AKM diadaptasi dari tes PISA termasuk didalamnya adalah domain numerasi, sehingga 4 domain tersebut adalah bilangan, geometri dan pengukuran, aljabar, serta data dan ketidakpastian. Keempat domain tersebut akan dipergunakan dalam penelitian ini dikarenakan fokus penelitian adalah pengembangan soal untuk AKM numerasi. Berikut adalah penjabaran dari keempat konten domain AKM numerasi.

##### 1) Bilangan

AKM menetapkan bahwa domain bilangan terbagi menjadi tiga subdomain yakni (1) representasi, yang meliputi representasi bilangan cacah, bilangan bulat, pecahan, desimal, irasional, berpangkat, dan notasi ilmiah; (2) sifat urutan, yang meliputi perbandingan dan urutan dari berbagai jenis bilangan; dan (3) operasi, yang meliputi operasi hitung bilangan (Pusmenjar, 2021a, pp. 80–81). Secara keseluruhan, pelaksanaan AKM numerasi ada pada tiga jenjang pendidikan yakni SD/Sederajat, SMP/Sederajat, dan SMA/SMK/Sederajat. Karena domain bilangan hanya diujikan pada jenjang SD/Sederajat dan SMP/Sederajat, maka penelitian ini

yang memiliki fokus pengembangan soal AKM numerasi pada jenjang SMA tidak mengikutsertakan konten domain bilangan.

2) Geometri dan pengukuran

Domain geometri dan pengukuran pada AKM numerasi terbagi menjadi tiga subdomain yakni (1) bangun geometri, yang meliputi bangun datar, konsep volume, luas permukaan pada bangun ruang, serta konsep kesebangunan; (2) pengukuran, yang meliputi pengukuran panjang, berat, waktu, volume, debit, serta penggunaan satuan luas tidak baku dan baku; serta (3) penalaran spasial, yang meliputi penggunaan arah, sistem koordinat petak, serta sistem koordinat kartesius (Pusmenjar, 2021a, p. 81). Tabel 5 adalah distribusi domain geometri dan pengukuran pada jenjang SMA/SMK/Sederajat yang menjadi fokus penelitian pada pengembangan soal ini (Pusmenjar, 2020, p. 93).

**Tabel 5.** *Distribusi Domain Geometri dan Pengukuran*

Subdomain	Kompetensi
Bangun geometri	1. Memahami dan menggunakan perbandingan trigonometri 2. Menghitung volume dan luas permukaan limas segi-n, kerucut, dan bola
Pengukuran	Tidak ada
Penalaran spasial	Tidak ada

3) Aljabar

Domain aljabar pada AKM numerasi terbagi menjadi tiga subdomain yakni (1) persamaan dan pertidaksamaan, yang meliputi penyelesaian persamaan sederhana hingga sistem persamaan linear tiga variabel; (2) relasi dan fungsi termasuk pola bilangan, yang meliputi pengenalan pola bergambar dan objek, serta penyelesaian

masalah dengan konsep fungsi linear dan eksponensial; serta (3) rasio dan proporsi, yang meliputi rasio atau skala dan aritmatika sosial (Pusmenjar, 2021a, p. 81). Tabel 6 adalah distribusi domain aljabar pada jenjang SMA/SMK/Sederajat yang menjadi fokus penelitian pada pengembangan soal ini (Pusmenjar, 2020, p. 93).

**Tabel 6.** *Distribusi Domain Aljabar*

Subdomain	Kompetensi
Persamaan dan pertidaksamaan	<ol style="list-style-type: none"> <li>1. Menyelesaikan persamaan dan pertidaksamaan kuadrat</li> <li>2. Menyelesaikan sistem persamaan linear dua atau tiga variabel</li> </ol>
Relasi dan fungsi (termasuk pola bilangan)	<ol style="list-style-type: none"> <li>1. Memahami barisan aritmatika dan geometri</li> <li>2. Memahami fungsi kuadrat dan grafiknya, serta sifat-sifatnya</li> </ol>
Rasio dan proporsi	Tidak ada

#### 4) Data dan ketidakpastian

Domain data dan ketidakpastiaan pada pelaksanaan AKM numerasi terbagi menjadi dua subdomain yakni (1) data dan representasinya, yang meliputi penyajian data sederhana menggunakan turus dan diagram hingga pengevaluasian data yang lebih kompleks dan ukuran penyebaran data; serta (2) ketidakpastian dan peluang, yang meliputi pengenalan kejadian yang mungkin dan tidak mungkin hingga menghitung dan menggunakan peluang kejadian majemuk (Pusmenjar, 2021a, pp. 81–82). Tabel 7 adalah distribusi domain aljabar pada jenjang SMA/SMK/Sederajat yang menjadi fokus penelitian pada pengembangan soal ini (Pusmenjar, 2020, p. 93).

**Tabel 7.** *Distribusi Domain Data dan Ketidakpastian*

Subdomain	Kompetensi
Data dan representasinya	Menentukan ukuran penyebaran (jangkauan, simpangan, dan variasi) serta menggunakannya dalam konteks yang bervariasi
Ketidakpastian dan peluang	Memahami dan menggunakan sifat-sifat peluang kejadian

**c. Komponen Konteks Pada Numerasi**

Konteks dalam suatu masalah merupakan aspek pemecahan masalah yang memberi tambahan tuntunan pada pemecahan masalah itu sendiri (OECD, 2019, p. 87). Dengan adanya konteks maka dapat dijangkau kepentingan individu seluas mungkin dari berbagai situasi. Numerasi yang menjadi subjek pengujian tentu memiliki beberapa konteks masalah yang ingin diujikan. Tabel 8 adalah distribusi konteks numerasi dari perspektif LANTITE, PISA, hingga AKM.

**Tabel 8.** *Distribusi Konteks Numerasi*

LANTITE (Thai et al., 2021, p. 75)	PISA (OECD, 2019, p. 88)	AKM (Pusmenjar, 2020, p. 74)
Personal	Personal	Personal
Pekerjaan	Pekerjaan	Sosial-budaya
Pendidikan	Sosial	Saintifik
-	Saintifik	-

Ketika konteks dari ketiga pengujian tersebut dibandingkan, dapat diketahui bahwa ketiganya memiliki persamaan dan perbedaan. Hal yang perlu diketahui bahwa penentuan dan kedalaman konteks juga bergantung dengan subjek pengujian. Untuk PISA dan AKM yang subjek pengujiannya adalah peserta didik maka kedalaman konteks tidak begitu sama seperti LANTITE dimana subjek pengujinya adalah calon pendidik

atau mahasiswa pendidikan guru. Karena fokus penelitian pengembangan soal ini mengarah ke pengujian AKM, maka di bawah ini akan dijabarkan konteks-konteks yang dipergunakan pada AKM.

1) Personal

Konteks personal pada AKM ditujukan pada aktivitas seseorang, keluarga, atau sekelompok orang. Untuk lebih rincinya konteks personal meliputi kegiatan-kegiatan seperti persiapan makan, belanja, bermain, berolahraga, kesehatan pribadi, transportasi pribadi, perjalanan, keuangan pribadi, mengenai hobi, cita-cita, hingga cara individu dalam mengambil keputusan (OECD, 2019, p. 88). Tujuan dari penggunaan konteks personal adalah agar peserta didik mengenali peran matematika dalam kehidupan pribadi dan sehari-hari (Pusmenjar, 2021a, p. 68).

2) Sosial-budaya

Konteks sosial-budaya yang dimaksudkan pada AKM adalah masalah-masalah komunitas atau masyarakat seperti pada sistem pemungutan suara, transportasi publik, pemerintahan, demografi, periklanan, statistika, dan ekonomi nasional (Pusmenjar, 2020, p. 76). Tujuan dari penggunaan konteks sosial-budaya adalah agar peserta didik mengenali peran matematika dalam hidup berkomunitas atau sosialis meskipun pada konteks ini fokusnya pada perspektif masyarakat bukan personal (Pusmenjar, 2021a, p. 69).

3) Sainifik

Konteks saintifik pada AKM berkaitan dengan implementasi matematika di alam semesta serta isu-isu yang berkaitan dengan sains dan teknologi, seperti tentang cuaca atau iklim, ekologi, masalah medis, ruang angkasa, genetika, dan terkait matematika sendiri (OECD, 2019, p. 88; Pusmenjar, 2021a, p. 70). Konteks yang terkait dengan keilmuan matematika disebut konteks intra-matematika, sedangkan terkait dengan keilmuan lain disebut ekstra-matematika (Pusmenjar, 2020, p. 77).

#### d. Komponen Level Kognitif Pada Numerasi

Kognitif sebagai suatu aktivitas mental individu dalam memperoleh suatu pengetahuan baru dengan melewati tahap menghubungkan, menilai, hingga mempertimbangkan suatu fenomena juga terdapat pada numerasi. Dengan adanya level kognitif pada numerasi, maka dapat diketahui aktivitas mental apa yang menjadi tolak ukur dalam numerasi itu sendiri. Tabel 9 adalah distribusi level kognitif numerasi dari perspektif LANTITE, PISA, hingga AKM.

**Tabel 9.** *Distribusi Level Kognitif Numerasi*

<b>LANTITE</b> (Thai et al., 2021, p. 75)	<b>PISA</b> (OECD, 2019, p. 77)	<b>AKM</b> (Pusmenjar, 2020, p. 81)
Mengidentifikasi	Memformulasikan situasi secara matematis	Pengetahuan dan pemahaman
Menerapkan	Menggunakan konsep, fakta, prosedur, dan penalaran matematika	Penerapan
Mengevaluasi dan mengkomunikasikan	Menginterpretasi, menerapkan, dan mengevaluasi hasil matematika	Penalaran

Ketika level kognitif dari ketiga pengujian tersebut dibandingkan, dapat diketahui bahwa ketiganya memiliki kriteria kognitif yang berbeda meskipun beberapa ada yang sama. Perbedaan tersebut didasari atas berbedanya definisi dasar pada numerasi dari ketiga pengujian di atas. Karena fokus penelitian pengembangan soal ini mengarah ke pengujian AKM, maka di bawah ini akan dijabarkan level kognitif yang dipergunakan pada AKM.

1) Pengetahuan dan pemahaman

Level kognitif pengetahuan dan pemahaman atau disebut dengan *knowing* ini merupakan kemampuan dasar dalam mengetahui dan memahami fakta, konsep, proses, serta prosedur dalam matematika (Pusmenjar, 2021a, p. 74). Aspek yang mencakup level kognitif *knowing* adalah mengingat, mengidentifikasi, mengklasifikasikan, menghitung, mengambil/memperoleh, dan mengukur (Pusmenjar, 2020, p. 81)

2) Penerapan

Level kognitif penerapan atau *applying* merupakan kemampuan dalam menerapkan pengetahuan dan pemahaman tentang fakta, proses, konsep, serta prosedur matematika dalam menyelesaikan masalah (Pusmenjar, 2021a, p. 76). Aspek yang mencakup level kognitif *applying* adalah memilih strategi, menyatakan atau membuat model, menerapkan atau melaksanakan, dan menafsirkan (Pusmenjar, 2020, p. 83).

3) Penalaran

Level kognitif penalaran atau *reasoning* merupakan kemampuan bernalar peserta didik dalam menganalisis data dan informasi, menarik kesimpulan, serta memperluas pemahaman dalam kondisi yang baru (Pusmenjar, 2021a, p. 78). Aspek yang mencakup level kognitif *reasoning* adalah menganalisis, memadukan, mengevaluasi, menyimpulkan, serta membuat justifikasi (Pusmenjar, 2020, p. 85).

#### **4. Penilaian Berbasis Komputer (CBT)**

Teknologi yang semakin berkembang pesat berdampak pada pemanfaatan teknologi komputer pada bidang pendidikan, khususnya dalam hal penilaian. Pelaksanaan penilaian dikenal menjadi dua bentuk yakni *Paper Based Test* (PBT) dan *Computer Based Test* (CBT). PBT merupakan bentuk penilaian yang pelaksanaannya secara konvensional menggunakan kertas, sedangkan CBT menggunakan komputer. Pada penelitian ini instrumen soal

yang dikembangkan juga dirancang menggunakan CBT sehingga pada sub-bab ini fokus pembahasan adalah CBT.

CBT merupakan sistem penilaian berbantuan komputer yang memiliki tujuan dalam membantu pendidik melakukan evaluasi baik penskoran, pelaksanaan tes hingga keefektifan serta keefisiensi pelaksanaannya. Kelebihan dari CBT yakni bisa menghemat biaya pelaksanaan penilaian, kemudahan administrasi, keakurasian pada hasil yang tinggi, penilaian serta pelaporan ujian yang lebih mudah diakses (Supriyati et al., 2021, pp. 117–118). Dari kelebihan ini, CBT memiliki hasil yang lebih konsisten daripada PBT. Pada implementasinya CBT terbagi menjadi dua tipe yakni CBT berbasis *web* serta CBT berbasis *desktop*. Pada penelitian ini CBT yang digunakan adalah berbasis *web* dengan pertimbangan kekurangan yang ada pada CBT *desktop* dimana peneliti harus mengatur ulang konfigurasi aplikasi ujian *online* untuk setiap kali mengikuti ujian (Efendi et al., 2021, p. 2). Selain itu butuh waktu yang cukup lama dalam mempersiapkan CBT berbasis *desktop*. Maka dari itu CBT berbasis *web* yang digunakan untuk penelitian ini berbantuan *Learning Management System (LMS) Moodle*.

LMS merupakan aplikasi kompleks yang memberikan penawaran media dan alat untuk memfasilitasi peserta didik dalam pembelajaran online. Sedangkan LMS Moodle merupakan aplikasi *e-learning* berbasis *open source*. Meskipun memiliki hak cipta, Moodle memberi kebebasan dalam menyalin, menggunakan, serta memodifikasi. Selain itu Moodle dapat langsung dipergunakan tanpa harus modifikasi pada *Unix, Linux, Windows*, serta *Mac* (Simanullang & Rajagukguk, 2020, p. 2). Sebagai suatu *software* yang didesain untuk kegiatan belajar mengajar berbasis internet serta *web* (Manurung & Rajagukguk, 2019, pp. 8–9), Moodle telah menjadi alternatif LMS yang paling populer. Hal tersebut diperjelas Widyaningsih et al. (2021, p. 52) dimana pendidik dapat memberikan umpan balik secara langsung kepada peserta didik serta menampilkan jawaban yang benar dari jawaban yang salah oleh peserta didik. Hal ini menjadi salah satu keuntungan dari penggunaan Moodle dalam penilaian pendidikan. Oleh karena itu pengembangan soal pada penelitian ini

berbasis komputer menggunakan CBT *web* berbantuan LMS Moodle dengan menggunakan 4 paket soal, dan diantara paket tersebut terdapat item-item soal yang berperan sebagai *anchor*.

## **B. Kajian Penelitian yang Relevan**

Mendukung kajian teori yang telah diuraikan di atas, maka beberapa penelitian yang terkait dengan pengembangan soal AKM numerasi berbasis komputer untuk kelas XI SMA adalah sebagai berikut:

Penelitian yang dilakukan Pramadya et al. (2019) dengan judul *self-assessment profile on statistics using computer-based mathematical summative test*, menginformasikan bahwa penelitian dengan pendekatan kualitatif ini mendeskripsikan profil penilaian diri peserta didik pada tes sumatif berbasis komputer. Teknik pengambilan data yakni dengan tes matematika sumatif (materi statistika) dan wawancara. Subjek penelitian sebanyak 110 peserta didik dari dua SMA di Tasikmalaya. Tes matematika sumatif tersebut terdiri atas soal-soal pilihan ganda dengan 70 menit waktu pengerjaan. Untuk mengkomputerisasikan tes, penelitian ini berbantuan *Learning Management System (LMS)*. Persamaan penelitian ini dengan penelitian yang akan dilakukan yakni penggunaan CBT dengan bantuan LMS untuk pengujian. Sedangkan perbedaannya adalah pada jenis dan pendekatan penelitian dimana penelitian ini dengan deskriptif kualitatif sedangkan penelitian yang akan dilakukan berfokus pada pengembangan.

Penelitian yang dilakukan Simanullang & Rajagukguk (2020) dengan judul *learning management system based on moodle to improve students learning activity* merupakan penelitian kuasi eksperimen, dimana peneliti melihat pengaruh LMS Moodle terhadap aktivitas belajar peserta didik secara *online*. Memilih LMS Moodle merupakan langkah penelitian yang dilakukan setelah peneliti melakukan analisis kebutuhan. Kemudian Moodle yang akan digunakan didesain dan selanjutnya dilakukan validasi ahli kepada dua ahli media. Hasil validasi menunjukkan bahwa 91,67% LMS Moodle yang telah didesain valid dengan kriteria sangat baik. Beberapa fitur aktivitas belajar yang ada pada desain LMS Moodle ini yakni *video*, forum diskusi, *chat*, materi, dan kuis. Pada penelitian ini teknik pengambilan data berupa observasi yang dilakukan peneliti selama aktivitas belajar

berbantuan LMS Moodle. Terdapat 10 indikator observasi dengan rata-rata hasil observasi terhadap aktivitas belajar peserta didik yakni pada rentang persentase 83% - 90%. Hal tersebut menunjukkan bahwa masing-masing indikator observasi ada pada kategori sangat aktif partisipasi peserta didik dalam aktivitas belajar. Oleh karena itu kesimpulan dari penelitian ini adalah terdapat peningkatan aktivitas belajar sesuai dengan tujuan penelitian yang diharapkan. Persamaan penelitian ini dengan penelitian yang akan dilakukan adalah penggunaan LMS Moodle untuk pelaksanaan penelitian, selain itu kedua penelitian sama-sama melakukan validasi ahli terhadap LMS Moodle yang dipergunakan. Sedangkan perbedaannya terletak pada jenis dan fokus penelitian yang mana pada penelitian ini merupakan kuasi eksperimen, sedangkan penelitian yang akan dilakukan merupakan penelitian pengembangan. Disamping itu, LMS Moodle yang didesain untuk penelitian ini berupa fitur *video*, forum diskusi, *chat*, materi, dan kuis, namun untuk penelitian yang akan dikembangkan hanya pada satu fitur yakni kuis atau penilaian.

Penelitian yang dilakukan Cahyanovianty & Wahidin (2021) dengan judul analisis kemampuan numerasi peserta didik kelas VIII dalam menyelesaikan soal asesmen kompetensi minimum, menginformasikan bahwa peserta didik siap melaksanakan pembaruan sistem penilaian pendidikan di Indonesia, yang mulanya UN kemudian diubah menjadi AKM. Penelitian yang ditujukan kepada 100 peserta didik kelas VIII di SMPN 7 Tambun Selatan, Bekasi, Jawa Barat pada tahun ajaran 2020/2021 menunjukkan bahwa peserta didik lebih menyukai soal AKM yang berbasis konteks karena mudah dipahami dan berkaitan dengan kehidupan sehari-hari, sehingga dapat membantu peserta didik dengan mudah mengeksplor diri dalam belajar matematika. Selain itu, penelitian yang berorientasi menganalisis kemampuan numerasi peserta didik dengan soal AKM memiliki fokus acuan pada 3 komponen yakni konten, konteks, dan proses kognitif. Hasil lain menginformasikan bahwa 75% peserta didik dominan memiliki kemampuan numerasi tingkat sedang, sedangkan 14% dan 11% berturut-turut kecenderungan kemampuan numerasi peserta didik tingkat tinggi dan rendah. Persamaan penelitian ini dengan penelitian yang akan dilakukan adalah 3 komponen utama yang digunakan dalam soal AKM yakni konten, konteks, dan proses kognitif. Sedangkan

perbedaannya terletak pada tujuan utama penelitian, yang mana penelitian ini bertujuan untuk menganalisis atau mendeskripsikan kemampuan numerasi peserta didik berdasarkan soal AKM. Namun untuk penelitian yang akan peneliti lakukan adalah mengembangkan soal AKM numerasi.

Penelitian yang dilakukan Heru et al. (2021) dengan judul *design of supplementary mathematics module for preparation of minimum competency assessment for fifth grade elementary school students* memiliki tujuan untuk mengembangkan modul matematika metode *discovery learning* yang terstandarisasi AKM. Model pengembangan yang digunakan adalah 4D Thiagarajan yang meliputi *define, design, develop, dan disseminate*. Pada penelitian ini uji coba dilakukan sebanyak dua kali yakni uji lapangan terbatas untuk melihat respon peserta didik terhadap produk yang telah dikembangkan dan uji lapangan operasional pada subjek penelitian yakni peserta didik kelas V SD Negeri 3 Mendo Barat tahun ajaran 2020/2021. Teknik pengambilan data pada penelitian ini yakni angket dan tes. Angket dipergunakan untuk memperoleh data kepraktisan modul dari uji coba lapangan terbatas, sedangkan tes untuk memperoleh nilai akhir peserta didik pada uji lapangan operasional. Hasil penelitian menunjukkan bahwa modul yang telah dikembangkan telah valid dan praktis dengan persentase 91,32%, serta rata-rata nilai akhir peserta didik sebesar 87,05 dengan kategori sangat baik. Produk modul yang telah terfinalisasi kemudian disosialisasikan kepada pendidik matematika yang ada di sekolah. Persamaan penelitian ini dengan penelitian yang dikembangkan adalah apa yang dikembangkan terstandarisasi dengan AKM karena harapan dari penelitian ini dan yang akan dikembangkan merupakan cerminan dari AKM itu sendiri, terlebih pada numerasi. Selain itu persamaan terletak pada teknik pengumpulan data yang mana menggunakan angket dan tes. Angket untuk kepraktisan dan tes untuk capaian hasil akhir. Sedangkan perbedaannya terletak pada kompleksitas apa yang dikembangkan, dimana pada penelitian ini modul yang dikembangkan selain meliputi instrumen tes juga terdapat perangkat pembelajaran berupa rencana pelaksanaan pembelajaran (RPP) metode *discovery learning*, selain itu perbedaan terletak pada subjek dan pelaksanaan uji coba, dimana pada penelitian ini subjek penelitian untuk jenjang SD sedangkan penelitian yang akan dikembangkan

ditujukan untuk jenjang SMA. Kemudian untuk uji coba pada penelitian ini dilakukan sebanyak dua kali, namun untuk penelitian yang akan dikembangkan adalah dengan satu kali uji coba.

Penelitian yang dilakukan Machromah et al. (2021) dengan judul *minimum competency assessment: designing tasks to support students' numeracy*, menggunakan model pengembangan ADDIE dalam mengembangkan tugas, namun dengan fokus hanya pada 3 tahap yakni analisis, desain, dan pengembangan. Tugas yang dikembangkan berbasis konteks dan proses kognitif, serta tervalidasi oleh 4 *expert judgment* dari segi konten, *layout*, dan bahasa. Selain itu, teknik pengumpulan data yang digunakan selama proses pengembangan adalah observasi, wawancara, dokumentasi, kuesioner, serta *focus group discussion* (FGD) untuk pengumpulan data tambahan. Hasil pengembangan tugas menunjukkan bahwa kerangka AKM sejalan dengan kerangka kerja PISA yang meliputi konten, konteks, dan tingkat kognitif. Berawal dari hal tersebut, pengembangan tugas yang ditujukan untuk kelas VIII jenjang SMP pada penelitian ini menggunakan kerangka kerja PISA untuk mendesain hingga menghasilkan suatu tugas yang mendukung numerasi peserta didik. Sebanyak tiga tugas yang dinyatakan valid oleh *expert judgment*, yakni (1) konstruksi rangka sebagai konteks ilmiah, (2) pemanfaatan ruang kosong di bawah tangga sebagai konteks sosial budaya, dan (3) jalur terpendek di taman hiburan sebagai tugas pribadi. Setiap tiga tugas tersebut terdiri atas tiga pertanyaan dalam berbagai tingkat kognitif yakni pemahaman, aplikasi, dan penalaran. Persamaan penelitian ini dengan penelitian yang akan dilakukan yakni pengembangan instrumen berupa tugas atau soal yang berbasis pada numerasi peserta didik. Sedangkan perbedaannya terletak pada model pengembangan yang digunakan, dimana pada penelitian ini menggunakan model ADDIE yang berfokus pada analisis, desain, dan pengembangan. Namun, untuk penelitian selanjutnya menggunakan model pengembangan Retnawati (2016a, p. 3) yang terdiri atas 9 langkah. Selain itu, untuk penelitian ini tugas yang dikembangkan untuk kelas VIII jenjang SMP namun untuk yang selanjutnya ditujukan untuk kelas XI jenjang SMA.

Penelitian yang dilakukan Megawati & Sutarto (2021) dengan judul *analysis numeracy literacy skills in terms of standardized math problem on a minimum*

*competency assessment*, menginformasikan bahwa penelitian deskriptif dengan pendekatan kualitatif ini bertujuan untuk menganalisis kemampuan literasi numerasi peserta didik, mengeksplor faktor-faktor yang dapat mempengaruhi kemampuan literasi numerasi, hingga melakukan evaluasi pembelajaran yang dapat meningkatkan kemampuan literasi numerasi peserta didik. Teknik pengambilan data yakni dengan wawancara, kuesioner, dan tes matematika terstandarisasi AKM. Subjek penelitian terdiri atas wakil kepala sekolah bagian kurikulum, pendidik matematika, serta 36 peserta didik kelas XI MIPA 3 di SMA Negeri 16 Semarang. Wawancara tidak terstruktur ditujukan untuk wakil kepala sekolah bagian kurikulum, pendidik matematika, serta beberapa peserta didik dari kelas XI MIPA 3, dengan pedoman wawancara yang disusun sendiri oleh peneliti namun pertanyaan dapat berkembang sesuai dengan kondisi selama wawancara. Instrumen tes terdiri atas 30 pertanyaan sesuai standar AKM dan sudah tervalidasi oleh dosen dan pendidik matematika. Tes dikembangkan sesuai dengan komponen konteks dan konten AKM. Namun untuk bentuk soal, pada instrumen tes ini hanya terdapat 4 bentuk yakni pilihan majemuk, pilihan majemuk kompleks, isian singkat, dan uraian. Sedangkan untuk kuesioner terdapat dua bentuk yang berbeda kegunaan, yakni kuesioner untuk menganalisis kemampuan literasi numerasi peserta didik dan kuesioner untuk mengeksplor faktor-faktor yang dapat mempengaruhi kemampuan literasi numerasi. Persamaan penelitian ini dengan penelitian yang akan dilakukan yakni pengembangan tes numerasi sesuai dengan standar AKM terutama pada bagian konten dan konteks nya. Sedangkan perbedaannya adalah pada fokus penelitian yang lebih mengarah ke kualitatif deskriptif atau menganalisis kemampuan literasi numerasi peserta didik, selain itu pada penelitian ini bentuk soal yang dikembangkan hanya 4 bentuk soal, menghilangkan bentuk soal menjodohkan. Sedangkan pada penelitian yang akan dilakukan bentuk soal sesuai dengan standar AKM yakni ada 5 yang mengikutsertakan bentuk soal menjodohkan.

Penelitian yang dilakukan Widyaningsih et al. (2021) dengan judul *the development of the HOTS test of physics based on modern test theory: question modeling through e-learning of moodle LMS*, menggunakan model pengembangan ADDIE serta menghasilkan seperangkat instrumen soal sebanyak 24 item pilihan

majemuk materi Fisika. Soal yang telah dikembangkan orientasinya adalah HOTS atau mengukur kemampuan tingkat tinggi mahasiswa program studi Pendidikan Fisika di Universitas Papua. Instrumen divalidasi oleh 3 pakar yakni ahli fisika, fisikawan, serta dosen. Kualitas instrumen yang telah dikembangkan dianalisis menggunakan Rasch Model dengan bantuan program *Quest*. Pengambilan data dilakukan secara *online* berbantuan LMS Moodle. Persamaan penelitian ini dengan penelitian yang akan dilakukan yakni penggunaan LMS Moodle untuk merepresentasikan pengembangan instrumen serta analisis kualitas yang digunakan sama-sama menggunakan teori respon item. Sedangkan perbedaannya terletak pada penggunaan model pengembangan dan ruang lingkup instrumen yang dikembangkan, dimana penelitian ini dipergunakan untuk materi Fisika sedangkan penelitian yang akan dilakukan ruang lingkupnya ada pada AKM numerasi.

Penelitian yang dilakukan Purnomo et al. (2022) dengan judul *development of instrument numeracy skills test of minimum competency assessment (MCA) in Indonesia*, menginformasikan bahwa instrumen tes numerasi berbasis AKM yang telah dikembangkan untuk peserta didik kelas VIII SMP internasional berbasis pondok pesantren di Jawa Timur valid, praktis, dan 93,45% sangat layak digunakan. Penelitian ini menggunakan dua langkah dalam mengembangkan instrumen, yakni (1) tahap awal, dimana peneliti melakukan analisis kebutuhan dari pendidik dan peserta didik, serta meninjau ulang kemampuan numerasi berdasarkan kerangka kerja AKM, (2) tahap evaluasi formatif, dimana mencakup analisis instrumen oleh peneliti dalam evaluasi diri, validasi kepada 4 *expert judgment* bidang matematika, serta uji lapangan. Hasil uji lapangan menunjukkan bahwa instrumen yang telah dikembangkan berpotensi memiliki pengaruh terhadap kemampuan numerasi peserta didik sebesar 75,69%. Selain itu, kemampuan numerasi peserta didik tergolong baik dengan 72,92% memiliki kemampuan numerasi sedang, tinggi, dan sangat tinggi. Persamaan penelitian ini dengan penelitian yang akan dilakukan yakni mengembangkan instrumen yang sama yakni numerasi yang berbasis AKM. Sedangkan perbedaannya terletak pada langkah pengembangan, subjek uji lapangan, serta formasi *expert judgment* dimana pada penelitian ini 4 *expert judgment* berlatar belakang ahli bidang matematika, namun untuk penelitian yang

akan dilakukan *expert judgment* selain dari ahli bidang matematika juga ahli pengukuran.

### **C. Kerangka Pikir**

Numerasi merupakan salah satu dari enam literasi dasar yang harus selalu ditingkatkan peserta didik agar mampu menghadapi segala tantangan pada abad ke-21. Pentingnya numerasi dapat diketahui dari banyaknya pelaksanaan penilaian baik skala nasional hingga internasional yang mengujikan numerasi. PISA sebagai penilaian skala internasional untuk menguji literasi matematika atau bisa disebut numerasi menunjukkan bahwa hasil numerasi untuk peserta Indonesia dari periode tahun 2000 hingga 2018 tergolong rendah dan berada pada ranking 1 hingga 8 terbawah dari seluruh negara peserta.

Hasil numerasi yang demikian tidak membuat pemerintah Indonesia diam begitu saja, tahun 2017 pemerintah di bawah naungan Kemendikbud membuat terobosan baru bernama Gerakan Literasi Nasional (GLN) yang didalamnya memuat literasi matematika atau numerasi. Berawal dari sana pemerintah memiliki harapan agar peserta didik di Indonesia terbiasa berliterasi. Selain itu pada tahun 2021 Kemdikbudristek juga membuat terobosan baru terkait pelaksanaan penilaian skala nasional dimana mengganti UN dengan Asesmen Nasional (AN) yang salah satu aspek ujinya yakni AKM, dengan numerasi sebagai salah satu kompetensi dasar yang diuji.

Periode pertama pelaksanaan AKM menunjukkan bahwa hasil numerasi belum merata untuk ke-34 Provinsi di seluruh jenjang pendidikan. Jenjang SD/Sederajat satu Provinsi hasilnya jauh di bawah kompetensi minimum dan 33 Provinsi lain hasilnya di bawah kompetensi minimum. Jenjang SMP/Sederajat 32 Provinsi hasilnya di bawah kompetensi minimum dan dua Provinsi lain sudah mencapai kompetensi minimum. Jenjang SMA/SMK/Sederajat 30 Provinsi hasilnya di bawah kompetensi minimum dan 4 Provinsi lain sudah mencapai kompetensi minimum. Hasil yang cukup mencengangkan ini menjadi bahan evaluasi pemerintah karena berdasarkan harapan di awal, pelaksanaan AKM dipergunakan untuk mengetahui sejauh mana kemampuan peserta didik pada beberapa proses kognitif yang diukur, yakni pemahaman, penerapan, serta penalaran. Maka dari itu, dari awal

pemerintah memvariasikan bentuk soal AKM yang meliputi pilihan majemuk, pilihan majemuk kompleks, menjodohkan, isian singkat, serta uraian.

Fakta lapangan terkait ketersediaan variasi bentuk soal dari berbagai sumber untuk persiapan pelaksanaan AKM menunjukkan persentase di bawah 50%, yang artinya belum cukup tersedia. Selain itu pendidik berupaya untuk mengembangkan instrumen AKM numerasi sendiri namun kenyataannya instrumen tersebut tidak dapat dipergunakan untuk mengukur numerasi sesuai dengan kriteria yang ada pada AKM. Berdasarkan permasalahan tersebut perlunya dilakukan pengembangan soal AKM numerasi agar dapat membantu pendidik dalam menyediakan kebutuhan soal, terlebih untuk melatih peserta didik dalam mempersiapkan pelaksanaan AKM numerasi. Fokus pengembangan soal ada pada jenjang SMA kelas XI dikarenakan penelitian-penelitian terdahulu mengembangkan hal yang sama namun untuk jenjang SD dan SMP.

Kemudian sejak adanya *Covid-19*, pelaksanaan penilaian pendidikan serba dilaksanakan secara *online*, hal tersebut tentu sudah menjadi hal yang lumrah terkait adanya alat ukur sebagai bentuk penilaian berbasis komputer, serta mempertegas bahwa teknologi sudah mampu berdampingan dengan pelaksanaan pendidikan. AKM pada implementasinya dilaksanakan dengan berbasis komputer, sehingga pengembangan soal yang akan dilakukan juga mengarah pada bentuk *Computer Based Test* (CBT) sebagai bentuk dukungan gelombang revolusi industri 4.0. CBT yang dimaksudkan dalam penelitian ini adalah CBT berbasis *web* berbantuan LMS Moodle.

Pengembangan soal AKM numerasi berbasis komputer untuk kelas XI SMA diawali dengan menentukan konstruk instrumen. Konstruk instrumen yang dimaksud didasarkan pada kerangka kerja AKM numerasi yang mencakup komponen konten (geometri dan pengukuran, aljabar, serta data dan ketidakpastian), konteks (personal, sosial-budaya, serta saintifik), dan proses kognitif (*knowing*, *applying*, serta *reasoning*). Selain itu spesifikasi soal yang dikembangkan juga memuat 5 variasi bentuk soal yakni pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, serta uraian. Hal tersebut dilakukan guna instrumen

yang telah dikembangkan dapat memberikan hasil pengukuran yang akurat sesuai dengan tujuan pengembangan.

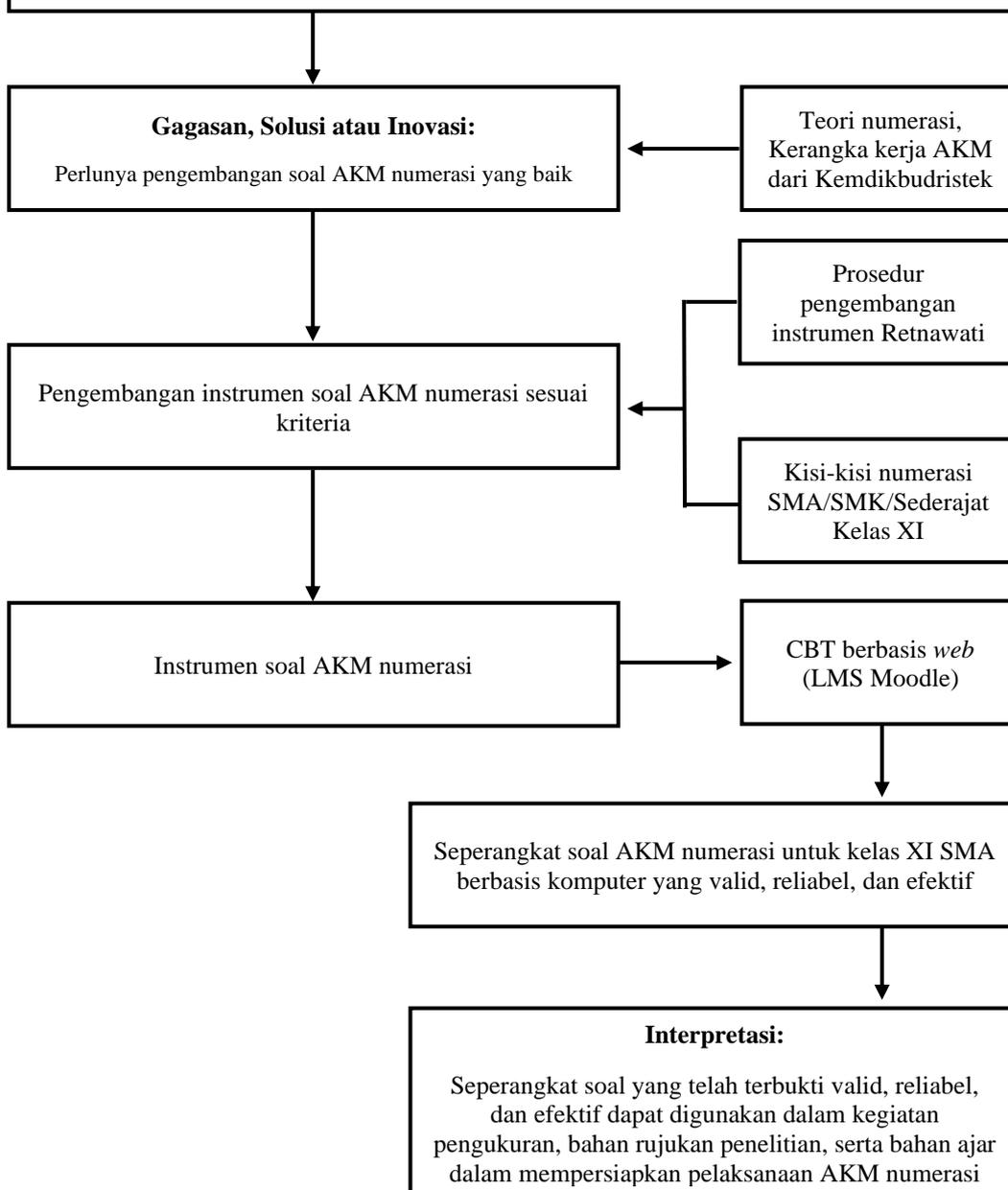
Instrumen yang telah dikembangkan kemudian dibuktikan kualitasnya berdasarkan indeks validitas, estimasi reliabilitas, kesalahan pengukuran, tingkat kesukaran item soal, serta kesetaraan antar paket soalnya. Indeks validitas menunjukkan sejauh mana instrumen yang telah dikembangkan memberikan hasil pengukuran yang akurat dan terpercaya sesuai dengan tujuan pengukuran. Indeks validitas yang dimaksud untuk kualitas instrumen ini mencakup indeks validitas isi dan validitas konstruk. Estimasi reliabilitas dilakukan untuk mengetahui keandalan atau keajegan instrumen dalam memberikan informasi terkait hasil pengukuran. Kesalahan pengukuran diperlukan untuk memahami kesalahan acak yang mempengaruhi skor responden dalam merespons instrumen (Retnawati, 2016a, p. 98). Estimasi tingkat kesukaran item dilakukan untuk menunjukkan item mana saja yang telah dikembangkan tergolong sukar, sedang dan mudah. Tingkat kesetaraan soal dilakukan untuk mengetahui persamaan penyetaraan antar paket soal yang telah dikembangkan.

Instrumen yang telah terbukti kualitasnya kemudian dianalisis profil peserta didiknya dalam menyelesaikan soal AKM numerasi. Analisis profil dilakukan dengan mengestimasi capaian hasil uji coba peserta didik secara menyeluruh menggunakan acuan yang dikemukakan (Mardapi, 2008, p. 123). Memperkuat estimasi yang dilakukan sebelumnya, pada poin profil ini distribusi hasil uji coba peserta didik disetiap kategori juga dideskripsikan.

Bagian terakhir dari penelitian pengembangan ini adalah estimasi keefektifan soal AKM numerasi yang telah dikembangkan serta diketahui kualitas dan profil peserta didik sebagai subjek uji cobanya. Efektivitas diestimasi berdasarkan kriteria perhitungan (Riduwan, 2010, p. 15) yang berguna untuk mengetahui kepraktisan dan kemanfaatan soal yang telah dikembangkan sehingga bisa dipergunakan secara meluas. Hasil penelitian ini diharapkan menjawab pertanyaan penelitian dan menjadi solusi atas masalah yang diajukan. Secara singkat kerangka berpikir tersebut dapat disajikan dalam bagan berikut ini.

### Permasalahan:

1. Kemampuan numerasi peserta didik di Indonesia tergolong sangat rendah jika dibandingkan dengan negara-negara peserta PISA dari setiap periode pelaksanaannya;
2. Hasil pemetaan numerasi untuk setiap Provinsi di setiap jenjang sekolah menunjukkan bahwa mayoritas kemampuan numerasi peserta didik di bawah kompetensi minimum, akan tetapi beberapa Provinsi menunjukkan hasil yang cukup bagus dimana kemampuan numerasi sudah mencapai kompetensi minimum;
3. Belum cukupnya ketersediaan soal AKM numerasi jenjang SMA/Sederajat dari berbagai sumber yang sesuai dengan kriteria AKM.



#### **D. Pertanyaan Penelitian**

Berdasarkan kerangka pikir di atas, pertanyaan penelitian yang diajukan adalah sebagai berikut.

1. Bagaimana konstruk soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
  - a. Bagaimana spesifikasi soal AKM numerasi yang dikembangkan?
  - b. Bagaimana penyesuaian kompetensi dan indikator soal AKM numerasi yang dikembangkan?
  - c. Bagaimana kisi-kisi soal AKM numerasi yang dikembangkan?
2. Bagaimana kualitas soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
  - a. Bagaimana validasi isi soal AKM numerasi yang dikembangkan?
  - b. Bagaimana validasi konstruk soal AKM numerasi yang dikembangkan?
  - c. Bagaimana tingkat kesukaran item dari soal AKM numerasi yang dikembangkan?
  - d. Bagaimana estimasi reliabilitas soal AKM numerasi yang dikembangkan?
  - e. Bagaimana kesalahan pengukuran dari soal AKM numerasi yang dikembangkan?
3. Bagaimana kesetaraan soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan?
  - a. Bagaimana persamaan penyetaraan perangkat soal AKM numerasi paket A sampai paket D dengan metode Rerata & Rerata?
4. Bagaimana profil peserta didik dalam menyelesaikan soal AKM numerasi untuk kelas XI SMA berbasis komputer yang dikembangkan?
  - a. Bagaimana capaian hasil uji coba peserta didik secara menyeluruh dalam menyelesaikan soal AKM numerasi berbasis komputer?
  - b. Bagaimana distribusi hasil uji coba peserta didik dalam menyelesaikan soal AKM numerasi berbasis komputer?
5. Bagaimana keefektifan produk soal AKM numerasi berbasis komputer yang telah dikembangkan?

- a. Bagaimana kepraktisan produk soal AKM numerasi berbasis komputer yang telah dikembangkan ditinjau dari perspektif peserta didik?
- b. Bagaimana kemanfaatan produk soal AKM numerasi berbasis komputer yang telah dikembangkan ditinjau dari perspektif pendidik?

## **BAB III**

### **METODE PENELITIAN**

#### **A. Model Pengembangan**

Penelitian ini memiliki orientasi untuk menghasilkan suatu produk. Produk yang dimaksud berupa instrumen soal AKM numerasi berbasis komputer untuk kelas XI SMA. Dengan demikian, jenis penelitian yang tepat untuk penelitian ini adalah penelitian pengembangan (RD atau *Research Development*). Pada penelitian pengembangan ini, peneliti melakukan pengkajian pendahuluan secara sistematis, melakukan pengembangan soal AKM numerasi berbasis komputer, dan melakukan analisis karakteristik, kualitas soal, serta keefektifan soal berbasis komputer yang telah dikembangkan berdasarkan data uji coba. Model pengembangan instrumen yang digunakan diadopsi dari model yang dikemukakan Retnawati (2016a, p. 3) dengan sembilan langkah, yakni (1) menentukan tujuan penyusunan instrumen soal, (2) mencari teori yang relevan atau menentukan cakupan materi, (3) menyusun indikator dan matriks instrumen soal, (4) menyusun item soal, (5) validasi isi, (6) revisi berdasarkan masukan validator, (7) melakukan uji coba, (8) melakukan analisis, dan (9) merakit instrumen soal. Pengembangan soal didasarkan atas analisis kebutuhan di lapangan berupa kajian teoritis mengenai AKM numerasi. Kemudian pembuatan produk soal disesuaikan dengan kompetensi dari AKM numerasi jenjang SMA/SMK/Sederajat.

#### **B. Prosedur Pengembangan**

Penelitian pengembangan ini dilakukan tanpa menguji hipotesis, tanpa mengkorelasikan atau membandingkan variabel, sehingga prosedur pengembangan yang digunakan mengacu pada langkah-langkah yang dikemukakan Retnawati (2016a, p. 3) sebagaimana berikut:

##### **1. Menentukan Tujuan Penyusunan Instrumen Soal**

Tujuan penyusunan instrumen soal menjadi pondasi awal dalam penelitian pengembangan karena berguna untuk mengkonstruksi instrumen, bentuk instrumen, penskoran, serta pemaknaan hasil. Retnawati (2016, p.

4) menjelaskan bahwa tujuan penyusunan ini harus selaras dengan tujuan penelitian. Oleh karena itu, secara general tujuan dari penelitian ini adalah untuk mengembangkan soal AKM numerasi berbasis komputer untuk kelas XI SMA yang valid dan reliabel.

## **2. Mencari Teori yang Relevan atau Menentukan Cakupan Materi**

Mencari teori yang relevan atau menentukan cakupan materi merupakan langkah kedua dari prosedur pengembangan pada penelitian ini. Soal yang dikembangkan merupakan soal AKM numerasi yang ditujukan untuk peserta didik kelas XI SMA. Sehingga cakupan materi yang dikembangkan mengikuti kisi-kisi yang diujikan pada AKM numerasi yang meliputi domain aljabar, geometri dan pengukuran, serta data dan kepastian. Karena penelitian ini dibuat berbasis komputer (CBT), maka pada langkah kedua juga ditentukan jenis CBT yang digunakan yakni CBT *online* berbasis *web* berbantuan LMS Moodle. Sehingga peneliti perlu melakukan *hosting*.

## **3. Menyusun Indikator dan Matriks Instrumen Soal**

Indikator pencapaian kompetensi pada penelitian ini merupakan bentuk khusus dari kompetensi-kompetensi yang ada pada AKM numerasi. Penyusunan indikator pencapaian kompetensi didasarkan pada cakupan materi dan kompetensi AKM numerasi yang telah ditetapkan. Tujuan dari disusunnya indikator pencapaian kompetensi adalah untuk memudahkan peneliti dalam mengkonstruksi matriks distribusi item serta kisi-kisi instrumen. Selanjutnya, matriks serta kisi-kisi instrumen menjadi pedoman peneliti dalam mengembangkan item soal. Selain menyusun indikator pencapaian kompetensi item soal, pada langkah ini peneliti juga menyusun *storyboard* media CBT yang dipergunakan sebagai media pelaksanaan penelitian.

## **4. Menyusun Item Soal**

Menyusun item soal merupakan langkah keempat dalam prosedur pengembangan instrumen soal. Item soal disusun berdasarkan indikator pencapaian kompetensi yang telah dibuat dan ada pada kisi-kisi instrumen.

Item soal yang disusun sebanyak 4 paket dengan masing-masing paket terdapat 16-17 item dengan 4 item *anchor*, dengan secara keseluruhan untuk (1) distribusi domain geometri dan pengukuran sebanyak 18 item, aljabar sebanyak 14 item, serta data dan ketidakpastian sebanyak 22 item; (2) distribusi konteks personal sebanyak 15 item, sosial budaya sebanyak 25 item, serta saintifik sebanyak 14 item; dan (3) distribusi level kognitif pemahaman (*knowing*) sebanyak 13 item, penerapan (*applying*) sebanyak 29 item, serta penalaran (*reasoning*) sebanyak 12 item. Selain menyusun item soal, peneliti juga mengimplementasikan *storyboard* atau mengkonstruksi media CBT sehingga layak pakai.

## 5. Validasi Isi

Validasi merupakan hal penting yang harus dilakukan sebelum soal yang telah dikembangkan diuji cobakan. Pada langkah ini validasi terjadi dua kali, yakni validasi isi soal AKM numerasi dan validasi media CBT. Sehingga tujuan dari validasi isi adalah untuk mengetahui tingkat kesahihan item-item soal yang telah dikembangkan dan mengetahui kelayakan media CBT untuk pengujian. Tidak hanya item-item soal dan media CBT saja yang dilakukan validasi, akan tetapi kisi-kisi instrumen soal, alternatif jawaban, serta rubrik penskoran. Validasi isi soal AKM numerasi dilakukan oleh 6 ahli yakni dua orang ahli pengukuran, satu orang ahli matematika, dan tiga orang praktisi atau guru matematika. Sedangkan validasi media dilakukan oleh dua ahli media.

Pada langkah validasi ini para ahli menelaah item soal numerasi dan media CBT secara kuantitatif dan kualitatif. Telaah secara kuantitatif dilakukan dengan memberikan skor pada lembar validasi dengan skala interval 1 hingga 4. Sedangkan telaah secara kualitatif dilakukan dengan memberi masukan dan saran terhadap item soal yang dikembangkan. Kemudian formula Aiken digunakan untuk analisis hasil telaah secara kuantitatif dari para ahli.

## **6. Revisi Berdasarkan Masukan Validator**

Revisi dilakukan setelah melakukan validasi isi soal AKM numerasi dan validasi media CBT kepada para ahli. Revisi didasarkan atas masukan dan saran validator baik mengenai item soal, rubrik alternatif jawaban, model penskoran, tata bahasa yang dipergunakan, hingga keberlayakan media CBT. Setelah dilakukan revisi, soal dikonsultasikan ulang kepada validator untuk memastikan bahwa soal AKM numerasi beserta media CBT telah valid.

## **7. Melakukan Uji Coba**

Uji coba dilakukan untuk mendapatkan data empirik mengenai validitas konstruk, tingkat kesukaran, estimasi reliabilitas, kesalahan pengukuran, serta kesetaraan antar paket. Soal diuji coba kepada subjek penelitian yakni peserta didik jenjang SMA Negeri kelas XI di Kota Yogyakarta. Dengan dilakukannya uji coba, dapat diketahui soal-soal yang telah disusun memenuhi kualitas ataupun yang belum memenuhi kualitas item soal yang baik, serta keefektifan media CBT yang dipergunakan, sehingga kemudian dapat dilakukan perbaikan ulang.

## **8. Melakukan Analisis**

Analisis dari hasil uji coba dilakukan dengan pendekatan teori respon butir berbantuan program R. Analisis yang dilakukan pada langkah ini yaitu, validitas konstruk, tingkat kesukaran, estimasi reliabilitas, kesalahan pengukuran, serta estimasi persamaan kesetaraan antar paket soal. Sedangkan untuk profil peserta didik berupa data kuantitatif berasal dari skor hasil uji coba. Skor diinterpretasi sesuai dengan kriteria penilaian AKM yang telah ditetapkan, yakni mahir, cakap, dasar, dan perlu intervensi khusus. Hasil interpretasi tersebut menjadi tolak ukur kemampuan peserta didik SMA kelas XI di Kota Yogyakarta dalam menyelesaikan soal AKM numerasi yang telah dikembangkan.

## **9. Merakit Instrumen Soal**

Merakit instrumen soal merupakan langkah terakhir dalam prosedur pengembangan soal. Pada tahap ini item soal berbasis LMS Moodle yang telah dianalisis dan diperbaiki dirakit ulang menjadi sebuah perangkat soal yang telah memiliki karakteristik item soal yang baik, valid, reliabel, dan efektif dipergunakan. Setelah seperangkat instrumen soal dirakit dengan baik, maka hasil rakitan tersebut dapat digunakan untuk pengukuran serta dimanfaatkan secara umum diberbagai daerah di Indonesia.

### **C. Desain Uji Coba Produk**

Uji coba produk dilakukan untuk mengetahui seberapa layak produk yang dikembangkan. Desain uji coba produk meliputi desain uji coba, subjek uji coba, teknik dan instrumen pengumpulan data, serta teknik analisis data. Berikut adalah deskripsi terkait desain uji coba produk dari pengembangan soal AKM numerasi berbasis komputer untuk kelas XI SMA.

#### **1. Desain Uji Coba**

Pada penelitian ini, langkah pertama yang dilakukan peneliti adalah mendesain dengan detail langkah-langkah pengembangan guna memperoleh data empirik untuk mengetahui kualitas soal. Kemudian uji coba dilakukan pada peserta didik SMA Negeri kelas XI di Kota Yogyakarta. Berasal dari pelaksanaan uji coba ini validitas konstruk, tingkat kesukaran, reliabilitas, kesalahan pengukuran, persamaan penyetaraan antar paket soal, profil peserta didik, serta keefektifan LMS Moodle yang digunakan dapat terestimasi dengan baik. Hasil akhir dari desain ini berupa seperangkat soal AKM numerasi berbasis komputer yang telah direvisi sehingga diperoleh seperangkat soal yang lebih baik.

#### **2. Subjek Uji Coba**

Subjek uji coba merupakan peserta didik di kelas XI SMA Negeri di Kota Yogyakarta tahun ajaran 2022/2023. Pemilihan subjek uji coba didasarkan pada kebutuhan pengembangan. Pada penelitian ini, pemilihan subjek uji coba merupakan keterbatasan penelitian dimana subjek uji coba seharusnya dilakukan secara nasional di berbagai daerah di Indonesia,

namun kemudian ditetapkan hanya pada satu kota dalam Provinsi D.I. Yogyakarta yang dapat dijangkau peneliti. Selain itu, berdasarkan data dari Dinas Pendidikan Pemuda dan Olahraga D.I. Yogyakarta, hasil AKM numerasi tahun 2021 untuk jenjang SMA/SMK/Sederajat sudah mencapai kompetensi minimum. Hal tersebut menunjukkan tingkat numerasi yang cakap serta cukup banyak peserta didik berada pada level mahir. Dengan pertimbangan tersebut, pemilihan subjek uji coba pada penelitian ini diharapkan hasilnya akan sejalan dengan hasil AKM numerasi 2021 serta memperjelas potret kemampuan numerasi peserta didik.

Sekolah yang dipilih sebagai subjek uji coba adalah SMA Negeri di Kota Yogyakarta yang telah mengikuti pelaksanaan AKM 2021. Banyaknya sekolah yang digunakan disesuaikan dengan banyaknya subjek yang dibutuhkan untuk mengerjakan soal yang telah dikembangkan, terlebih pada pengembangan ini terdapat empat paket soal numerasi. Disamping itu, sekolah juga bersedia dilakukan pengambilan data pada peserta didiknya. Ukuran minimal sampel yang layak untuk setiap paketnya sebanyak 250-500 (Şahin & Aml, 2017, p. 322). Berkaitan dengan ukuran sampel tersebut, peneliti kemudian mendiskusikan kembali dengan pakar psikometrik, sehingga menjadikan ukuran sampel sebagai keterbatasan penelitian. Diskusi tersebut menghasilkan keputusan bahwa per paket soal dilakukan pengujian pada 300 responden. Sehingga total responden yang diperlukan dalam penelitian ini tidak kurang dari 1200 responden. Teknik sampling yang digunakan adalah *purposive sampling*. Jumlah subjek uji coba sebanyak 1570 peserta didik yang terdiri atas 321 dari SMAN 1 Yogyakarta, 300 dari SMAN 2 Yogyakarta, 232 dari SMAN 3 Yogyakarta, 257 dari SMAN 5 Yogyakarta, 214 dari SMAN 6 Yogyakarta, dan 246 dari SMAN 11 Yogyakarta.

### **3. Teknik dan Instrumen Pengumpulan Data**

#### **a. Teknik Pengumpulan Data**

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah tes AKM numerasi dan kuesioner. Tes AKM numerasi terdiri atas 4 paket soal dengan masing-masing paket terdapat 16-17 soal dengan 4 item *anchor*. Materi soal disesuaikan dengan kompetensi pada AKM numerasi jenjang SMA/SMK/Sederajat yang mana terdapat tiga konten, yaitu (1) aljabar, (2) geometri dan pengukuran, serta (3) data dan ketidakpastian. Disamping itu, terdapat lima subkonten, yaitu (1) persamaan dan pertaksamaan, (2) relasi dan fungsi (termasuk pola bilangan), (3) bangun geometri, (4) data dan representasinya, serta (5) ketidakpastian dan peluang. Sedangkan kuesioner digunakan untuk mengetahui keefektifan LMS Moodle dengan melihat respon subjek uji coba utama dan tambahan yakni peserta didik dan pendidik matematika.

#### **b. Instrumen Pengumpulan Data**

Instrumen pengumpulan data dalam penelitian ini berupa lembar validasi, *draft* tes yang terkomputerisasi pada LMS Moodle, serta angket. Lembar validasi merupakan lembar penilaian instrumen yang berupa tabel skor relevansi item soal dengan indikator. Lembar validasi diberikan kepada para ahli atau validator untuk dinilai isi instrumennya sebelum diujicobakan. Apabila terdapat masukan dan saran dari validator yang mengharuskan peneliti melakukan perbaikan, maka terlebih dahulu diperbaiki isi instrumennya sebelum diujicobakan. *Draft* tes yang berupa kisi-kisi, petunjuk pengerjaan, item-item soal AKM numerasi, serta rubrik penskoran dipergunakan untuk memperoleh respon atau jawaban peserta didik yang kemudian menjadi data dalam mengetahui kualitas soal yang telah dikembangkan. Sedangkan angket merupakan instrumen penilaian keefektifan LMS Moodle yang dipergunakan peserta didik dalam menyelesaikan tes. Angket terdiri atas seperangkat pernyataan dengan skor penilaian berupa skala likert. Angket diberikan kepada peserta didik untuk menilai kepraktisan soal,

sedangkan angket yang lain diberikan kepada pendidik matematika untuk menilai kemanfaatan soal AKM numerasi berbasis komputer yang telah dikembangkan.

#### 4. Teknik Analisis data

##### a. Analisis Hasil Validitas Isi

Pada penelitian ini, validitas isi dipergunakan untuk seluruh instrumen pengumpulan data, dan dilakukan sebelum instrumen diujicobakan. Validitas isi dari suatu instrumen ditentukan oleh kesepakatan para ahli atau validator. Para ahli meyakini bahwa suatu instrumen valid apabila instrumen mengukur penguasaan kemampuan yang didefinisikan dalam domain ataupun konstruk psikologi yang diukur (Retnawati, 2016a, p. 18). Instrumen validitas isi menggunakan lembar validasi dengan skala interval 1 hingga 4. Skala 1 bermakna sangat tidak sesuai, skala 2 bermakna tidak sesuai, skala 3 bermakna sesuai, dan skala 4 bermakna sangat sesuai. Skala tersebut menunjukkan kesesuaian item dengan indikator yang hendak diukur. Kesepakatan para ahli diolah menggunakan formula *Aiken* sebagai berikut (Retnawati, 2016a, p. 18):

$$V = \frac{\sum s}{n(c - 1)} \quad (1)$$

Keterangan:

- $V$  = indeks kesepakatan para ahli mengenai validitas item
- $s$  =  $r - I_0$
- $r$  = skala penilaian yang diberikan para ahli
- $I_0$  = skala penilaian validitas terendah
- $c$  = skala penilaian validitas tertinggi
- $n$  = jumlah para ahli/validator

Indeks  $V$  berada pada rentang nilai 0 hingga 1. Semakin mendekati nilai 1, maka validitas dari suatu item semakin tinggi. Begitupun saat indeks  $V$  mendekati nilai 0, maka validitas itemnya

semakin rendah. Pada penelitian ini, kategori penilaian validitas isi menggunakan kategori Retnawati (2016a, p. 19), dimana jika indeks  $V \leq 0,4$  maka validitasnya kurang, jika indeks  $V$  diantara  $0,4 - 0,8$  maka validitasnya sedang, dan jika indeks  $V \geq 0,8$  maka validitasnya tinggi. Sedangkan data kualitatif diperoleh dari masukan dan saran para ahli setelah dilakukannya validitas isi. Oleh karena itu, sekiranya peneliti memperbaiki instrumen sesuai dengan masukan dan saran hingga memperoleh instrumen yang siap diujicobakan kepada subjek uji coba yang telah ditetapkan.

**b. Analisis Hasil Validitas Konstruk**

Data hasil uji coba peserta didik dianalisis menggunakan *Confirmatory Factor Analysis* (CFA), dimana butir instrumen dibuat berdasarkan indikator dari teori yang sudah kuat. Sehingga analisis CFA dalam penelitian ini bertujuan untuk mengkonfirmasi kisi-kisi pada level domain numerasi. Analisis CFA dilakukan menggunakan program R dengan terdapat dua hal yang akan dianalisis, yakni validasi konstruk instrumen, serta validasi butir. Validasi konstruk instrumen bertujuan untuk melihat kesesuaian model instrumen dengan data empiris. Pembuktian validitas konstruk ini menggunakan kriteria *Goodness of Fit Test* (GOF) dengan kriteria nilai dapat dilihat pada Tabel 10 (Hair et al., 2013; Zainudin et al., 2019, p. 605).

**Tabel 10.** *Kriteria Goodness of Fit Test (GOF)*

Ukuran GOF	Batas Nilai
GFI	$\geq 0,90$
RMR	$\leq 0,50$
RMSEA	$\leq 0,80$
NNFI	$\geq 0,90$
AGFI	$\geq 0,90$
CFI	$\geq 0,90$
IFI	$\geq 0,90$

Pembuktian validitas butir dilakukan dengan melihat nilai *loading factor*. Suatu butir dikatakan valid secara konstruk jika nilai *loading factor* lebih dari 0,3 (Apriandi et al., 2022, p. 1495; Hair et al., 2013; Wijono & Mardapi, 2016, p. 237).

### **c. Analisis Tingkat Kesukaran Item dengan Pendekatan IRT Data Campuran**

Hasil uji coba dapat dilakukan analisis dengan pendekatan teori respon item (*item response theory/IRT*) dengan data campuran dikotomus dan politomus apabila telah memenuhi uji asumsi yang disyaratkan pada pendekatan tersebut. Uji asumsi yang dimaksud yakni asumsi unidimensi, asumsi independensi lokal, dan asumsi invariansi parameter. Berikut adalah penjelasan dari ketiga uji asumsi tersebut.

#### a) Uji Asumsi Unidimensi

Salah satu cara untuk uji asumsi unidimensi adalah dengan menggunakan analisis faktor untuk melihat nilai eigen pada matriks varians kovarians inter-item (Retnawati, 2014, p. 4). Tujuan uji asumsi unidimensi adalah menunjukkan bahwa hanya ada satu komponen yang dominan dalam data hasil uji coba.

#### b) Uji Asumsi Independensi Lokal

Hasil uji unidimensi berpengaruh terhadap uji asumsi independensi lokal, karena asumsi independensi lokal otomatis terbukti apabila data hasil uji coba memenuhi asumsi unidimensi (Retnawati, 2014, p. 7).

#### c) Uji Asumsi Invariansi Parameter

Uji asumsi invariansi parameter terbagi menjadi dua jenis, yakni invariansi parameter item dan invariansi parameter kemampuan. Asumsi parameter item dibuktikan dengan melakukan estimasi parameter tingkat kesukaran item pada peserta didik yang mengikuti uji coba dengan kelompok yang berbeda, misalnya kelompok berdasarkan jenis kelamin, tempat tinggal, status ekonomi sosial, dan lain-lain (Retnawati, 2014, p. 7). Sedangkan untuk asumsi

invariansi parameter kemampuan dibuktikan dengan mengestimasi parameter kemampuan peserta didik menggunakan item nomor genap dan item nomor ganjil (Retnawati, 2014, p. 9). Hasil dari uji invariansi parameter berupa diagram pencar, sehingga apabila titik-titik pada diagram pencar mendekati garis yang melewati titik asal dengan gradien 1, maka dapat diasumsikan bahwa parameter item atau kemampuan yang diukur invarian (Retnawati, 2014, p. 8).

Setelah data memenuhi uji asumsi IRT maka dapat dilakukan estimasi parameter item. Pada penelitian ini parameter item yang digunakan adalah tingkat kesukaran item ( $b$ ). Tingkat kesukaran item menunjukkan jumlah sifat laten atau kemampuan peserta tes yang diperlukan untuk memiliki probabilitas 0,50 dalam menjawab item soal dengan benar. Parameter  $b$  memiliki rentang nilai dari -2,00 s/d +2,00 yang artinya parameter  $b$  tergolong baik karena memiliki indeks kesukaran item yang sedang (Aida et al., 2017, p. 133; DOĞRUÖZ & AKIN ARIKAN, 2020, p. 52). Selain itu Retnawati (2014, p. 17) juga menjelaskan bahwa nilai parameter  $b$  yang mendekati -2,00 atau kurang dari -2,00 maka indeks kesukaran itemnya sangat rendah, sedangkan apabila nilai parameter  $b$  mendekati +2,00 atau lebih dari +2,00 maka indeks kesukaran itemnya sangat tinggi untuk suatu kelompok peserta tes.

#### **d. Analisis Estimasi Reliabilitas**

Koefisien reliabilitas merupakan koefisien kestabilan atau keajegan dari suatu hasil pengukuran (Retnawati, 2016a, p. 84). Sehingga apabila hasil pengukuran memiliki koefisien reliabilitas tinggi, maka kecil kesalahan yang terjadi pada hasil pengukuran yang diperoleh, begitupun sebaliknya. Pada penelitian ini koefisien reliabilitas diperoleh dari perhitungan reliabilitas *Cronbach's Alpha*. Berikut adalah formula *Alpha* yang digunakan untuk mengestimasi koefisien reliabilitas yang telah diujicobakan (Retnawati, 2016a, p. 91):

$$\alpha = \left( \frac{k}{k-1} \right) \left( 1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right) \quad (2)$$

Keterangan:

$\alpha$  = koefisien reliabilitas instrumen

$k$  = banyak item pertanyaan dalam instrumen

$\sum \sigma_i^2$  = jumlah varians item instrumen

$\sigma_t^2$  = varians skor total

Koefisien reliabilitas pada umumnya ada pada rentang -1,00 hingga 1,00. Menurut Retnawati (2016a, p. 85), koefisien tinggi menunjukkan reliabilitas tinggi, sedangkan koefisien suatu skor tes rendah maka reliabilitas suatu instrumen juga rendah. Sedangkan Guilford (1956, p. 145) mengkategorikan koefisien realibilitas sebagai berikut:

**Tabel 11.** *Kategori Koefisien Reliabilitas*

Skor Total	Kategori
$0,80 < r_{xx'} \leq 1,00$	Sangat tinggi
$0,60 < r_{xx'} \leq 0,80$	Tinggi
$0,40 < r_{xx'} \leq 0,60$	Sedang
$0,20 < r_{xx'} \leq 0,40$	Rendah
$-1,00 \leq r_{xx'} \leq 0,20$	Sangat rendah

Keterangan:

$r_{xx'}$  = koefisien reliabilitas

**e. Analisis Kesalahan Baku Pengukuran (SEM)**

Kesalahan baku pengukuran memiliki hubungan yang berbanding terbalik dengan fungsi informasi tes. Fungsi informasi tes merupakan jumlah dari banyaknya fungsi informasi item dari suatu tes yang dapat diestimasi menggunakan formula sebagai berikut (Hambleton & Swaminathan, 1985, p. 94):

$$I_i(\theta) = \sum_{i=1}^n \frac{D^2 e^{D(\theta-b_i)}}{(1 + e^{D(\theta-b_i)})^2} \quad (3)$$

Keterangan:

$I_i(\theta)$  = fungsi informasi tes

$D$  = faktor skala

$e$  = bilangan natural mendekati 2,718

$\theta$  = tingkat kemampuan subjek

$b_i$  = indeks kesukaran item ke-i

Estimasi kesalahan baku pengukuran dapat diestimasi menggunakan formula berikut (Hambleton & Swaminathan, 1985, p. 104):

$$SEM(\theta) = \frac{1}{\sqrt{I_i(\theta)}} \quad (4)$$

Keterangan:

$SEM(\theta)$  = estimasi kesalahan baku pengukuran

$I_i(\theta)$  = fungsi informasi tes

#### f. Analisis Penyetaraan

Menurut Kolen & Brennan (2014) penyetaraan merupakan proses statistik untuk menyetarakan skor dari dua atau lebih tes yang instrumennya memiliki spesifikasi yang sama. Tujuan dari penyetaraan adalah menghasilkan skor yang dapat saling menggantikan antar tes satu dengan tes lainnya (Retnawati, 2014, p. 93). Pada penelitian ini untuk melakukan penyetaraan dilakukan pembuktian secara statistik dengan IRT. Pada IRT, penyetaraan merupakan proses menempatkan skor dari dua atau lebih bentuk tes paralel ke skala skor umum (Hambleton et al., 1991; Yusron et al., 2020, p. 3).

Pada penelitian ini jenis penyetaraan yang dilakukan adalah penyetaraan horizontal, dimana instrumen tes yang disetarakan mengukur kelas yang sama yakni kelas XI SMA. Tes yang saling dibandingkan diberikan kepada peserta tes yang memiliki kemampuan

sama. Desain penyetaraan yang digunakan adalah desain tes dengan butir bersama (*anchor*), serta metode penyetaraannya adalah metode Rerata & Rerata. Menurut Loyd & Hoover (1980, p. 180) metode Rerata & Rerata memiliki tujuan untuk menentukan konstanta yang melibatkan dua parameter yakni parameter daya pembeda dan tingkat kesukaran. Sehingga konstanta  $\alpha$  dan  $\beta$  dapat dihitung dengan rerata dari parameter yang terlibat. Berikut adalah formula yang digunakan dalam metode Rerata & Rerata untuk memperoleh konstanta penyetaraan  $\alpha$  dan  $\beta$  (Retnawati, 2014, p. 106)

$$\alpha = \frac{\bar{a}_1}{\bar{a}_2} \quad (5)$$

$$\beta = \bar{b}_2 - \alpha \bar{b}_1 \quad (6)$$

Keterangan:

$\bar{b}_1$  &  $\bar{b}_2$  = rerata indeks kesukaran butir bersama tes 1 dan tes 2

$\bar{a}_1$  &  $\bar{a}_2$  = rerata indeks daya beda butir bersama tes 1 dan tes 2

$\alpha$  &  $\beta$  = konstanta penyetaraan

Apabila skala tes 1 disetarakan dengan skala tes 2 maka hubungan parameter butir dan kemampuan peserta untuk dua skala tes dapat dinyatakan sebagai berikut (Retnawati, 2014, p. 104).

$$\theta_{1i}^* = \alpha \theta_{1i} + \beta \quad (7)$$

$$\alpha_{1j}^* = \frac{\alpha_{1j}}{\alpha} \quad (8)$$

$$b_{1j}^* = \alpha b_{1j} + \beta \quad (9)$$

$$c_{1j}^* = c_{1j} \quad (10)$$

Keterangan:

$\alpha_{1j}, b_{1j}, c_{1j}$  = parameter butir untuk butir  $j$  pada skala tes 1

$\alpha_{1j}^*, b_{1j}^*, c_{1j}^*$	= parameter butir untuk butir $j$ pada skala tes 1 setelah disetarakan dengan tes 2
$\theta_{1i}$	= kemampuan peserta $i$ pada skala tes 1
$\theta_{1i}^*$	= kemampuan peserta $i$ pada skala tes 1 setelah disetarakan dengan tes 2
$\alpha$ & $\beta$	= konstanta penyetaraan

### g. Analisis Profil Berdasarkan Hasil Pengukuran Uji Coba

*Output* dari uji coba adalah data respons jawaban peserta didik yang kemudian dilakukan pengukuran untuk mengestimasi kemampuan setiap peserta didik. Data dianalisis sehingga diketahui capaian hasil peserta didik secara menyeluruh dalam menyelesaikan soal AKM numerasi dan kemudian dikategorikan. Selain diketahui capaian hasil secara menyeluruh maka dapat diketahui pula distribusi hasil uji coba untuk setiap kategori kemampuan peserta didik.

Estimasi kemampuan peserta didik yang diperoleh dikonversi ke dalam skala 0 hingga 100 menggunakan rumus *T-Score* (Kurpius & Stafford, 2006, p. 84).

$$T = 10(Z) + 50 \quad (11)$$

Nilai yang diperoleh kemudian dikategorikan berdasarkan rumus yang dikemukakan oleh (Mardapi, 2008, p. 123) serta dikombinasikan dengan tingkat kompetensi yang menjadi acuan pada rapor pendidikan AKM (Kemendikbudristek, 2022).

**Tabel 12.** *Tingkat Kompetensi yang Dicapai Peserta Didik*

Interval Kemampuan	Tingkat Kompetensi
$\theta > \bar{X} + 1,5.SD$	Mahir
$\bar{X} < \theta \leq \bar{X} + 1,5.SD$	Cakap
$\bar{X} - 1,5.SD < \theta \leq \bar{X}$	Dasar
$\theta \leq \bar{X} - 1,5.SD$	Perlu intervensi khusus

Keterangan:

$$\bar{X} = \text{rata-rata ideal} \left( \frac{\text{skor maksimum} + \text{skor minimum}}{2} \right)$$

$$SD = \text{standar deviasi ideal} \left( \frac{\text{skor maksimum} - \text{skor minimum}}{6} \right)$$

$\theta$  = kemampuan peserta didik

Keempat tingkat kompetensi tersebut merupakan atribut peserta didik dalam pelaporan hasil AKM numerasi berdasarkan ketentuan Kemendikbudristek. Keempat tingkat kompetensi tersebut yakni mahir, cakap, dasar, dan perlu intervensi khusus. Kriteria keempat tingkat kompetensi tersebut telah dideskripsikan pada Bab II.

Selain itu, estimasi kemampuan peserta didik dipergunakan dalam menetapkan indeks numerasi yang telah dicapai peserta didik dalam rentang nilai 1,00 – 3,00 sesuai acuan pada rapor pendidikan oleh Kemendikbudristek. Indeks numerasi yang ditetapkan terdapat empat indeks, yakni di atas kompetensi minimum (rentang nilai 2,10 – 3,00), mencapai kompetensi minimum (rentang nilai 1,80 – 2,09), di bawah kompetensi minimum (rentang nilai 1,40 – 1,79), dan jauh di bawah kompetensi minimum (rentang nilai 1,00 – 1,39). Kriteria keempat indeks numerasi tersebut juga telah dideskripsikan pada Bab II.

#### **h. Analisis Keefektifan Produk**

Memberikan angket kepada peserta didik dan pendidik serta menganalisis responnya merupakan langkah yang dilakukan untuk mengetahui keefektifan soal AKM numerasi berbasis komputer yang telah dikembangkan. Pada penelitian ini keefektifan produk menggambarkan sejauh mana tujuan telah terpenuhi. Instrumen soal AKM numerasi berbasis komputer yang telah dikembangkan memiliki keefektifan jika berfungsi secara tepat dalam mencapai tujuannya, dalam hal ini yakni mengungkap tingkat kepraktisan dan kemanfaatan soal.

Tingkat kepraktisan soal mengukur sejauh mana produk yang telah dikembangkan dapat digunakan atau memiliki kemudahan dalam

penggunaan. Angket kepraktisan soal ditujukan kepada peserta didik dengan skala penilaian interval 1 hingga 5. Tingkat kepraktisan yang dinilai mencakup empat aspek yakni, (1) tampilan, (2) konstruksi soal, (3) bahasa, dan (4) keberfungsian *tools* pada media.

Tingkat kemanfaatan soal mengukur sejauh mana produk yang telah dikembangkan bermanfaat terlebih untuk penggunaan lebih lanjut. Angket kemanfaatan soal ditujukan kepada pendidik matematika dengan skala penilaian interval 1 hingga 5. Tingkat kemanfaatan yang dinilai mencakup dua aspek yakni, (1) konten, dan (2) media.

Respon peserta didik dan pendidik kemudian dianalisis menggunakan rumus sebagai berikut untuk mengetahui kategori kepraktisan dan kemanfaatan produk yang telah dikembangkan.

$$N = \frac{k}{N_k} \times 100\% \quad (12)$$

Keterangan:

$N$  = persentase keefektifan

$k$  = skor hasil pengumpulan data

$N_k$  = skor keseluruhan nilai tertinggi

Kriteria yang digunakan dalam perhitungan kepraktisan dan kemanfaatan di atas adalah sebagai berikut (Riduwan, 2010, p. 15).

**Tabel 13.** *Kategori Keefektifan*

Interval Kriteria	Kategori
$81\% \leq N < 100\%$	Sangat praktis/ Sangat manfaat
$61\% \leq N < 80\%$	Praktis/ Bermanfaat
$41\% \leq N < 60\%$	Cukup praktis/ Cukup bermanfaat
$21\% \leq N < 40\%$	Tidak praktis/ Tidak manfaat
$N < 20\%$	Sangat tidak praktis/ Sangat tidak manfaat

## **BAB IV**

### **HASIL PENELITIAN DAN PEMBAHASAN**

#### **A. Hasil Penelitian**

##### **1. Hasil Pengembangan Produk Awal**

Pada penelitian ini produk yang dikembangkan berupa soal AKM numerasi kelas XI SMA berbasis komputer. Materi numerasi disesuaikan dengan kisi-kisi AKM numerasi nasional jenjang SMA/SMK/Sederajat yang meliputi domain aljabar, geometri dan pengukuran, serta data dan ketidakpastian. Soal AKM numerasi yang dikembangkan terdiri atas 5 jenis bentuk soal yakni pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian. Pada pengimplementasiannya, soal AKM numerasi disajikan dalam bentuk *Computerized Based Test* (CBT) berbasis *web* menggunakan LMS Moodle. Model pengembangan produk diadopsi dari model yang dikemukakan Retnawati (2016, p. 3) dengan sembilan langkah. Berikut pendeskripsian hasil pengembangan produk sesuai dengan sembilan langkah tersebut.

##### **a. Hasil Menentukan Tujuan Penyusunan Instrumen Soal**

Penelitian pengembangan ini didasarkan atas analisis kebutuhan berupa kajian teoritis mengenai keberadaan soal AKM numerasi sebagai bahan persiapan pelaksanaan Asesmen Nasional (AN). Tahun 2021 merupakan tahun pertama pelaksanaan AN sebagai pengganti UN. Sebagai salah satu upaya persiapan pelaksanaan AKM, para pendidik berupaya mengembangkan instrumen soal sesuai standar AKM terlebih numerasi namun hasil penelitian Purnomo et al. (2022, p. 637) menunjukkan hasil instrumen tersebut tidak dapat dipergunakan untuk mengukur numerasi peserta didik. Disamping itu, salah satu inovasi pelaksanaan AN yakni terdapat 5 bentuk soal dalam pengujian AKM yakni soal bentuk pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian. Dengan adanya inovasi baru tersebut ternyata di lapangan belum menunjukkan ketersediaan yang cukup bahan referensi soal AKM dengan presentasi

di bawah 50%, hal tersebut ditegaskan dalam penelitian (Widarti et al., 2022, p. 66). Berawal dari analisis kebutuhan di atas, diperlukannya penelitian pengembangan yang memiliki fokus pada keberadaan soal AKM numerasi.

Penelitian pengembangan ini mengerucut pada jenjang SMA/SMK/Sederajat dengan pertimbangan penelitian Machromah et al. (2021), Nuzulia & Gafur (2022), Purnomo et al. (2022), serta Sulistyani & Kusumawardana (2022) yang memiliki fokus pengembangan soal AKM numerasi di jenjang SD/Sederajat dan SMP/Sederajat saja. Karena pada pelaksanaan AKM terkomputerisasi, pada penelitian ini juga dikembangkan soal yang sudah terkomputerisasi atau dengan kata lain menggunakan media CBT. Sehingga tujuan penyusunan instrumen soal ini yakni untuk mengembangkan soal AKM numerasi berbasis komputer untuk kelas XI SMA yang baik. Kualitas soal yang baik diperoleh dengan membuktikan konstruksi intrumennya, estimasi reliabilitas, hingga estimasi keefektifan soal berbasis komputer yang telah dikembangkan.

**b. Hasil Mencari Teori yang Relevan atau Menentukan Cakupan Materi**

Instrumen yang dikembangkan merupakan soal AKM numerasi untuk kelas XI SMA. Dengan demikian, cakupan materi yang digunakan disesuaikan dengan kisi-kisi yang diujikan pada pelaksanaan AKM numerasi yang meliputi kesamaan pada domain, subdomain, serta kompetensi yang telah ditetapkan oleh Badan Standar, Kurikulum, dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi pada Prosedur Operasional Standar (POS) pelaksanaan Asesmen Nasional terutama AKM numerasi 2021.

**Tabel 14.** Kisi-kisi Numerasi SMA (Pusmenjar, 2020, p. 95)

Domain	Subdomain	Kompetensi
Aljabar	Persamaan dan pertaksamaan	Menyelesaikan persamaan dan pertaksamaan kuadrat
		Menyelesaikan sistem persamaan linear dua atau tiga variabel
Geometri dan Pengukuran	Bangun geometri	Memahami barisan aritmatika dan geometri
		Memahami fungsi kuadrat dan grafiknya, serta sifat-sifatnya
Data dan ketidakpastian	Data dan representasinya	Menggunakan perbandingan trigonometri
		Menghitung volume dan luas permukaan limas segi- $n$ , kerucut, dan bola
Data dan ketidakpastian	Ketidakpastian dan peluang	Menginterpretasikan penyajian data
		Menentukan ukuran penyebaran data (jangkauan, simpangan, dan variansi)
		Menggunakan sifat-sifat peluang suatu kejadian
		Menggunakan sifat-sifat peluang kejadian majemuk

Tabel 14 menginformasikan bahwa pada pelaksanaan AN 2021 untuk AKM numerasi kelas XI SMA/SMK/Sederajat memiliki tiga domain yang diuji yakni aljabar, geometri dan pengukuran, serta data dan ketidakpastian. Kemudian masing-masing dari domain tersebut memiliki subdomain hingga kompetensi-kompetensi seperti yang ada pada tabel di atas. Oleh karena itu, soal AKM numerasi yang akan dikembangkan pada penelitian ini mengacu pada kompetensi-kompetensi tersebut. Selanjutnya dilakukan penetapan sebanyak 4 paket (A/B/C/D) soal AKM numerasi yang dikembangkan. Dimana masing-masing paket terdiri atas 16-17 item yang didalamnya terdapat 4 item *anchor*. Jumlah item *anchor* ditetapkan dan disesuaikan dengan jumlah minimal yang disyaratkan yakni 20% dari total item tes serta aspek ukur yang terwakilkan (Supahar & Prasetyo, 2015, p. 106). Adanya item *anchor* dalam pengembangan soal ini sebagai penghubung antar paket soal saat dilakukan analisis kualitas item soal secara bersama-sama atau dengan kata lain bermanfaat dalam analisis penyetaraan (Hayati & Mardapi, 2014, p. 31; Santoso et al., 2019, p. 168).

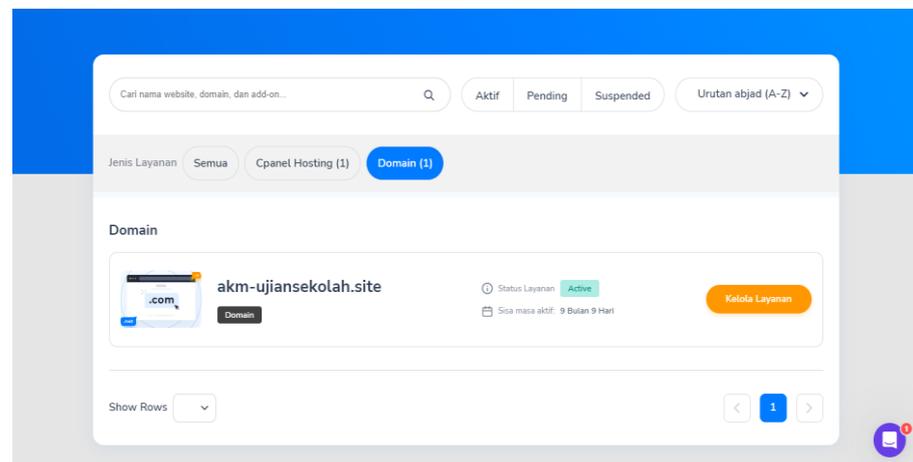
Waktu pengerjaan selama 2 jam pelajaran ( $2 \times 45$  menit), mengikuti jumlah jam pelajaran yang disediakan dalam satu hari untuk pembelajaran matematika SMA, serta mengingat bahwa subjek penelitian ini adalah peserta didik kelas XI SMA. Kemudian level kognitif yang dipergunakan dalam pengembangan soal ada tiga yakni pengetahuan (*knowing*), penerapan (*applying*), dan penalaran (*reasoning*). Sedangkan konteks yang dipergunakan juga ada tiga yakni personal, saintifik, dan sosial budaya.

Tipe jawaban yang digunakan dalam pengembangan ini menyesuaikan bentuk soal yang digunakan, dimana untuk pilihan ganda, pilihan ganda kompleks, dan menjodohkan tipe jawabannya dengan pilihan jawaban benar yang sudah tersedia, sedangkan tipe jawaban isian singkat dan uraian menggunakan tipe jawaban benar

dengan isian singkat dan uraian atau dengan kata lain tidak disediakan pilihan jawaban benar oleh pengembang. Kemudian berkaitan dengan penskoran dan pemaknaan hasil, soal-soal yang dikembangkan terbagi menjadi dua penyekoran, mengingat bentuk soal yang digunakan lebih dari satu. Penyekoran dikotomi dimana jawaban benar bernilai 1 dan jawaban salah bernilai 0 untuk soal pilihan ganda dan isian singkat. Sedangkan penyekoran politomi untuk soal pilihan ganda kompleks, menjodohkan, dan uraian dimana jawaban benar dan sesuai dengan kriteria kunci jawaban bernilai 2, jawaban benar namun kurang memenuhi kriteria bernilai 1, dan jawaban salah bernilai 0.

Selain menentukan cakupan materi hingga penetapan jenis penskoran yang digunakan dalam penelitian pengembangan ini, peneliti juga melakukan *hosting* sehingga diperoleh laman [www.akm-ujiansekolah.site](http://www.akm-ujiansekolah.site) yang kemudian dipergunakan untuk pembuatan media CBT berbantuan LMS Moodle. Tampilan domain dari *hosting* yang dilakukan dapat dilihat pada Gambar 1.

**Gambar 1.** Tampilan Domain dari Hosting



### c. Hasil Menyusun Indikator dan Matriks Instrumen Soal

#### 1) Penyusunan Indikator dan Matriks Instrumen Soal

Indikator item soal disusun dan disesuaikan dengan kompetensi-kompetensi AKM numerasi seperti pada Tabel 13 dan menggunakan kata kerja operasional sehingga dapat diukur. Berikut adalah cuplikan

indikator item soal dari kisi-kisi pengembangan soal AKM numerasi. Untuk lebih lengkapnya terlampir pada Lampiran 1B.

**Gambar 2.** *Cuplikan Kisi-kisi Instrumen Soal Tahap Awal*

Kompetensi	Materi	Indikator Soal	Kode Soal	Bentuk Soal
Menyelesaikan persamaan dan pertidaksamaan kuadrat	Persamaan kuadrat	Diberikan suatu persamaan kuadrat, peserta didik menafsirkan akar-akar persamaan kuadrat untuk menyelesaikan masalah yang tersajikan	SOAL01	Pilihan ganda kompleks
		Diberikan suatu persamaan kuadrat, peserta didik menyatakan nilai dari salah satu koefisien persamaan kuadrat yang hilang	SOAL02	Pilihan ganda
	Pertaksamaan kuadrat	Diberikan pertaksamaan bentuk kuadrat, peserta didik menentukan batas nilai dari pertaksamaan tersebut	SOAL03	Isian singkat
		Diberikan konteks masalah yang berkaitan dengan pertaksamaan kuadrat, peserta didik menyatakan himpunan penyelesaiannya	SOAL04	Pilihan ganda
Menyelesaikan sistem persamaan linear dua atau tiga variabel	SPLDV	Diberikan masalah kontekstual yang berkaitan dengan sistem pertidaksamaan linear dua variabel, peserta didik menentukan model matematika dari masalah tersebut	SOAL05	Pilihan ganda
	SPLTV	Diberikan masalah kontekstual yang berkaitan dengan sistem persamaan linear tiga variabel, peserta didik menyimpulkan hasil dari ketiga variabel yang diperoleh	SOAL06	Menjodohkan

Gambar 2 di atas menginformasikan bahwa pada tahap awal kisi-kisi dibuat hanya terdiri atas kompetensi, materi, indikator soal, kode soal, dan bentuk soal. Indikator soal disusun sebanyak 54 indikator dengan indikator soal domain aljabar sebanyak 14 indikator, domain geometri dan pengukuran sebanyak 18 indikator, serta domain data dan ketidakpastian sebanyak 22 indikator. Selain itu, proporsi bentuk soal yang didapat setelah pembentukan kisi-kisi tahap awal ini yakni soal pilihan ganda sebanyak 12 item, pilihan ganda kompleks 12 item, menjodohkan 12 item, isian singkat 12 item, dan uraian 6 item. Sedangkan matriks instrumen soal yang telah disusun dapat dilihat pada Lampiran 1A.

## 2) Penyusunan *Storyboard* Media CBT

Media CBT yang digunakan dalam penelitian ini berbasis *web* berbantuan LMS Moodle yang menggunakan bahasa pemrograman *Hypertext Preprocessor* (PHP) dan sistem basis data *Structured Query Language* (SQL) (Haręzlak & Werner, 2012, p. 52; Luna et al., 2017, p. 90). Setelah menyusun indikator item soal peneliti juga menyusun *storyboard* sebelum media CBT dikembangkan. Berikut adalah *storyboard* yang dihasilkan pada langkah ini.

## Halaman Utama

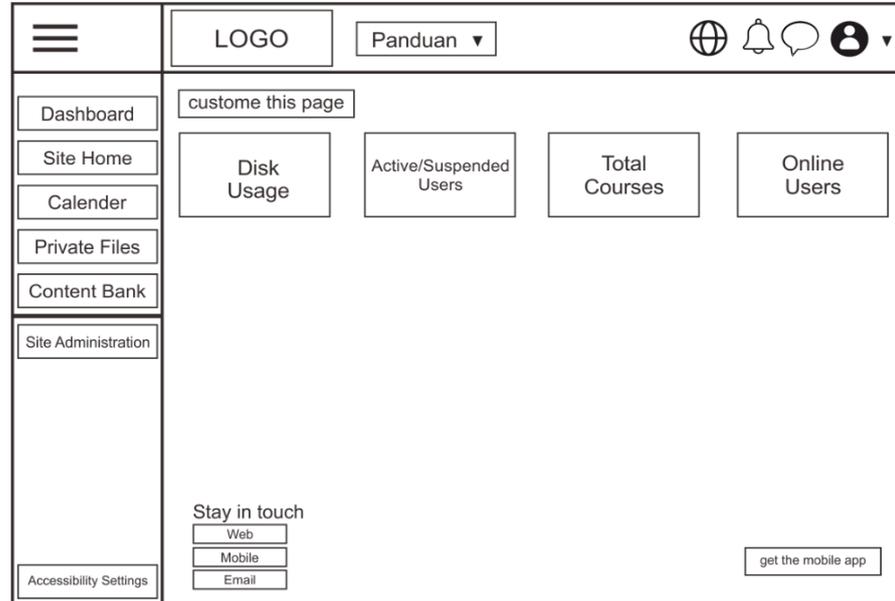
**Gambar 3.** *Storyboard Tampilan Halaman Utama*

The storyboard shows a rectangular frame representing the main page. At the top left, there is a box labeled 'LOGO'. To its right is a box labeled 'Panduan' with a downward-pointing triangle. At the top right is a globe icon. The main content area is divided into two sections. On the left, the text reads 'TRY OUT' in large bold letters, followed by 'Asesmen Kompetensi Minimum Numerasi' in smaller text. On the right, there is a box titled 'access to the platform' containing four input fields: 'username', 'password', 'log in', and 'new account'. Below the 'log in' field is the text 'forgotten your username or password?'. At the bottom left, there is a section titled 'Stay in touch' with three stacked boxes labeled 'Web', 'Mobile', and 'Email'. At the bottom right, there is a box labeled 'get the mobile app'.

Halaman utama media CBT mencakup logo, panduan (pendidik dan peserta didik), pengaturan bahasa, *background* selamat datang di laman media CBT dengan tulisan “Try Out Asesmen Kompetensi Minimum Numerasi”, kolom *username* dan *password* yang harus diisi pengguna ketika akan mengakses soal AKM numerasi berbasis komputer. Terdapat tiga jenis pengguna dalam penelitian pengembangan ini yakni administrator, pendidik, dan peserta didik.

## Halaman Administrator

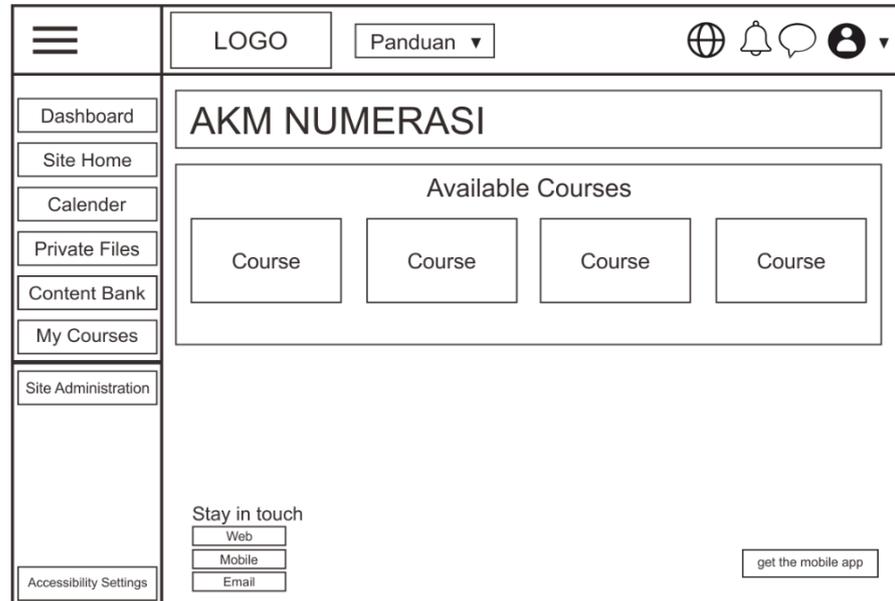
**Gambar 4.** *Storyboard Tampilan Halaman Administrator*



Halaman administrator merupakan halaman yang dimiliki secara penuh oleh admin yang dapat mengakses segala fitur yang ada pada media CBT. Dilihat dari Gambar 4, halaman administrator memiliki 6 menu utama yakni *dashboard*, *site home*, *calender*, *private files*, *content bank*, dan *site administration*. Bagian *header* selain terdapat logo dan panduan (pendidik dan peserta didik) juga terdapat pengaturan bahasa (Bahasa Indonesia dan Bahasa Inggris), notifikasi, pesan, serta informasi pengguna. Sedangkan pada *footer* terdapat informasi mengenai laman *web* pengujian AKM numerasi, nomor telpon dan email administrator. Hal penting yang membedakan akses administrator dengan pengguna lain salah satunya yakni admin dapat mengetahui data terkait *disk usage*, *active/suspended users*, *total courses*, dan *total online users*.

## Halaman Pendidik

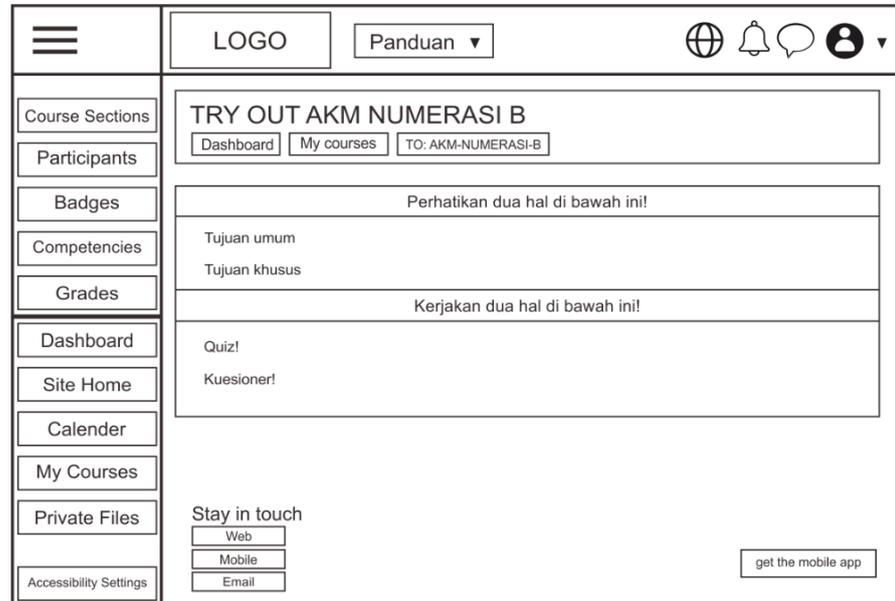
**Gambar 5.** *Storyboard Tampilan Halaman Pendidik*



Halaman pendidik merupakan halaman yang dimiliki oleh para pendidik yang sudah diberi akses oleh administrator. Terkait adanya *courses* pendidik tidak dapat membuat *courses* namun bisa mengedit apapun yang ada pada *course*. Gambar 5 menunjukkan bahwa tampilan pada halaman pendidik terdapat menu *dashboard*, *site home*, *calender*, *private files*, *content bank*, *my courses*, dan *site administration*.

## Halaman Peserta Didik

**Gambar 6.** Storyboard Tampilan Halaman Peserta Didik



Halaman peserta didik merupakan halaman yang dimiliki peserta didik, dimana hanya dapat digunakan untuk mengerjakan tes yang telah disediakan oleh admin dan pendidik. Pada LMS Moodle peserta didik dapat diberi akses untuk melihat skor yang diperoleh. Asal sistem penskoran sudah diatur admin atau pendidik. Gambar 6 menunjukkan bahwa tampilan pada halaman peserta didik mencakup menu *dashboard*, *site home*, *calender*, *my courses*, dan *private files*. Pada menu *my courses* berisi *course sections*, *participants*, *badges*, *competencies*, dan *grades*.

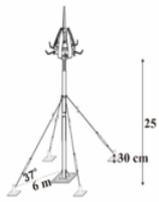
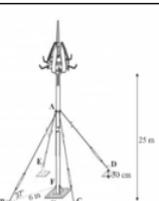
### d. Hasil Menyusun Item Soal

#### 1) Penyusunan Item Soal

Penyusunan item soal didasarkan atas kisi-kisi instrumen yang telah tersusun di langkah sebelumnya. Sebanyak 54 soal telah dikembangkan dengan memperhatikan pedoman penulisan soal yang baik. Pada langkah ini dibuat kartu soal untuk tiap item yang mencakup informasi mengenai indikator soal, item soal, kunci jawaban hingga komponen-komponen lain yang memiliki keterkaitan dengan item soal. Dengan adanya kartu soal ini diharapkan dapat

mempermudah proses validasi para ahli. Berikut sampel kartu soal yang telah disusun pada langkah penyusunan item soal.

**Gambar 7. Sampel Kartu Soal AKM Numerasi**

Jenjang Pendidikan : SMA Kelas/Semester : XI/Genap		Tahun Ajaran : 2022/2023 Penyusun : Ihda Mutimmatul Fitriyah, S.Pd.	
<b>Kode Soal</b> SOAL16	<b>Nomor Soal</b> <b>1</b> PAKET A	<b>Item Soal</b> <b>Tiang Penyangga Penangkal Petir</b> Tiang penyangga penangkal petir merupakan pipa penagak yang berfungsi untuk menaikkan posisi <i>head</i> terminal penangkal petir elektrostatik agar mendapatkan level yang tinggi, karena semakin tinggi level penangkal petir maka sangat berarti untuk mendapatkan radius perlindungan ( <i>Sumber: pakarpetir.co.id</i> ). Tiang penyangga penangkal petir memiliki 4 kawat penyangga yang dipasang simetris pada tiang utamanya, dan tampak pada gambar berikut. Panjang kawat yang diperlukan untuk menyangga tiang adalah .... (Keterangan: $\sin 37^\circ = \frac{3}{5}$ ; $\cos 37^\circ = \frac{4}{5}$ ; dan $\tan 37^\circ = \frac{3}{4}$ ) A. 14,4 m B. 19,2 m C. 30 m D. 32 m E. 40 m	
<b>Domain Numerasi</b> Geometri dan pengukuran	<b>Bentuk Soal</b> Pilihan ganda		
<b>Subdomain Numerasi</b> Bangun geometri	<b>Konteks</b> Saintifik	<b>Kunci Jawaban</b> Ditanya terkait total kawat yang diperlukan untuk menyangga tiang sama halnya dengan total panjang AB, AC, AD, dan AE dimana keempatnya memiliki nilai yang sama. Apabila ABF merupakan segitiga siku-siku maka: $\cos 37^\circ = \frac{BF}{AB}$	
<b>Kompetensi</b> Menggunakan perbandingan trigonometri pada masalah kontekstual	<b>Level Kognitif</b> Knowing		
<b>Indikator Soal</b> Diberikan suatu masalah kontekstual yang merepresentasikan bangun segitiga siku-siku, peserta didik menentukan sisi miring segitiga dengan rasio trigonometri	<b>Kesukaran</b> Mudah	 <p><math>\frac{4}{5} = \frac{6}{AB}</math> <math>AB = 7,5 m</math> Panjang AB = AC = AD = AE = 7,5 m Jadi, total kawat yang diperlukan untuk menyangga tiang adalah <math>4 \times 7,5 = 30 m</math> (C)</p> <p><b>Pengecoh,</b> Opsinya A. 14,4 m berfungsi menjadi pengecoh dimana peserta didik salah menggunakan konsep rasio trigonometri: <math>\sin 37^\circ = \frac{AB}{BF}</math> <math>\frac{3}{5} = \frac{AB}{6} \rightarrow AB = 3,6 m</math> sehingga total kawat yang diperlukan <math>4 \times 3,6 = 14,4 m</math> Opsinya B. 19,2 m berfungsi menjadi pengecoh dimana peserta didik salah menggunakan konsep rasio trigonometri: <math>\cos 37^\circ = \frac{AB}{BF}</math> <math>\frac{4}{5} = \frac{AB}{6} \rightarrow AB = 4,8 m</math> sehingga total kawat yang diperlukan <math>4 \times 4,8 = 19,2 m</math> Opsinya D. 32 m berfungsi menjadi pengecoh dimana peserta didik salah menggunakan konsep rasio trigonometri: <math>\tan 37^\circ = \frac{BF}{AB}</math> <math>\frac{3}{4} = \frac{6}{AB} \rightarrow AB = 8 m</math> sehingga total kawat yang diperlukan <math>4 \times 8 = 32 m</math></p>	
		Opsinya E. 40 m berfungsi menjadi pengecoh dimana peserta didik salah menggunakan konsep rasio trigonometri: $\sin 37^\circ = \frac{BF}{AB}$ $\frac{3}{5} = \frac{6}{AB} \rightarrow AB = 10 m$ sehingga total kawat yang diperlukan $4 \times 10 = 40 m$	

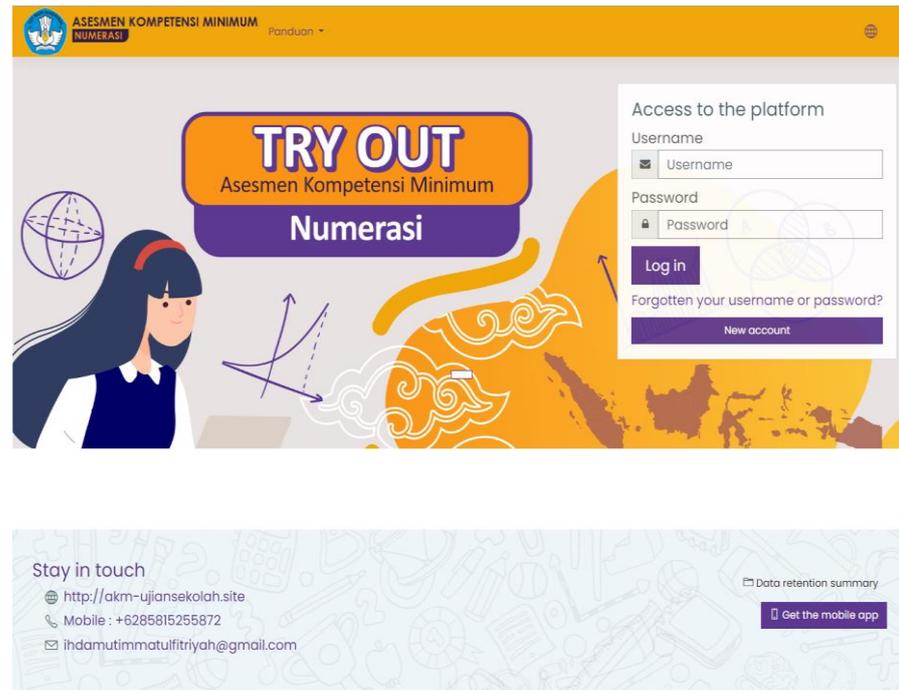
## 2) Pengimplementasian Storyboard Media CBT

Media CBT untuk pengujian soal AKM numerasi dapat diakses pada laman [www.akm-ujiansekolah.site](http://www.akm-ujiansekolah.site) kapanpun dan dimanapun.

Bisa diakses dengan laptop, tab, ataupun *handphone* pengguna. *Storyboard* yang telah disusun kemudian diimplementasikan sehingga dihasilkan tampilan media CBT seperti berikut.

### Halaman Utama

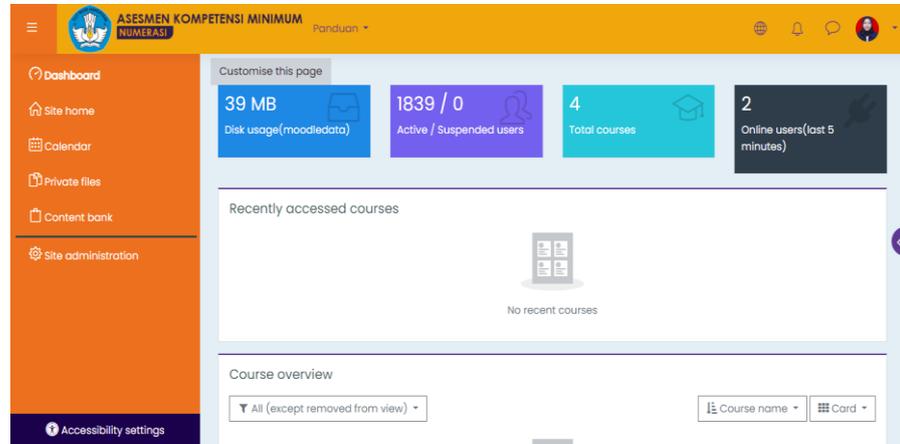
**Gambar 8.** *Tampilan Halaman Utama*



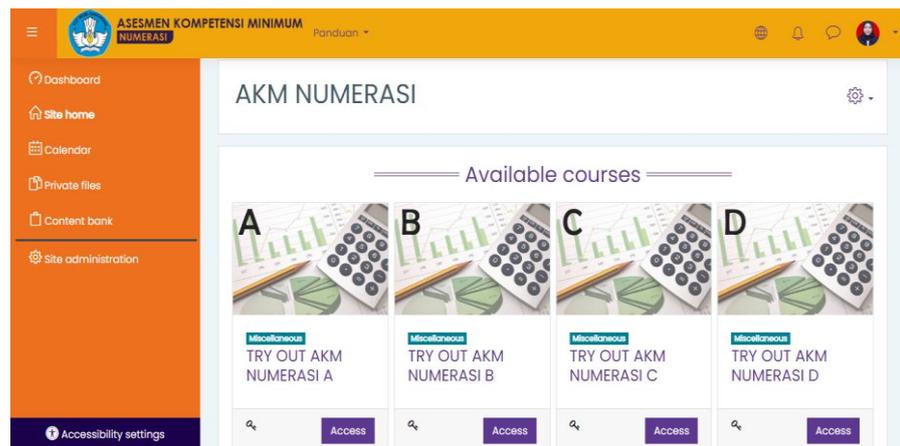
*Storyboard* halaman utaman yang telah dirancang di langkah sebelumnya kemudian diimplementasikan sehingga menghasilkan halaman utama yang nampak seperti Gambar 8. Terdapat logo, panduan, pengaturan bahasa, *background* tulisan “Try Out Asesmen Kompetensi Minimum Numerasi”, kolom *username*, *password*, serta informasi narahubung kepemilikan media CBT. Halaman utama merupakan halaman awal yang ditampilkan ketika pengguna mengakses laman *web* pengujian.

## Halaman Administrator

**Gambar 9.** *Tampilan Menu Dashboard*



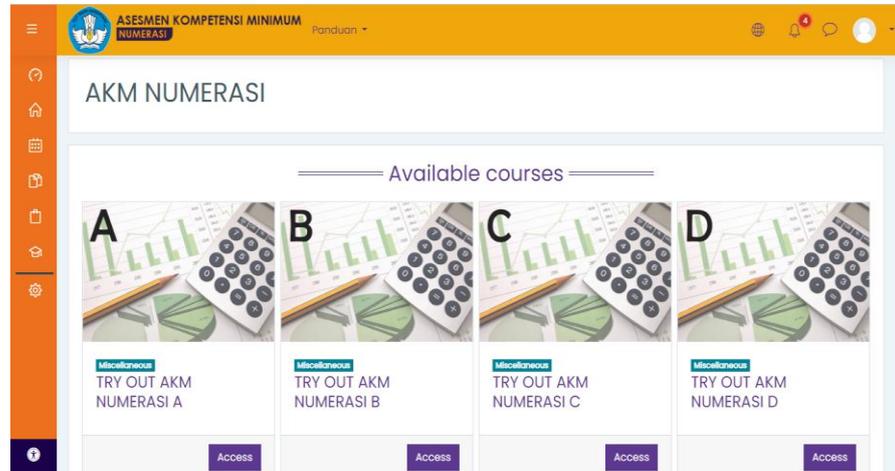
**Gambar 10.** *Tampilan Menu Site Home*



Gambar 9 merupakan salah satu tampilan yang hanya bisa dimiliki oleh administrator, terkait seberapa banyak data yang sudah digunakan, total pengguna aktif atau ditangguhkan, jumlah *course*, serta berapa banyak pengguna yang sedang *online*. Sedangkan Gambar 10 menunjukkan jumlah *course* yang dapat diakses para pengguna. Terkait pembuatan *course* hanya administrator yang dapat menambahkan atau menghapus *course*. Selain itu hanya administrator yang dapat mengatur pendidik dan peserta didik untuk dapat mengakses *course* yang telah dibuat.

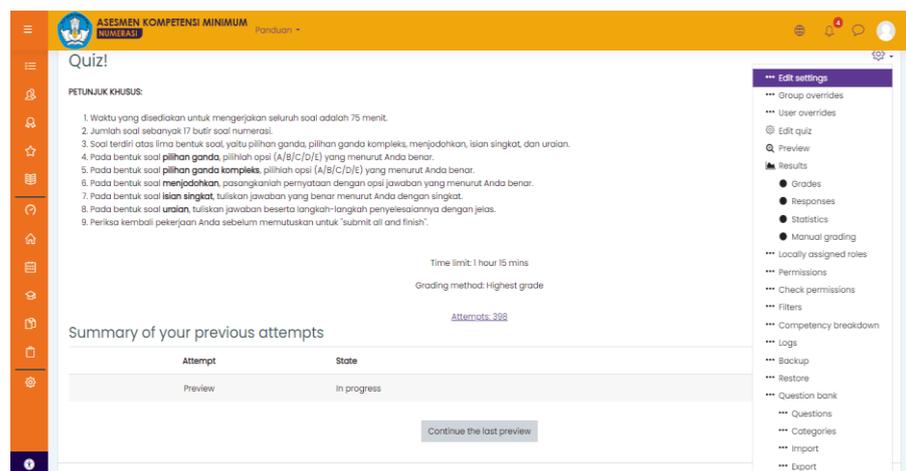
## Halaman Pendidik

**Gambar 11.** Tampilan Menu Site Home Pendidik

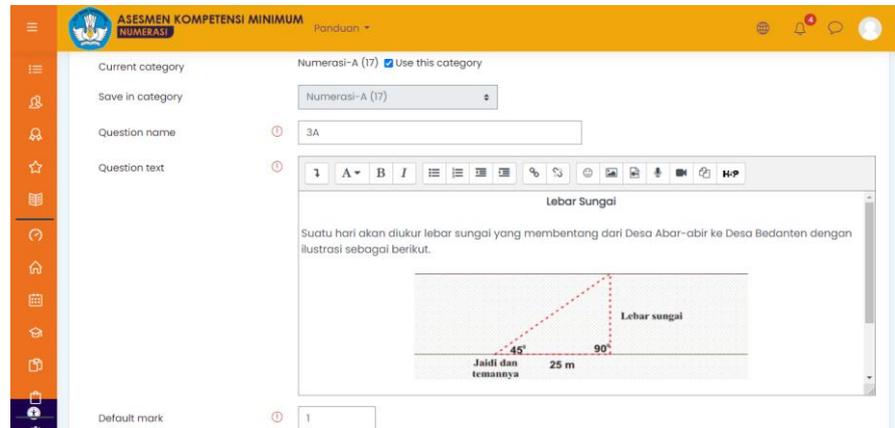


Gambar 11 merupakan tampilan menu *site home* yang berisi *course-course* yang dapat diakses pendidik. Hak akses *course-course* ini administrator berikan kepada pendidik. Untuk penelitian ini administrator memberikan akses sepenuhnya di keempat *course* yakni AKM numerasi paket A hingga D.

**Gambar 12.** Tampilan Edit Quiz



**Gambar 13.** *Tampilan Edit Soal*



Gambar 12 dan Gambar 13 menunjukkan keuntungan yang didapat dari pendidik dimana dapat mengubah pengaturan di dalam *course*, dalam penelitian ini terdapat menu Quiz yang ada dalam *course*. Sehingga pendidik dapat mengubah teks soal numerasi jika ada yang salah ketik atau mengubah soal.

### **Halaman Peserta Didik**

**Gambar 14.** *Tampilan Aktivitas dalam Course*

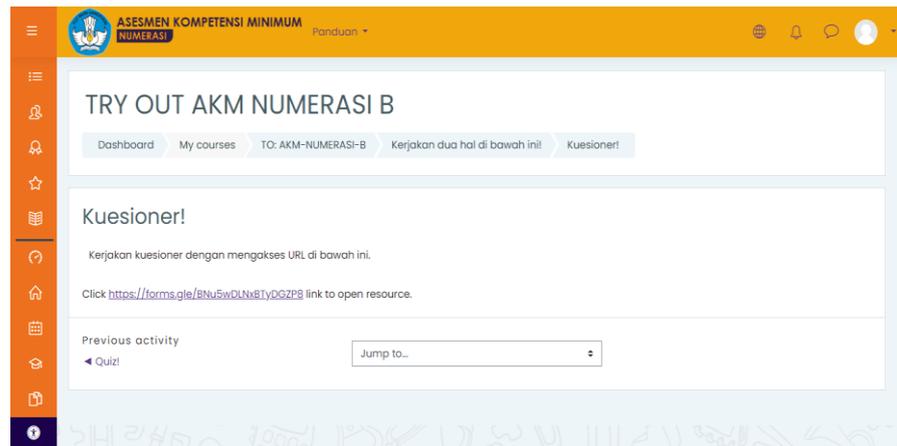


Gambar 14 menunjukkan aktivitas yang ada dalam *course* untuk penelitian ini yakni Quiz dan Kuesioner.

**Gambar 15.** Tampilan Soal Numerasi dalam Aktivitas Quiz

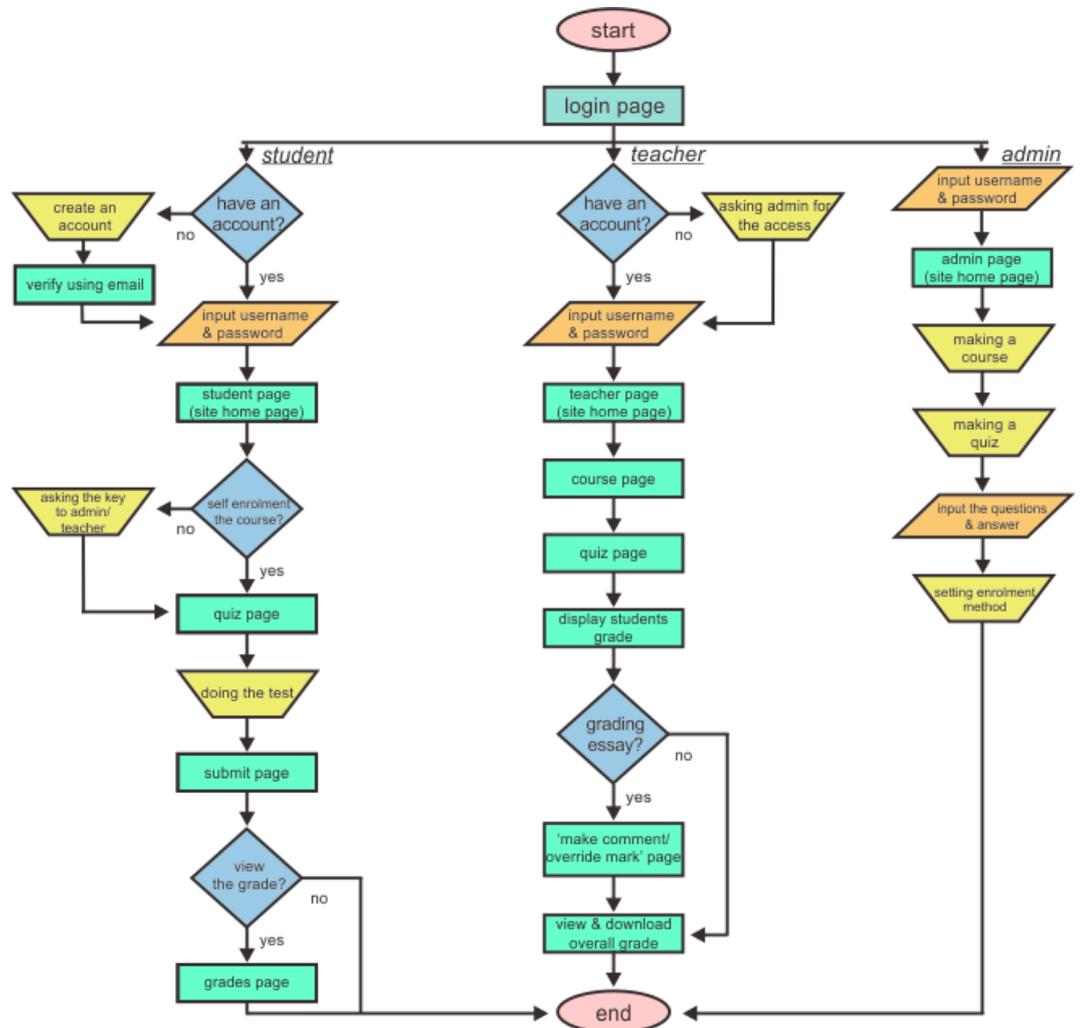


**Gambar 16.** Tampilan dalam Aktivitas Kuesioner



Gambar 15 dan Gambar 16 menunjukkan isi di dalam aktivitas Quiz dan Kuesioner, yang mana keduanya harus diselesaikan peserta didik dalam durasi pengujian. Setelah pengimplementasian *storyboard* sehingga menjadi media CBT yang utuh maka langkah selanjutnya adalah proses validasi kepada para ahli. Terakhir untuk *flowchart* alur LMS Moodle dapat dilihat pada Gambar 17.

**Gambar 17.** Flowchart LMS Moodle AKM Numerasi



**e. Hasil Validasi Isi**

Pembuktian validitas isi merupakan langkah selanjutnya yang dilakukan melalui para ahli atau *expert judgment*. Pada langkah ini validasi terjadi dua kali yakni pembuktian validasi isi soal AKM numerasi dan validasi isi media CBT.

**1) Validasi Isi Soal AKM Numerasi**

Validasi isi soal AKM numerasi dilakukan oleh 6 ahli yakni dua dosen ahli pengukuran, satu dosen ahli materi, dan tiga pendidik matematika SMA yang dapat dilihat pada Tabel 15 berikut.

**Tabel 15.** *Nama-nama Validator Isi Soal AKM Numerasi*

No	Nama Validator	Spesifikasi Keahlian
1	Dr. Syukrul Hamdi, M.Pd.	Ahli Pengukuran
2	Dr. Kana Hidayati, M.Pd.	Ahli Pengukuran
3	Dr. Siti Lailiyah, M.Si.	Ahli Materi
4	Risa Utaminingsih, M.Pd.	Praktisi/Pendidik Matematika
5	Itsna Lailatul Mas'udah, M.Pd	Praktisi/Pendidik Matematika
6	Durrotun Nabilah, M.Pd	Praktisi/Pendidik Matematika

Proses validasi meliputi telaah kisi-kisi instrumen soal numerasi, alternatif jawaban hingga rubrik penskorannya. Secara kualitatif, para validator memberi masukan dan saran baik secara keseluruhan atau pada masing-masing item soal. Masukan dan saran yang diperoleh dipergunakan peneliti untuk melakukan revisi terhadap soal AKM numerasi. Validator juga memberikan skor sebagai penilaian secara kuantitatif untuk validasi ini, karena dalam langkah ini diperlukan pengestimasian indeks validitas. Hasil validasi isi dari masing-masing validator terlampir dalam Lampiran 2A. Hasil penskoran yang diperoleh terangkum dalam indeks validitas pada Tabel 16 yang diestimasi dengan formula Aiken's V, sedangkan kriteria ditetapkan berdasarkan banyaknya kategori penilaian (penelitian ini terdapat 4 kategori jawaban) dan banyaknya validator (terdapat 6 validator) sehingga untuk tingkat signifikansi 5% maka kriteria item soal dikatakan valid sebesar 0,78 (Aiken, 1985, p. 134). Berdasarkan kriteria kevalidan tersebut untuk validasi isi penelitian ini diperoleh seluruh item soal yang valid karena memiliki Aiken's V paling rendah 0,78.

**Tabel 16.** *Perhitungan Hasil Aiken's V*

No Soal	Kode Soal	Aiken's V	Kriteria	No Soal	Kode Soal	Aiken's V	Kriteria
1A	SOAL16	0,83	Valid	15B	SOAL25	0,89	Valid
2A	SOAL02	0,83	Valid	16B	SOAL15	0,83	Valid
3A	SOAL18	0,78	Valid	1C	SOAL05	0,94	Valid
4A	SOAL41	0,94	Valid	2C	SOAL40	0,89	Valid
5A	SOAL45	0,78	Valid	3C	SOAL10	0,89	Valid
6A	SOAL11	0,94	Valid	4C	SOAL42	0,89	Valid
7A	SOAL33	0,94	Valid	5C	SOAL01	0,89	Valid
8A	SOAL52	0,89	Valid	6C	SOAL39	0,89	Valid
9A	SOAL49	0,83	Valid	7C	SOAL53	0,89	Valid
10	SOAL12	0,89	Valid	8C	SOAL17	0,89	Valid
11	SOAL31	0,83	Valid	9C	SOAL44	0,94	Valid
12	SOAL24	0,83	Valid	14C	SOAL47	0,94	Valid
13	SOAL03	0,94	Valid	15C	SOAL30	0,89	Valid
14A	SOAL21	0,94	Valid	16C	SOAL19	0,83	Valid
15A	SOAL43	0,89	Valid	17C	SOAL29	0,89	Valid
16A	SOAL20	0,94	Valid	1D	SOAL51	1,00	Valid
17A	SOAL08	0,89	Valid	2D	SOAL26	0,89	Valid
1B	SOAL22	0,94	Valid	3D	SOAL23	0,89	Valid
2B	SOAL04	0,83	Valid	4D	SOAL34	0,89	Valid
3B	SOAL13	1,00	Valid	5D	SOAL27	0,83	Valid
4B	SOAL32	0,83	Valid	6D	SOAL09	0,83	Valid
5B	SOAL54	0,94	Valid	7D	SOAL28	0,78	Valid
6B	SOAL48	0,83	Valid	8D	SOAL46	0,94	Valid
7B	SOAL37	0,94	Valid	9D	SOAL35	0,94	Valid
8B	SOAL06	0,94	Valid	14D	SOAL50	1,00	Valid
9B	SOAL36	0,94	Valid	15D	SOAL14	0,94	Valid
14B	SOAL38	1,00	Valid	16D	SOAL07	1,00	Valid

## 2) Validasi Media CBT

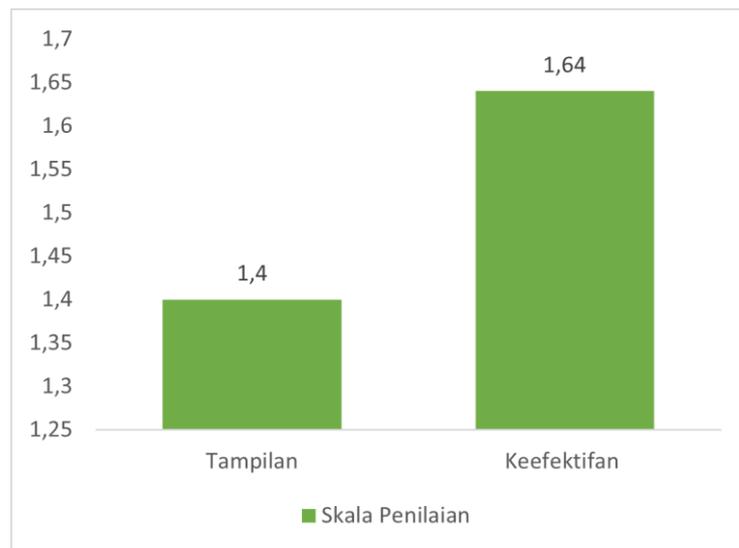
Media CBT yang telah disiapkan dalam penelitian ini divalidasi oleh dua ahli yakni dua dosen yang ahli media/IT serta pengukuran yang dapat dilihat pada Tabel 17 berikut.

**Tabel 17.** Nama-nama Validator Media CBT

No	Nama Validator	Spesifikasi Keahlian
1	Prof. Ir. Moh. Khairudin, M.T., Ph.D., IPU., ASEAN. Eng.	Ahli Media/IT
2	Dr. Risky Setiawan, M.Pd.	Ahli Media & Pengukuran

Hasil validasi media CBT diperoleh dari rata-rata skor pada masing-masing aspek penilaian yang dapat dilihat pada Gambar 18.

**Gambar 18.** Hasil Validasi Media CBT



Gambar 18 diperoleh rata-rata skor penilaian sebesar 1,40 untuk aspek tampilan dengan kategori cukup layak, dan pada aspek keefektifan penggunaan skor rata-rata sebesar 1,64 dengan kategori cukup layak. Secara keseluruhan penilaian media CBT diperoleh 1,52 dengan kategori cukup layak. Hasil validasi media dari masing-masing validator terlampir dalam Lampiran 2B. Berdasarkan hasil validasi media CBT ini dapat ditarik kesimpulan bahwa media CBT yang

mencakup pengujian soal AKM numerasi yang telah peneliti kembangkan layak digunakan namun dengan beberapa revisi.

**f. Hasil Revisi Berdasarkan Masukan Validator**

Hasil penelaahan soal AKM numerasi dan keberlayakan media CBT dari para validator saat validasi menunjukkan bahwa instrumen soal AKM numerasi yang telah dikembangkan serta media CBT yang akan dipergunakan perlu dilakukan perbaikan. Perbaikan dilakukan berdasarkan masukan dan saran baik yang disampaikan langsung ke peneliti saat bimbingan validasi dan atau tertulis pada surat keterangan validasi. Masukan dan saran yang diberikan validator berkaitan dengan soal AKM numerasi meliputi: (1) penggunaan dan konsistensi istilah yang kurang lazim digunakan pada perangkat soal, (2) kurangnya ilustrasi tambahan dan kebermaknaan stimulus soal, (3) beberapa soal yang petunjuk atau intruksinya kurang jelas, (4) tata kalimat yang kurang sesuai EYD, dan (5) perlu peninjauan ulang kesesuaian item soal dengan konteks soal dan kesukaran yang ditetapkan. Sedangkan masukan dan saran yang diberikan validator berkaitan dengan media CBT meliputi: (1) penyesuaian kembali pemilihan *background* dengan tema, (2) pengadaan petunjuk media tambahan untuk pendidik dan administrator, (3) penegasan tujuan tes dan keterangan petunjuk pengisian Quiz dan Kuesioner yang ada pada media, dan (4) perbaikan resolusi gambar yang tertera dalam soal. Setelah dilakukan perbaikan, peneliti mengkonsultasikan kembali ke validator guna memperoleh instrumen soal AKM numerasi yang valid serta media CBT yang cukup layak digunakan.

**g. Hasil Melakukan Uji Coba**

Soal AKM numerasi yang telah tervalidasi kemudian disusun menjadi empat paket soal, dimana paket A dan C berisi masing-masing 17 item soal serta paket B dan D berisi masing-masing 16 soal. Uji coba soal AKM numerasi berbasis komputer dilakukan di kelas XI dari 6 Sekolah Menengah Atas Negeri (SMAN) di Kota Yogyakarta

meliputi SMAN 1 Yogyakarta, SMAN 2 Yogyakarta, SMAN 3 Yogyakarta, SMAN 5 Yogyakarta, SMAN 6 Yogyakarta, dan SMAN 11 Yogyakarta. Subjek uji coba untuk paket A sebanyak 392 peserta didik, paket B sebanyak 395 peserta didik, paket C sebanyak 386 peserta didik, dan paket D sebanyak 397 peserta didik. Rincian sampel pada masing-masing sekolah tercantum pada Tabel 18.

**Tabel 18.** *Subjek Uji Coba Penelitian*

No	Sekolah	Jumlah Subjek				Total
		Paket	Paket	Paket	Paket	
		A	B	C	D	
1	SMAN 1 Yogyakarta	79	80	81	81	321
2	SMAN 2 Yogyakarta	74	78	74	74	300
3	SMAN 3 Yogyakarta	61	56	54	61	232
4	SMAN 5 Yogyakarta	64	66	61	66	257
5	SMAN 6 Yogyakarta	54	52	54	54	214
6	SMAN 11 Yogyakarta	60	63	62	61	246
	<b>Total</b>	<b>392</b>	<b>395</b>	<b>386</b>	<b>397</b>	<b>1570</b>

Selain melakukan uji coba soal AKM numerasi berbasis komputer pada peserta didik, para pendidik yang mengajar matematika di kelas XI dari keenam sekolah juga dijadikan subjek penelitian untuk memperoleh data terkait keefektifan soal AKM numerasi berbasis komputer yang telah dikembangkan. Jumlah pendidik yang dijadikan subjek penelitian yakni sebanyak 9 orang.

#### **h. Hasil Melakukan Analisis**

Hasil uji coba pada penelitian ini berupa respon peserta didik terhadap 16-17 item soal pada masing-masing paket serta respon pendidik terhadap angket keefektifan soal AKM numerasi berbasis komputer yang telah dikembangkan. Selanjutnya, peneliti melakukan penskoran pada respon peserta didik dan pendidik untuk dilakukan

analisis menggunakan pendekatan Teori Respon Butir (TRB) berbantuan aplikasi program R, SPSS, dan *Microsoft Excel*. Analisis dalam penelitian ini mencakup analisis validitas konstruk, estimasi reliabilitas, analisis butir soal, analisis keefektifan produk. Lebih lanjut hasil analisis disajikan dalam bagian berikutnya.

**i. Hasil Merakit Instrumen Soal**

Hasil validasi dan analisis kualitas butir soal menunjukkan bahwa keseluruhan soal AKM numerasi berbasis komputer yang telah dikembangkan memiliki karakteristik yang baik sehingga selanjutnya dilakukan perakitan soal kembali sebagai tes yang siap digunakan untuk mengumpulkan data penelitian. Perakitan soal AKM numerasi berbasis komputer tidak berbeda jauh dengan penyusunan soal pada langkah sebelum kegiatan uji coba. Hal tersebut dikarenakan seluruh soal yang dikembangkan memenuhi kecocokan dengan model serta memiliki validitas, reliabilitas dan karakteristik butir yang baik.

**2. Hasil Uji Coba Produk**

**a. Uji Validitas Konstruk**

Analisis validitas konstruk dilakukan melalui data respon hasil uji coba dengan *Confirmatory Factor Analysis* (CFA). CFA bersifat mengkonfirmasi kesesuaian model instrumen dengan data empiris di lapangan. Pembuktian kesesuaian model dengan CFA didasarkan pada kriteria *Goodness of Fit Test* (GOF). Nilai GOF yang dihasilkan dari instrumen yang dikembangkan diestimasi dengan bantuan aplikasi program R dan dapat dilihat pada Gambar 19.

**Gambar 19.** Nilai GOF Instrumen Soal AKM Numerasi

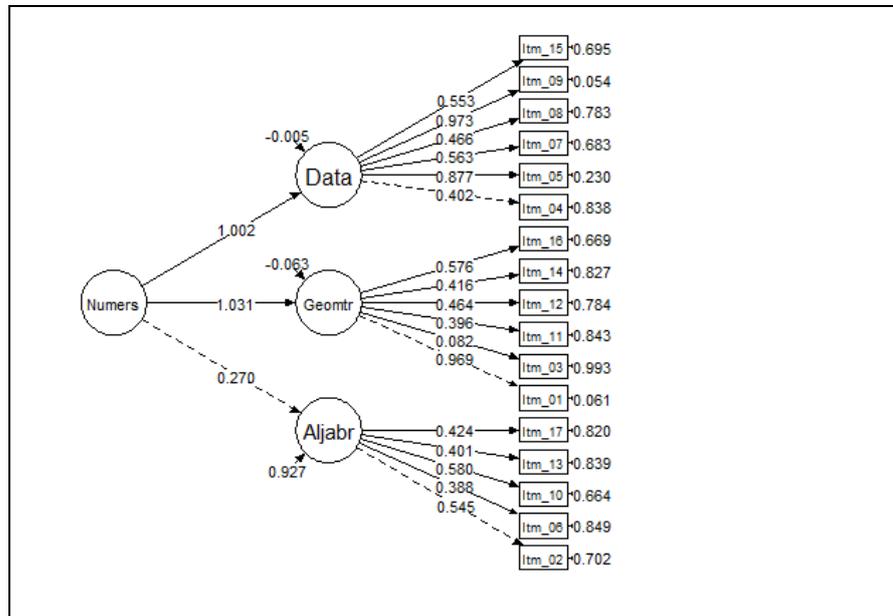
CFA_A			CFA_B		
Ukuran	GOF	Hasil Keterangan	Ukuran	GOF	Hasil Keterangan
Chi-square	764.320	Model Fit	Chi-square	977.310	Model Fit
df	485.000	Model Fit	df	551.000	Model Fit
p-value	0.059	Model Fit	p-value	0.071	Model Fit
RSMEA	0.051	Model Fit	RSMEA	0.041	Model Fit
GFI	0.960	Model Fit	GFI	0.920	Model Fit
RMR	0.032	Model Fit	RMR	0.044	Model Fit
NNFI	0.950	Model Fit	NNFI	0.950	Model Fit
AGFI	0.920	Model Fit	AGFI	0.940	Model Fit
CFI	0.970	Model Fit	CFI	0.960	Model Fit
IFI	0.970	Model Fit	IFI	0.960	Model Fit

Paket A

Paket B

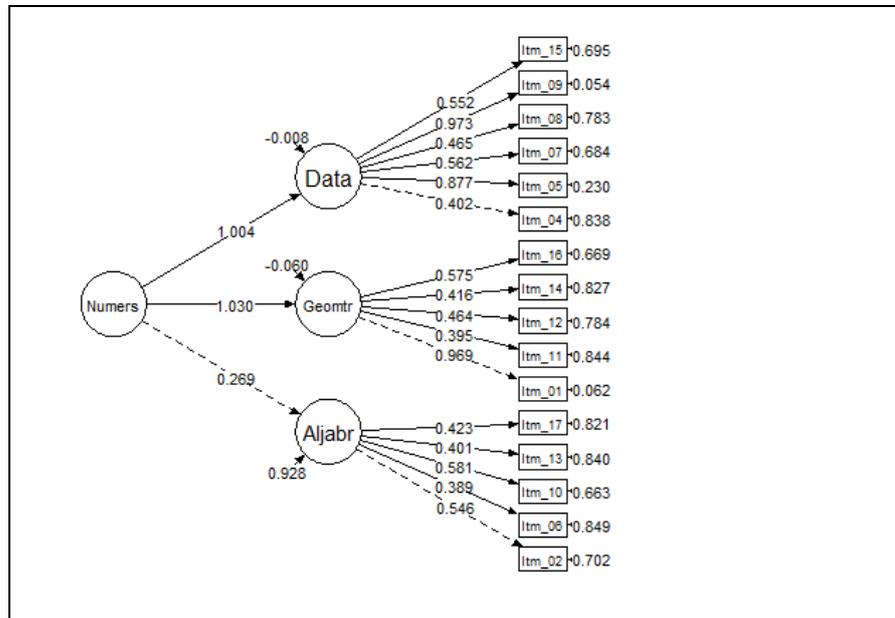


**Gambar 20.** Diagram Path CFA Soal Numerasi Paket A Tahap I



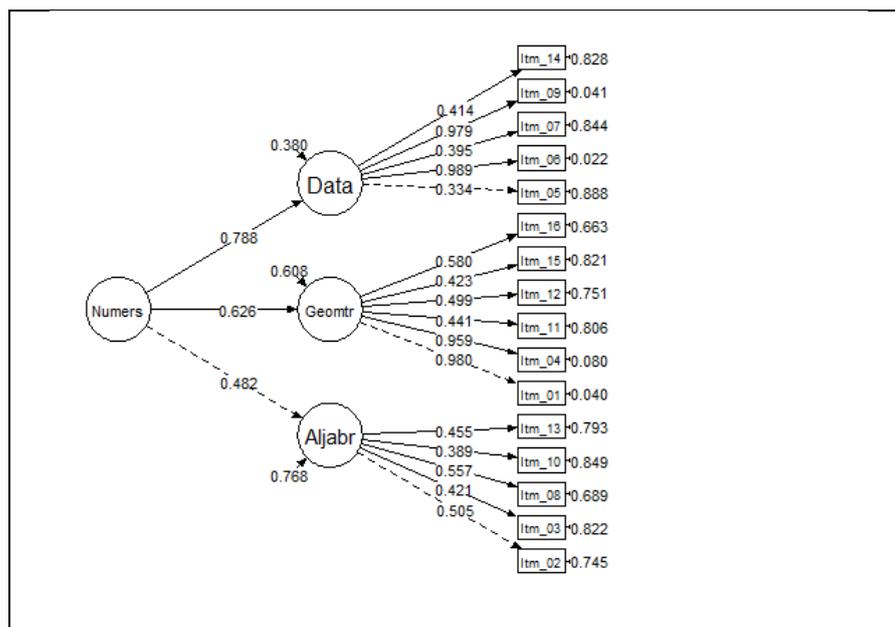
Berdasarkan Gambar 20, nilai *loading factor* seluruh item soal paket A berada pada rentang nilai 0,082 hingga 0,973. Hampir seluruh item memiliki nilai *loading factor* > 0,3 kecuali item nomor 3 dengan *loading factor* sebesar 0,082 < 0,3. Item nomor 3 menunjukkan bahwa item tersebut tidak valid terhadap model. Hal tersebut dapat terjadi akibat dari narasi soal yang kurang baik, materi soal kurang relevan, atau pilihan jawaban yang kurang baik sehingga menyebabkan data empiris tidak sesuai dengan model. Satu item soal tersebut kemudian dikeluarkan dan tidak disertakan dalam analisis lanjut yakni analisis teori respons butir. 16 item soal lain yang dinyatakan valid secara konstruk kemudian dianalisis CFA ulang sehingga hasilnya dapat dilihat pada Gambar 21.

**Gambar 21.** Diagram Path CFA Soal Numerasi Paket A Tahap II



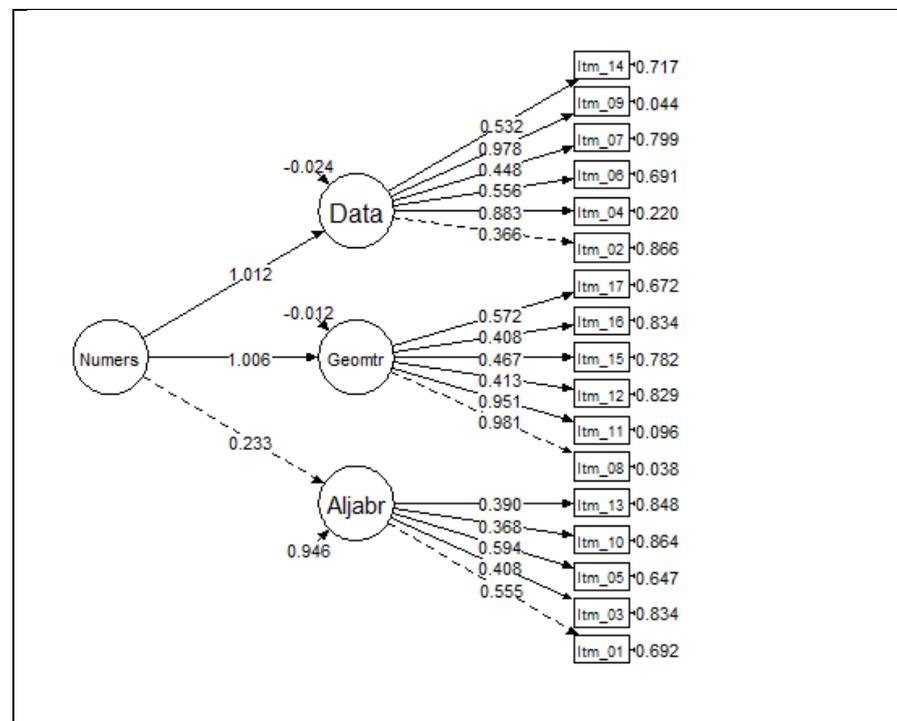
Gambar 21 menunjukkan bahwa analisis CFA kedua item soal paket A berada pada rentang nilai 0,389 hingga 0,973. Semua item soal valid secara konstruk dengan *loading factor* > 0,3 dan semua dapat dilanjut untuk analisis teori respons butir.

**Gambar 22.** Diagram Path CFA Soal Numerasi Paket B



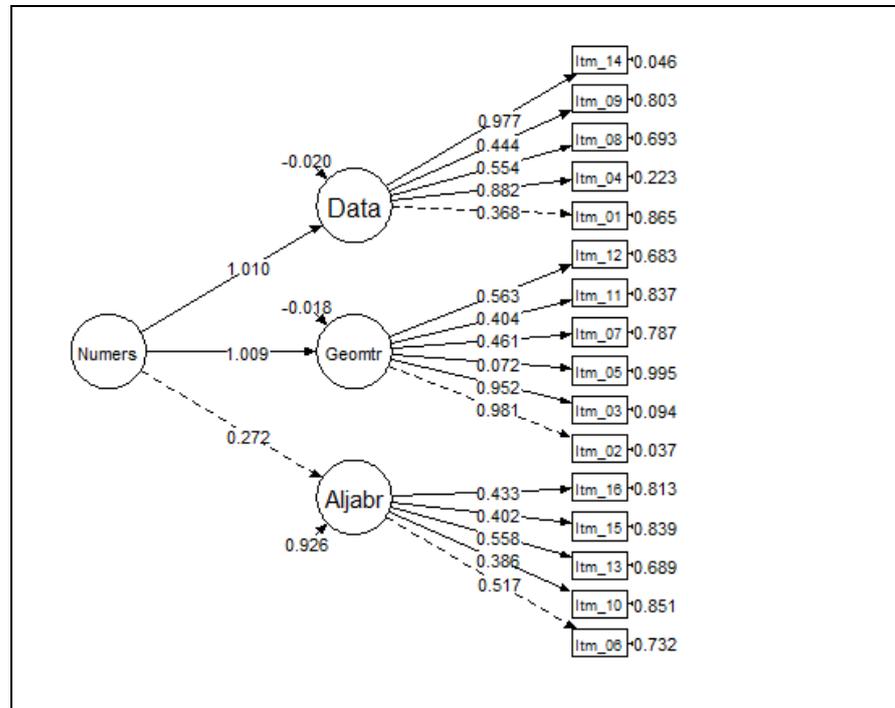
Berdasarkan Gambar 22, nilai *loading factor* seluruh item soal pada paket B berada pada rentang nilai 0,334 hingga 0,989. Seluruh item soal memiliki nilai *loading factor* > 0,3. Hal tersebut menunjukkan bahwa semua item soal pada paket B ini valid secara konstruk dan dapat dilanjut untuk analisis teori respons butir. Item soal valid secara konstruk dapat terjadi karena narasi soal yang baik, materi soal yang relevan, atau pilihan jawaban yang baik sehingga menyebabkan data empiris sesuai dengan model.

**Gambar 23.** Diagram Path CFA Soal Numerasi Paket C



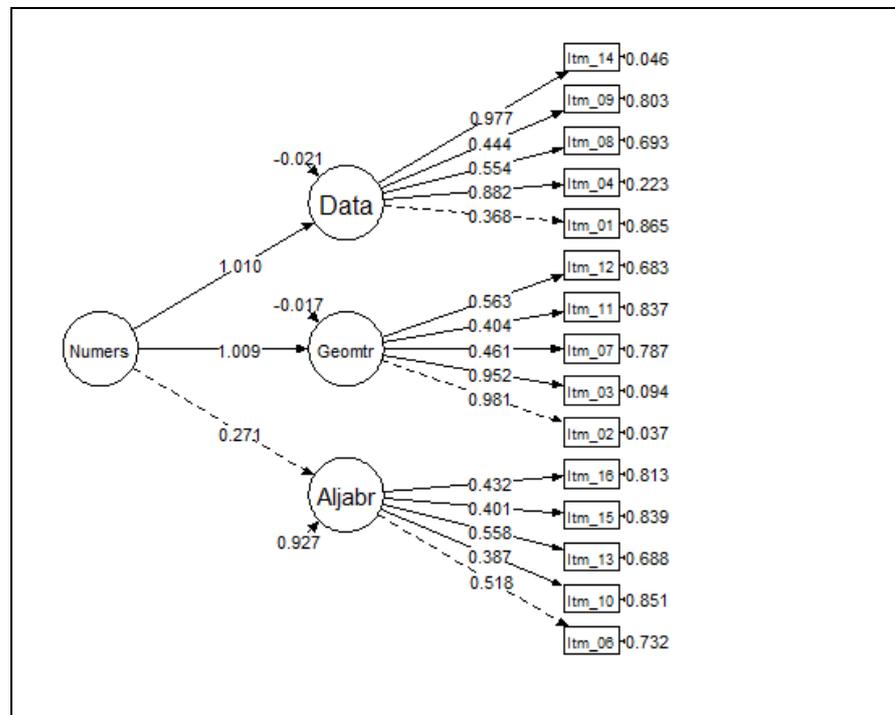
Berdasarkan Gambar 23, nilai *loading factor* seluruh item soal pada paket C berada pada rentang nilai 0,366 hingga 0,981. Seluruh item soal memiliki nilai *loading factor* > 0,3. Hal tersebut menunjukkan bahwa semua item soal pada paket C ini valid secara konstruk dan dapat dilanjut untuk analisis teori respons butir. Item soal valid secara konstruk dapat terjadi karena narasi soal yang baik, materi soal yang relevan, atau pilihan jawaban yang baik sehingga menyebabkan data empiris sesuai dengan model.

**Gambar 24.** Diagram Path CFA Soal Numerasi Paket D Tahap I



Berdasarkan Gambar 24, nilai *loading factor* seluruh item soal paket D berada pada rentang nilai 0,072 hingga 0,977. Hampir seluruh item memiliki nilai *loading factor*  $> 0,3$  kecuali item nomor 5 dengan *loading factor* sebesar  $0,072 < 0,3$ . Item nomor 5 menunjukkan bahwa item tersebut tidak valid terhadap model. Hal tersebut dapat terjadi akibat dari narasi soal yang kurang baik, materi soal kurang relevan, atau pilihan jawaban yang kurang baik sehingga menyebabkan data empiris tidak sesuai dengan model. Satu item soal tersebut kemudian dikeluarkan dan tidak disertakan dalam analisis lanjut yakni analisis teori respons butir. 15 item soal lain yang dinyatakan valid secara konstruk kemudian dianalisis CFA ulang sehingga hasilnya dapat dilihat pada Gambar 25.

**Gambar 25.** Diagram Path CFA Soal Numerasi Paket D Tahap II



Gambar 25 menunjukkan bahwa analisis CFA kedua item soal paket D berada pada rentang nilai 0,368 hingga 0,977. Semua item soal valid secara konstruk dengan *loading factor* > 0,3 dan semua dapat dilanjut untuk analisis teori respons butir.

**b. Uji Reliabilitas Instrumen**

Berikut adakah hasil uji reliabilitas instrumen soal AKM numerasi dari masing-masing paket soal.

**Gambar 26.** Hasil Uji Reliabilitas

Reliability Statistics		Reliability Statistics	
Cronbach's Alpha	N of Items	Cronbach's Alpha	N of Items
.770	17	.748	16
Paket A		Paket B	
Reliability Statistics		Reliability Statistics	
Cronbach's Alpha	N of Items	Cronbach's Alpha	N of Items
.756	17	.750	16
Paket C		Paket D	

Hasil analisis SPSS diketahui bahwa nilai reliabilitas *Cronbach's Alpha* untuk paket A sebesar 0,770, paket B sebesar 0,748, paket C sebesar 0,756, dan paket D sebesar 0,750. Keempat nilai uji reliabilitas tersebut berada pada rentang nilai reliabilitas 0,60 – 0,80 yang menyatakan bahwa instrumen andal/reliabel kategori tinggi. Sehingga dapat disimpulkan bahwa instrumen soal AKM numerasi dari keempat paket dikatakan reliabel.

**c. Uji Kualitas Soal**

Tahap analisis kualitas soal AKM numerasi yang telah diuji cobakan dianalisis menggunakan *item response theory* (IRT). Analisis IRT dilakukan pada item-item soal yang lolos pada uji validitas konstruk. Ada beberapa asumsi yang harus dipenuhi terlebih dahulu sebelum uji kualitas lebih lanjut. Asumsi yang dimaksud adalah unidimensi, independensi lokal, dan invariansi parameter.

**1) Uji Asumsi *Item Response Theory***

**a) Unidimensi**

Pembuktian unidimensi ditunjukkan melalui analisis faktor dengan melihat oleh nilai Nilai *Eigen*, *Total Variance Explained* dan *Scree Plot*. Asumsi unidimensi terpenuhi apabila instrumen soal yang diujikan memperlihatkan faktor yang lebih dominan dibandingkan faktor lain (Hambleton & Swaminathan, 1985, p. 16).

Analisis faktor dapat dilakukan apabila memenuhi syarat kecukupan sampel, ditunjukkan dengan nilai KMO lebih dari 0,5. Berdasarkan hasil uji KMO dan *Bartlett's Test* keempat paket soal memiliki nilai KMO lebih dari 0,5 yakni untuk paket A sebesar A sebesar 0,845, paket B sebesar 0,832, paket C sebesar 0,811, dan paket D sebesar 0,775. Hal tersebut menunjukkan bahwa jumlah sampel sudah memenuhi kelayakan analisis faktor.

Analisis faktor kemudian dilanjutkan untuk memperoleh nilai *Eigen* yang dapat dilihat melalui *Total Variance Explained* pada Gambar 27.

**Gambar 27.** Hasil *Total Variance and Eigenvalues Explained*

Component	Total	Initial Eigenvalues		Component	Total	Initial Eigenvalues	
		% of Variance	Cumulative %			% of Variance	Cumulative %
1	4.559	26.815	26.815	1	4.252	26.577	26.577
2	1.980	11.650	38.465	2	1.961	12.255	38.831
3	1.244	7.320	45.785	3	1.300	8.126	46.958
4	1.140	6.707	52.492	4	1.173	7.334	54.291
5	.996	5.861	58.353	5	.988	6.174	60.465
6	.916	5.389	63.742	6	.868	5.424	65.889
7	.793	4.666	68.408	7	.844	5.273	71.162
8	.753	4.427	72.834	8	.733	4.579	75.740
9	.717	4.220	77.054	9	.649	4.058	79.798
10	.678	3.987	81.041	10	.627	3.920	83.718
11	.628	3.694	84.734	11	.544	3.399	87.117
12	.601	3.538	88.272	12	.509	3.181	90.298
13	.559	3.287	91.559	13	.493	3.084	93.382
14	.438	2.578	94.137	14	.436	2.728	96.109
15	.372	2.190	96.327	15	.372	2.324	98.433
16	.348	2.048	98.376	16	.251	1.567	100.000
17	.276	1.624	100.000				

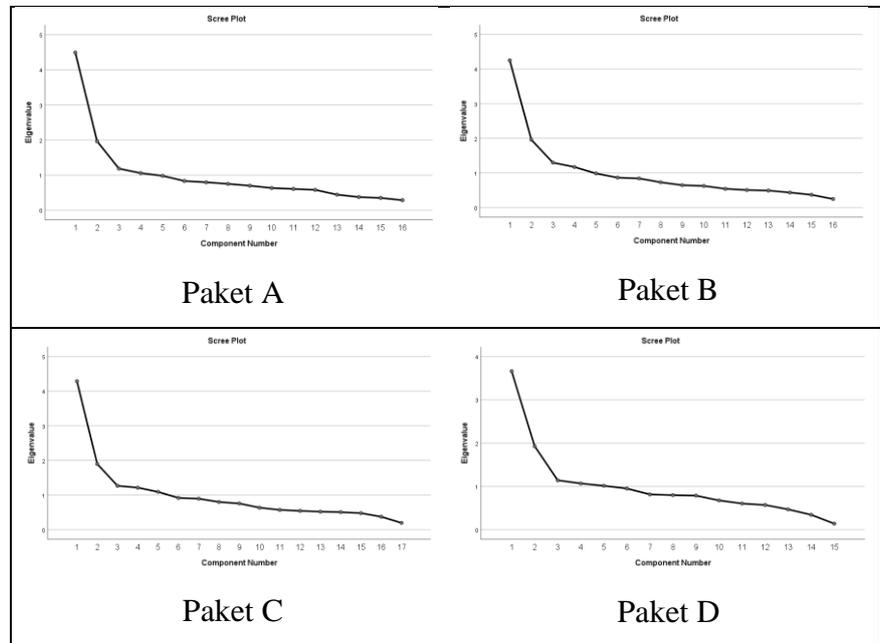
  

Component	Total	Initial Eigenvalues		Component	Total	Initial Eigenvalues	
		% of Variance	Cumulative %			% of Variance	Cumulative %
1	4.289	25.228	25.228	1	3.770	23.563	23.563
2	1.899	11.170	36.398	2	1.933	12.082	35.645
3	1.269	7.468	43.866	3	1.149	7.182	42.827
4	1.217	7.157	51.022	4	1.106	6.910	49.737
5	1.095	6.439	57.461	5	1.028	6.424	56.161
6	.921	5.419	62.881	6	.961	6.006	62.167
7	.899	5.286	68.166	7	.886	5.538	67.706
8	.803	4.725	72.891	8	.811	5.069	72.775
9	.760	4.470	77.361	9	.795	4.966	77.741
10	.638	3.755	81.116	10	.773	4.833	82.575
11	.574	3.378	84.494	11	.654	4.089	86.664
12	.546	3.212	87.706	12	.605	3.781	90.445
13	.523	3.077	90.782	13	.570	3.565	94.010
14	.510	2.998	93.780	14	.470	2.938	96.948
15	.481	2.828	96.608	15	.346	2.161	99.109
16	.377	2.218	98.826	16	.143	.891	100.000
17	.200	1.174	100.000				

Gambar 27 memperlihatkan bahwa setiap paket memiliki satu faktor yang dominan jika dibandingkan faktor lain. Nilai *Eigen* dari faktor yang dominan untuk paket A sebesar 4,494 dengan persentase kumulatif sebesar 28,085%, paket B sebesar 4,252 dengan persentase kumulatif sebesar 26,577%, paket C sebesar 4,289 dengan persentase kumulatif sebesar 25,228%, dan paket D sebesar 3,662 dengan persentase kumulatif sebesar

24,413%. Hasil dominasi satu faktor tersebut dapat diperjelas dengan melihat *Scree plot* Gambar 28.

**Gambar 28.** Hasil *Scree Plot* Uji Unidimensi



Pada seluruh paket soal terdapat satu curaman antara faktor 1 dengan faktor 2, sedangkan faktor 2 ke faktor 3 atau ke faktor lainnya yang mulai landai. Selain itu, untuk masing-masing paket nilai *Eigen* pada faktor 1 memiliki nilai 2 kali lipat lebih besar dibandingkan faktor 2. Oleh karena itu instrumen soal AKM numerasi pada masing-masing paket telah bersifat unidimensi.

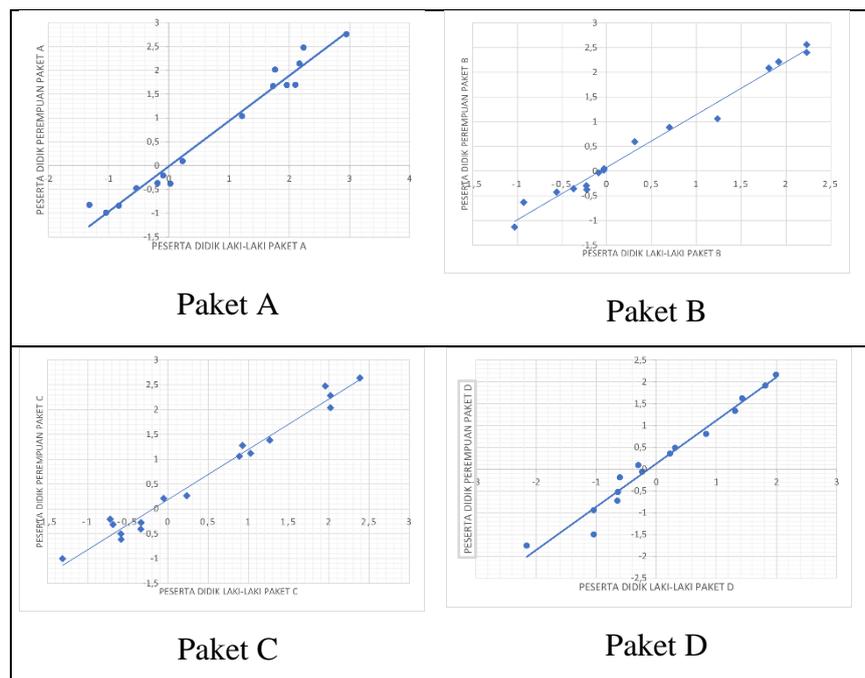
**b) Independensi Lokal**

Asumsi independensi lokal terpenuhi apabila jawaban peserta didik pada salah satu item tidak mempengaruhi jawabannya pada item soal yang lain. Asumsi independensi lokal otomatis terpenuhi apabila asumsi unidimensi juga terpenuhi (Retnawati, 2014, p. 7). Dikarenakan uji asumsi unidimensi telah terbukti bahwa keempat paket soal AKM numerasi bersifat unidimensi, maka asumsi independensi lokal untuk keempat paket juga telah terpenuhi.

### c) Invariansi Parameter

Invariansi parameter menunjukkan bahwa item-item soal tidak bergantung pada distribusi parameter kemampuan peserta tes dan parameter yang menjadi ciri peserta tes tidak bergantung ciri item-item soal. Uji invariansi parameter terdiri atas invariansi parameter butir dan parameter kemampuan. Pengujian ini dilakukan dengan melihat *scatter plot*, dimana jika titik-titik pada *scatter plot* mendekati garis yang melewati titik asal dengan gradien atau kemiringan sebesar 1 maka parameter-parameter tersebut invarian (Retnawati, 2014, p. 8). Berikut adalah hasil uji invariansi parameter yang dapat dilihat dari data uji coba.

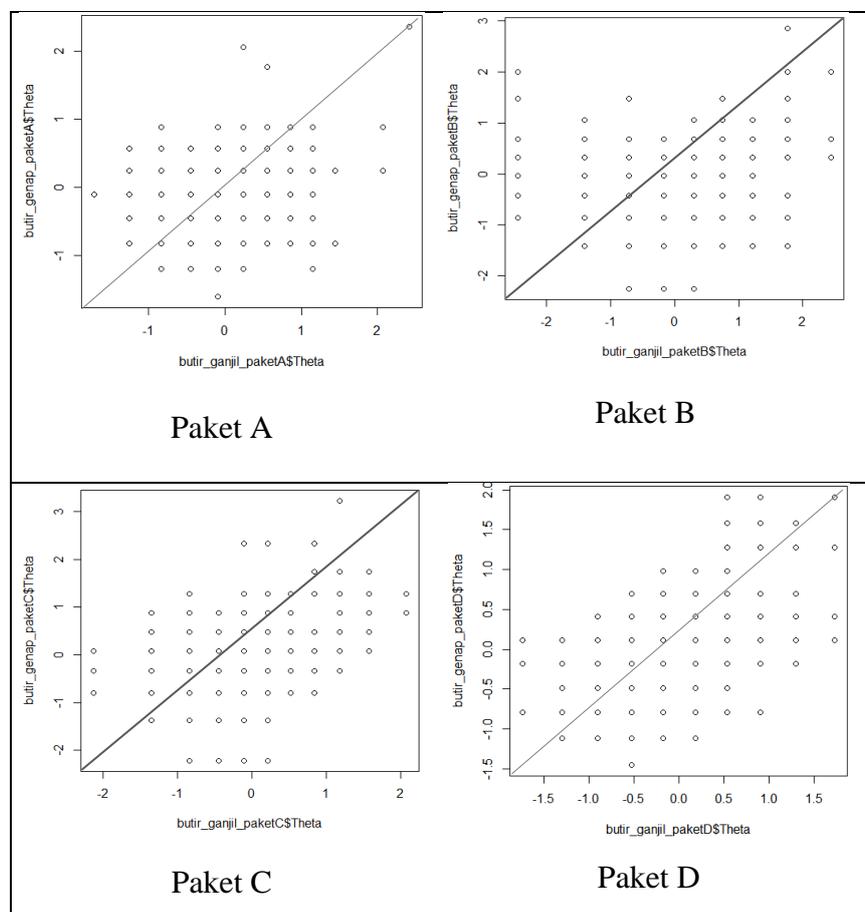
**Gambar 29.** Hasil Invariansi Parameter Butir



Asumsi invariansi parameter butir pembuktiannya ditujukan untuk parameter tingkat kesukaran dengan estimasi model *Rasch*. Pembuktian invariansi parameter butir ini dilakukan dengan membagi peserta didik menjadi dua kelompok yakni kelompok peserta didik berjenis laki-laki dan kelompok

perempuan. Gambar 29 menunjukkan bahwa tiap-tiap data disetiap paket memiliki posisi relatif dekat dengan garis pertolongan yang melewati titik asal dengan gradien atau kemiringan sebesar 1 (garis  $y = x$ ), sehingga parameter butir yang telah dikembangkan bersifat invarians. Selanjutnya untuk pembuktian invariansi parameter kemampuan dapat dilihat pada Gambar 30.

**Gambar 30.** Hasil Invariansi Parameter Kemampuan



Asumsi invariansi parameter kemampuan pembuktiannya ditujukan pada hasil estimasi kemampuan seluruh peserta didik dengan respon dari item-item soal ganjil dan item-item soal genap per masing-masing paket secara terpisah. Gambar 30 menunjukkan bahwa tiap-tiap data juga memiliki posisi relatif dekat dengan garis pertolongan yang melewati titik asal dengan

gradien atau kemiringan sebesar 1 (garis  $y = x$ ), sehingga parameter kemampuan yang telah dikembangkan bersifat invarians. Hal ini bermakna bahwa kemampuan peserta didik tidak berubah hanya karena karakteristik item soal seperti daya beda, tingkat kesukaran, tebakan semu, dll. Instrumen soal AKM numerasi yang telah dikembangkan memenuhi asumsi unidimensi, independensi lokal, serta invariansi parameter. Maka dari itu selanjutnya instrumen dilanjutkan analisis data campuran dengan pendekatan IRT.

## 2) Uji Data Campuran

### a) Uji Kecocokan Butir

Uji kecocokan butir/item *fit* merupakan indeks penentu suatu butir yang memenuhi persyaratan sebagai alat pengukur yang baik dan berfungsi secara optimal. Hal tersebut dapat diketahui melalui estimasi pola respon peserta didik terhadap butir-butir instrumen (Zi Yan & Heene, 2021, p. 54). Penelitian ini menggunakan uji data campuran politomus dan dikotomus sehingga model teori respons butir yang digunakan untuk analisis instrumen adalah *Partial Credit Model* (PCM) dan *Generalized Partial Credit Model* (GPCM) pada data politomus. Digunakan dua model tersebut karena penskoran untuk data politomus berupa kredit parsial. Sedangkan untuk data dikotomus model yang digunakan adalah model *Rasch*, model parameter logistik 2PL, dan 3PL.

Estimasi kecocokan butir pada penelitian ini tidak didasarkan pada nilai *Chi-square* karena menurut Tesio et al. (2023, p. 5) estimasi *Chi-square* bergantung pada ukuran sampel, terlebih sensitif untuk sampel besar. Uji kecocokan butir pada penelitian ini ditentukan berdasarkan nilai *infit* dan *outfit*. Nilai *infit* dan *outfit* dari instrumen soal AKM numerasi

diestimasi dengan bantuan aplikasi Program R yang hasilnya dapat dilihat pada Tabel 19.

**Tabel 19.** Uji Kecocokan Butir Paket A

DIKOTOMUS										
Kode	No	Rasch			2PL			3PL		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.	Outfit	Infit	Ket.
D1	1	0,817	0,876	<i>fit</i>	0,896	0,914	<i>fit</i>	0,915	1,035	<i>fit</i>
D2	2	0,787	0,843	<i>fit</i>	0,913	0,91	<i>fit</i>	0,014	0,029	<i>overfit</i>
D4	10	0,858	0,936	<i>fit</i>	0,905	0,934	<i>fit</i>	0,957	0,972	<i>fit</i>
D5	11	0,763	0,835	<i>fit</i>	0,863	0,869	<i>fit</i>	0,738	0,858	<i>mix</i>
D6	12	0,706	0,842	<i>mix</i>	0,647	1,009	<i>mix</i>	0,805	0,808	<i>fit</i>
D7	13	0,795	0,858	<i>fit</i>	0,44	0,912	<i>mix</i>	0,941	0,972	<i>fit</i>
D8	14	0,814	0,86	<i>fit</i>	0,907	0,908	<i>fit</i>	0,821	0,951	<i>fit</i>
D9	15	0,597	0,868	<i>mix</i>	0,218	1,023	<i>mix</i>	0,96	0,994	<i>fit</i>

POLITOMUS							
Kode	No	PCM			GPCM		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.
P1	4	0,988	0,993	<i>fit</i>	0,985	0,985	<i>fit</i>
P2	5	0,877	0,925	<i>fit</i>	0,942	0,958	<i>fit</i>
P3	6	0,927	0,948	<i>fit</i>	0,994	0,992	<i>fit</i>
P4	7	0,843	0,862	<i>fit</i>	0,924	0,931	<i>fit</i>
P5	8	0,626	0,807	<i>mix</i>	0,592	1,018	<i>mix</i>
P6	9	0,651	0,857	<i>mix</i>	0,415	1,056	<i>mix</i>
P7	16	0,845	0,854	<i>fit</i>	0,665	0,975	<i>mix</i>
P8	17	0,806	0,901	<i>fit</i>	0,68	0,941	<i>mix</i>

Suatu butir dikatakan *fit* apabila nilai *infit* dan *outfit* bernilai 0,75 s.d. 1,30, *underfit* apabila nilai *infit* atau *outfit* bernilai lebih dari 1,30, dan *overfit* apabila nilai *infit* atau *outfit* bernilai kurang dari 0,75 (Zi Yan & Heene, 2021, p. 242). Berdasarkan Tabel 19, untuk data dikotomus kecocokan item dengan model *Rasch* terdapat 6 item *fit* dan 2 item *mix*, model 2PL terdapat 5 item *fit* dan 3 item *mix*, dan model 3PL terdapat 6 item *fit*, 1 item *mix*, dan 1 item *overfit*. Seluruh item *mix fit* pada data dikotomus ini menunjukkan *fit* pada nilai *infit* dan *overfit* pada nilai *outfit*. Zi Yan & Heene (2021, p. 56) menjelaskan bahwa nilai *outfit* lebih sensitif terhadap *outlying* sehingga pertimbangan utama ada pada nilai *infit* daripada nilai *outfit*. Maka dari itu ketika terdapat item *mix fit* (*fit* pada nilai *infit*) dan item *overfit* maka yang diprioritaskan adalah item *mix fit*. Selain itu item akan tidak diprioritaskan apabila nilai *infit* atau *outfit* tidak pada rentang nilai 0,5 hingga 1,5 karena diluar

rentang nilai tersebut item tidak produktif dalam pengukuran dan tidak bisa untuk dipertahankan (Susac et al., 2018, p. 3).

Berdasarkan Tabel 19 untuk data politomus kecocokan item dengan model PCM terdapat 6 item *fit* dan 2 item *mix*, dan model GPCM terdapat 4 item *fit* dan 4 item *mix*. Seluruh item *mix fit* pada data dikotomus ini menunjukkan *fit* pada nilai *infit* dan *overfit* pada nilai *outfit*. Berdasarkan pertimbangan tersebut, untuk data dikotomus dan politomus paket A dapat dilakukan analisis teori respons butir lanjutan dengan model *Rasch* pada kedua data (model PCM merupakan model *Rasch* pada politomus).

**Tabel 20.** Uji Kecocokan Butir Paket B

DIKOTOMUS										
Kode	No	Rasch			2PL			3PL		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.	Outfit	Infit	Ket.
D1	1	0,872	0,921	<i>fit</i>	0,743	0,892	<i>mix</i>	0,981	1,041	<i>fit</i>
D2	2	0,863	0,888	<i>fit</i>	0,96	0,963	<i>fit</i>	0,865	0,865	<i>fit</i>
D3	3	0,964	0,963	<i>fit</i>	0,989	0,991	<i>fit</i>	0,951	1,005	<i>fit</i>
D4	10	0,908	0,949	<i>fit</i>	0,999	1	<i>fit</i>	0,965	0,972	<i>fit</i>
D5	11	0,576	0,774	<i>mix</i>	0,543	1,133	<i>mix</i>	0,973	1,303	<i>mix</i>
D6	12	0,633	0,792	<i>mix</i>	0,058	0,644	<i>overfit</i>	1,297	1,621	<i>mix</i>
D7	13	0,803	0,847	<i>fit</i>	0,957	0,958	<i>fit</i>	0,897	0,893	<i>fit</i>
D8	14	0,606	0,797	<i>mix</i>	0,217	1,193	<i>mix</i>	1,136	1,563	<i>mix</i>
D9	15	0,636	0,801	<i>mix</i>	0,3	1,081	<i>mix</i>	1,336	1,13	<i>mix</i>

POLITOMUS							
Kode	No	PCM			GPCM		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.
P1	4	0,945	0,961	<i>fit</i>	0,976	0,976	<i>fit</i>
P2	5	0,964	0,969	<i>fit</i>	0,976	0,986	<i>fit</i>
P3	6	0,897	0,915	<i>fit</i>	0,933	0,944	<i>fit</i>
P4	7	0,837	0,854	<i>fit</i>	0,788	0,829	<i>fit</i>
P5	8	0,668	0,717	<i>overfit</i>	0,635	0,742	<i>overfit</i>
P6	9	0,65	0,729	<i>overfit</i>	0,418	0,562	<i>overfit</i>
P7	16	0,953	0,914	<i>fit</i>	0,935	0,906	<i>fit</i>

Berdasarkan Tabel 20, untuk data dikotomus kecocokan item dengan model *Rasch* terdapat 5 item *fit* dan 4 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*), model 2PL terdapat 4 item *fit*, 4 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*), serta 1 item *overfit*, dan model 3PL terdapat 5 item *fit* dan 4 item *mix* (3 item *underfit* pada nilai *infit* dan *overfit* pada nilai

*output*). Karena pada model 3PL item *mix underfit* pada nilai *infit* (atau tidak ada yang *fit* disalah satu nilai *infit* atau *outfit*), maka analisis teori respons butir lanjutan dengan model *Rasch* pada data dikotomus.

Berdasarkan Tabel 20 untuk data politomus kecocokan item dengan model PCM terdapat 5 item *fit* dan 2 item *overfit*, dan model GPCM terdapat 5 item *fit* dan 2 item *overfit*. Karena kedua model memiliki komposisi yang sama, maka pertimbangannya disesuaikan dengan yang dinyatakan Susac et al. (2018, p. 3) bahwa item akan tidak diprioritaskan apabila nilai *infit* atau *outfit* tidak pada rentang nilai 0,5 hingga 1,5 karena diluar rentang nilai tersebut item tidak produktif dalam pengukuran dan tidak bisa untuk dipertahankan. Berdasarkan hal tersebut 1 item *overfit* pada model GPCM terdapat nilai kurang dari 0,5 sehingga keputusan penggunaan model untuk lanjut analisis teori respons butir pada data politomus ini adalah model PCM. Atau dengan kata lain untuk paket B model yang digunakan untuk lanjut analisis adalah model *Rasch*.

**Tabel 21.** Uji Kecocokan Butir Paket C

DIKOTOMUS											
Kode	No	Rasch			2PL			3PL			
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.	Outfit	Infit	Ket.	
D1	1	0,922	0,971	<i>fit</i>	0,956	0,966	<i>fit</i>	0,828	0,88	<i>fit</i>	
D2	2	1,034	1,156	<i>fit</i>	0,514	1,065	<i>mix</i>	0,84	0,879	<i>fit</i>	
D3	3	0,972	1,03	<i>fit</i>	0,892	0,936	<i>fit</i>	0,856	0,919	<i>fit</i>	
D4	10	0,914	0,92	<i>fit</i>	0,982	0,984	<i>fit</i>	0,924	0,93	<i>fit</i>	
D5	11	1,006	1,21	<i>fit</i>	0,554	1,087	<i>mix</i>	0,654	0,836	<i>mix</i>	
D6	12	0,944	0,998	<i>fit</i>	0,493	1,1	<i>mix</i>	0,744	0,857	<i>mix</i>	
D7	13	0,916	0,92	<i>fit</i>	0,982	0,983	<i>fit</i>	0,9	0,893	<i>fit</i>	
D8	14	0,942	0,994	<i>fit</i>	0,244	1,149	<i>mix</i>	0,713	0,799	<i>mix</i>	
D9	15	0,818	1,129	<i>fit</i>	0,679	0,904	<i>mix</i>	0,729	0,822	<i>mix</i>	

POLITOMUS							
Kode	No	PCM			GPCM		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.
P1	4	0,975	0,998	<i>fit</i>	0,983	0,984	<i>fit</i>
P2	5	0,915	0,934	<i>fit</i>	0,905	0,921	<i>fit</i>
P3	6	1,083	1,043	<i>fit</i>	0,986	0,997	<i>fit</i>
P4	7	0,721	0,78	<i>mix</i>	0,878	0,922	<i>fit</i>
P5	8	0,686	0,764	<i>mix</i>	0,742	0,834	<i>mix</i>
P6	9	0,846	0,874	<i>fit</i>	0,958	0,962	<i>fit</i>
P7	16	0,754	0,768	<i>fit</i>	0,126	0,614	<i>overfit</i>
P8	17	0,646	0,737	<i>overfit</i>	0,381	0,896	<i>mix</i>

Berdasarkan Tabel 21, untuk data dikotomus kecocokan item dengan model *Rasch* terdapat 9 item *fit*, model 2PL terdapat 4 item *fit* dan 5 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*), dan model 3PL terdapat 5 item *fit* dan 4 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*). Berdasarkan komposisi tersebut maka analisis teori respons butir lanjut dengan model *Rasch* pada data dikotomus.

Berdasarkan Tabel 21 untuk data politomus kecocokan item dengan model PCM terdapat 5 item *fit*, 2 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*) dan 1 item *overfit*, dan model GPCM terdapat 5 item *fit*, 2 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*) dan 1 item *overfit*. Karena kedua model memiliki komposisi yang sama, maka pertimbangannya disesuaikan dengan yang dinyatakan Susac et al. (2018, p. 3) bahwa item akan tidak diprioritaskan apabila nilai *infit* atau *outfit* tidak pada rentang nilai 0,5 hingga 1,5 karena diluar rentang nilai tersebut item tidak produktif dalam pengukuran dan tidak bisa untuk dipertahankan. Berdasarkan hal tersebut 2 item pada model GPCM terdapat nilai kurang dari 0,5 sehingga keputusan penggunaan model untuk lanjut analisis teori respons butir pada data politomus ini adalah model PCM. Atau dengan kata lain untuk paket C model yang digunakan untuk lanjut analisis adalah model *Rasch*.

**Tabel 22.** Uji Kecocokan Butir Paket D

Kode	No	DIKOTOMUS								
		<i>Rasch</i>			2PL			3PL		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.	Outfit	Infit	Ket.
D1	1	0,897	0,924	<i>fit</i>	0,958	0,961	<i>fit</i>	0,924	0,995	<i>fit</i>
D2	2	0,866	0,896	<i>fit</i>	0,908	0,924	<i>fit</i>	0,921	0,992	<i>fit</i>
D3	3	0,843	0,903	<i>fit</i>	0,975	0,978	<i>fit</i>	0,91	0,914	<i>fit</i>
D4	10	0,864	0,894	<i>fit</i>	0,992	0,992	<i>fit</i>	0,967	0,967	<i>fit</i>
D5	11	0,793	0,86	<i>fit</i>	0,843	0,961	<i>fit</i>	0,743	0,952	<i>mix</i>
D6	12	0,741	0,87	<i>mix</i>	0,007	0,186	<i>overfit</i>	0,961	0,99	<i>fit</i>
D7	13	0,906	0,926	<i>fit</i>	1	1	<i>fit</i>	0,987	0,987	<i>fit</i>
D8	14	0,84	0,882	<i>fit</i>	0,079	1,374	<i>mix</i>	0,866	1,005	<i>fit</i>
D9	15	0,807	0,85	<i>fit</i>	0,963	0,963	<i>fit</i>	0,941	0,945	<i>fit</i>

POLITOMUS							
Kode	No	PCM			GPCM		
Item	Item	Outfit	Infit	Ket.	Outfit	Infit	Ket.
P1	4	0,864	0,952	<i>fit</i>	0,916	0,976	<i>fit</i>
P3	6	0,83	1,063	<i>fit</i>	0,906	1,006	<i>fit</i>
P4	7	0,732	0,757	<i>mix</i>	0,737	0,786	<i>mix</i>
P5	8	0,699	0,733	<i>overfit</i>	0,835	0,706	<i>mix</i>
P6	9	0,72	0,752	<i>mix</i>	0,474	0,549	<i>overfit</i>
P7	16	0,811	0,875	<i>fit</i>	0,899	0,913	<i>fit</i>

Berdasarkan Tabel 22, untuk data dikotomus kecocokan item dengan model *Rasch* terdapat 8 item *fit* dan 1 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*), model 2PL terdapat 7 item *fit*, 1 item *mix* (*underfit* pada nilai *infit* dan *overfit* pada nilai *outfit*) dan 1 item *overfit*, dan model 3PL terdapat 8 item *fit* dan 1 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*). Berdasarkan komposisi tersebut antara model *Rasch* dan 3PL memiliki jumlah dan detail komposisi yang sama, maka analisis teori respons butir lanjut dengan dapat dengan model *Rasch* atau model 3PL pada data dikotomus. Namun karena pada paket sebelumnya model *Rasch* yang terpilih maka pada paket D akan digunakan model *Rasch* juga.

Berdasarkan Tabel 22 untuk data politomus kecocokan item dengan model PCM terdapat 3 item *fit*, 2 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*) dan 1 item *overfit*, dan model GPCM terdapat 3 item *fit*, 2 item *mix* (*fit* pada nilai *infit* dan *overfit* pada nilai *outfit*) dan 1 item *overfit*. Karena kedua model memiliki komposisi yang sama, maka pertimbangannya disesuaikan dengan yang dinyatakan Susac et al. (2018, p. 3) bahwa item akan tidak diprioritaskan apabila nilai *infit* atau *outfit* tidak pada rentang nilai 0,5 hingga 1,5 karena diluar rentang nilai tersebut item tidak produktif dalam pengukuran dan tidak bisa untuk dipertahankan. Berdasarkan hal tersebut 1 item pada model GPCM terdapat nilai kurang dari 0,5 sehingga keputusan penggunaan model untuk lanjut analisis teori respons butir pada

data politomus ini adalah model PCM. Atau dengan kata lain untuk paket D model yang digunakan untuk lanjut analisis adalah model *Rasch*. Secara keseluruhan paket model yang digunakan untuk lanjut analisis teori respons butir adalah model *Rasch*.

#### **b) Kurva Karakteristik Butir**

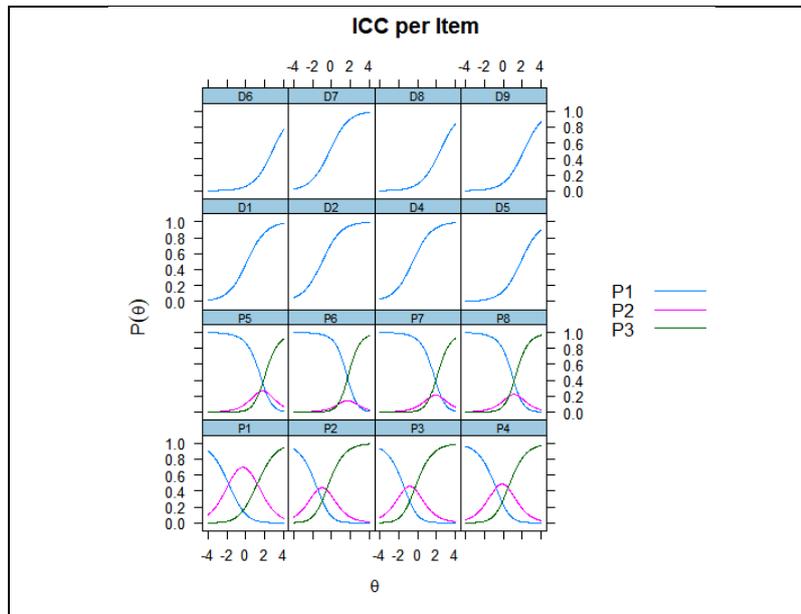
Parameter butir yang dianalisis dalam uji data campuran ini adalah tingkat kesukaran butir dengan estimasi model *Rasch*. Pada penskoran dikotomus model *Rasch* merupakan kasus khusus dari model IRT 1-PL. Sedangkan untuk penskoran politomus model *Rasch* yang digunakan adalah PCM (*Partial Credit Model*) yang merupakan model respon item keluarga *Rasch* yang mampu menangani respon item lebih dari dua kategori (Djidu et al., 2022, pp. 143–144).

Hasil estimasi parameter butir digambarkan pada grafik ICC (*Item Characteristic Curve*). Melalui grafik tersebut dapat diperoleh informasi mengenai probabilitas peserta didik menjawab benar pada tingkat kemampuannya masing-masing. Penskoran politomus probabilitas menjawab benar diestimasi dari nilai *thresholds* atau tingkat kesukaran item ( $b_i$ ) masing-masing lereng. Penskoran politomus dalam penelitian ini terdiri atas tiga kategori respon, sehingga terdapat 3 garis berbeda warna pada kurva item-item politomus. Perpotongan antara masing-masing garis merupakan ambang batas antar kategori atau *thresholds* atau tingkat kesukaran item ( $b_i$ ). Karena setiap item memiliki dua tingkat kesukaran item ( $b_1$  &  $b_2$ ), untuk penyimpulan tingkat kesukaran suatu item maka diperlukan *generalized item difficulty summaries* menggunakan estimasi *Location Indices* (LI).

Menurut Hambleton et al. (1991, p. 13) suatu item dikategorikan mudah jika nilai  $b < -2$ , kategori sedang jika

nilai  $-2 < b < 2$ , dan kategori sukar jika nilai  $b > 2$ . Estimasi karakteristik butir untuk paket A dapat dilihat pada Gambar 31 dan Tabel 23.

**Gambar 31.** *Item Characteristic Curve Paket A*



**Tabel 23.** *Nilai Parameter Butir Paket A*

Kode Item	No Item	$b_1$	$b_2$	$LI_{IRF}$	Kode Item	No Item	$b$
P1	4	-1,810	1,224	-0,293	D1	1	0,151
P2	5	-1,465	-0,555	-1,010	D2	2	-1,009
P3	6	-1,356	-0,309	-0,832	D4	10	-0,500
P4	7	-0,799	0,478	-0,160	D5	11	1,848
P5	8	2,112	1,470	1,791	D6	12	2,825
P6	9	2,787	0,601	1,694	D7	13	-0,214
P7	16	2,515	1,300	1,908	D8	14	2,377
P8	17	1,670	0,550	1,110	D9	15	2,155

Berdasarkan hasil uji data campuran paket A untuk tingkat kesukaran ( $b$ ) item dikotomus (D1 hingga D9 tanpa D3 karena sudah dieliminasi akibat tidak valid secara konstruk) berada pada kisaran -1,009 hingga 2,825. Sedangkan untuk item-item

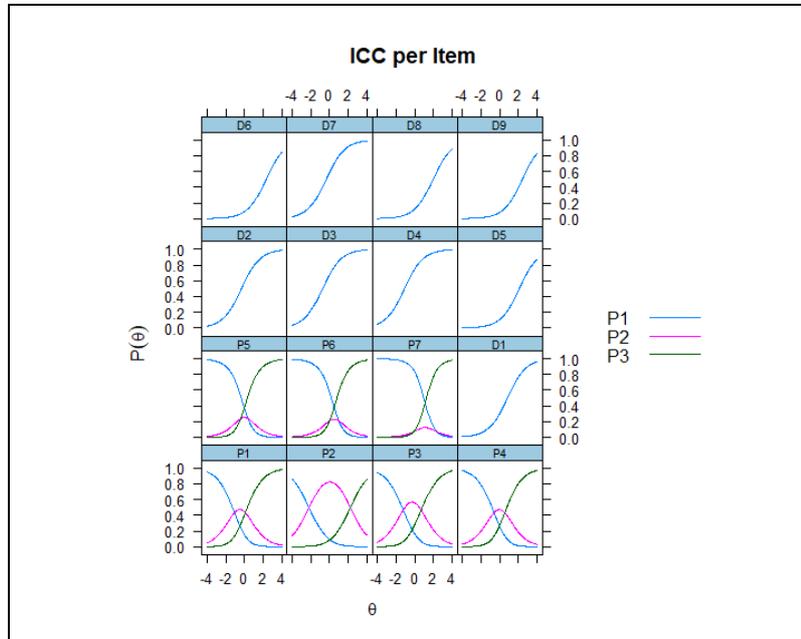
politomus (P1 hingga P8) tingkat kesukaran pertama ( $b_1$ ) berada pada kisaran -1,810 hingga 2,787 dan tingkat kesukaran kedua ( $b_2$ ) berada pada kisaran -0,555 hingga 1,470.

Pada P1 (item soal nomor 4) tingkat kesukaran atau *threshold* pertama ( $b_1$ ) sebesar -1,810 bermakna bahwa untuk dapat menyelesaikan kategori 2 setelah kategori 1 peserta didik harus memiliki kemampuan minimal -1,810. Sedangkan tingkat kesukaran atau *threshold* kedua ( $b_2$ ) sebesar 1,224 bermakna bahwa untuk dapat menyelesaikan kategori 3 setelah kategori 2 peserta didik harus memiliki kemampuan minimal 1,224. Kesukaran item sebesar -1,810 berarti tingkat kesukaran item sedang. Kesukaran item sebesar 1,224 berarti tingkat kesukaran item sedang. *Location Indices (LI)* pada P1 sebesar -0,293 yang menunjukkan bahwa ringkasan tingkat kesukaran untuk P1 memiliki tingkat kesukaran sedang.

Pada D1 (item soal nomor 1) tingkat kesukaran atau *threshold* ( $b$ ) sebesar 0,151 bermakna bahwa untuk menjawab benar item soal nomor 1 dengan probabilitas menjawab benar 50% diperlukan kemampuan minimal 0,151. Kesukaran item sebesar 0,151 berarti tingkat kesukaran item sedang. Untuk nilai tingkat kesukaran atau *threshold* masing-masing item di paket A dapat dilihat pada Tabel 23 di atas. Secara keseluruhan, tingkat kesukaran item pada paket A berada pada kategori sedang hingga sukar.

Estimasi karakteristik butir untuk paket B dapat dilihat pada Gambar 32 dan Tabel 24.

**Gambar 32.** Item Characteristic Curve Paket B



**Tabel 24.** Nilai Parameter Butir Paket B

Kode Item	No Item	$b_1$	$b_2$	$LI_{IRF}$	Kode Item	No Item	$b$
P1	4	-1,062	0,109	-0,477	D1	1	0,813
P2	5	-2,198	2,239	0,021	D2	2	-0,314
P3	6	-1,231	0,701	-0,265	D3	3	-0,747
P4	7	-0,630	0,522	-0,054	D4	10	-1,087
P5	8	0,396	-0,392	0,002	D5	11	2,088
P6	9	0,990	-0,018	0,486	D6	12	2,328
P7	16	2,378	-0,127	1,125	D7	13	-0,361
					D8	14	1,969
					D9	15	2,417

Berdasarkan hasil uji data campuran paket B untuk tingkat kesukaran ( $b$ ) item dikotomus (D1 hingga D9) berada pada kisaran -1,087 hingga 2,417. Sedangkan untuk item-item

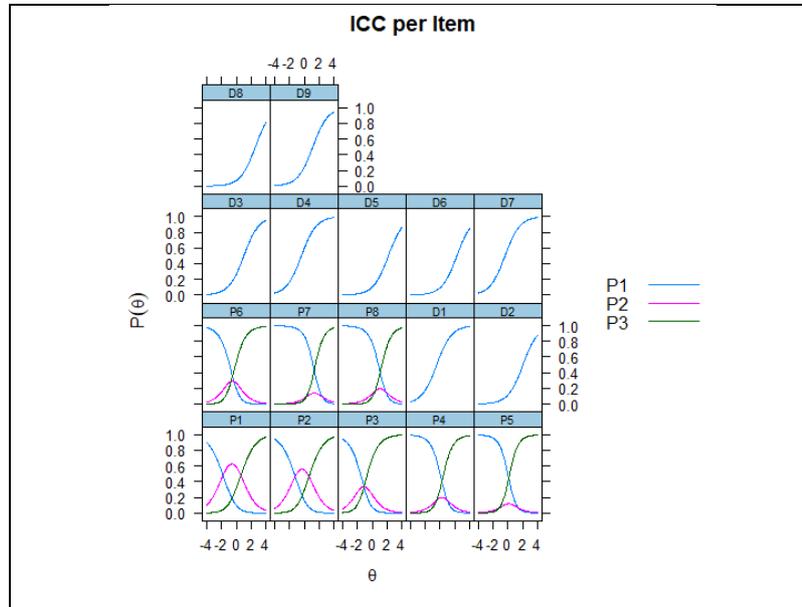
politomus (P1 hingga P7) tingkat kesukaran pertama ( $b_1$ ) berada pada kisaran -2,198 hingga 2,378 dan tingkat kesukaran kedua ( $b_2$ ) berada pada kisaran -0,392 hingga 2,239.

Pada P2 (item soal nomor 5) tingkat kesukaran atau *threshold* pertama ( $b_1$ ) sebesar -2,198 bermakna bahwa untuk dapat menyelesaikan kategori 2 setelah kategori 1 peserta didik harus memiliki kemampuan minimal -2,198. Sedangkan tingkat kesukaran atau *threshold* kedua ( $b_2$ ) sebesar 2,239 bermakna bahwa untuk dapat menyelesaikan kategori 3 setelah kategori 2 peserta didik harus memiliki kemampuan minimal 2,239. Kesukaran item sebesar -2,198 berarti tingkat kesukaran item mudah. Kesukaran item sebesar 2,239 berarti tingkat kesukaran item sukar. *Location Indices (LI)* pada P2 sebesar 0,021 yang menunjukkan bahwa ringkasan tingkat kesukaran untuk P2 memiliki tingkat kesukaran sedang.

Pada D2 (item soal nomor 2) tingkat kesukaran atau *threshold* ( $b$ ) sebesar -0,314 bermakna bahwa untuk menjawab benar item soal nomor 2 dengan probabilitas menjawab benar 50% diperlukan kemampuan minimal -0,314. Kesukaran item sebesar -0,314 berarti tingkat kesukaran item sedang. Untuk nilai tingkat kesukaran atau *threshold* masing-masing item di paket B dapat dilihat pada Tabel 24 di atas. Secara keseluruhan, tingkat kesukaran item pada paket B berada pada kategori mudah hingga sukar.

Estimasi karakteristik butir untuk paket C dapat dilihat pada Gambar 33 dan Tabel 25.

**Gambar 33.** *Item Characteristic Curve Paket C*



**Tabel 25.** *Nilai Parameter Butir Paket C*

Kode Item	No Item	$b_1$	$b_2$	$LI_{IRF}$	Kode Item	No Item	$b$
P1	4	-1,809	0,596	-0,607	D1	1	-0,445
P2	5	-1,224	0,632	-0,296	D2	2	2,029
P3	6	-1,118	-1,094	-1,106	D3	3	0,997
P4	7	0,958	-0,452	0,253	D4	10	-0,385
P5	8	1,442	-1,204	0,119	D5	11	2,181
P6	9	-0,342	-0,720	-0,531	D6	12	2,263
P7	16	2,459	0,217	1,338	D7	13	-0,385
P8	17	1,800	0,360	1,080	D8	14	2,538
					D9	15	1,145

Berdasarkan hasil uji data campuran paket C untuk tingkat kesukaran ( $b$ ) item dikotomus (D1 hingga D9) berada pada kisaran -0,445 hingga 2,538. Sedangkan untuk item-item politomus (P1 hingga P8) tingkat kesukaran pertama ( $b_1$ ) berada

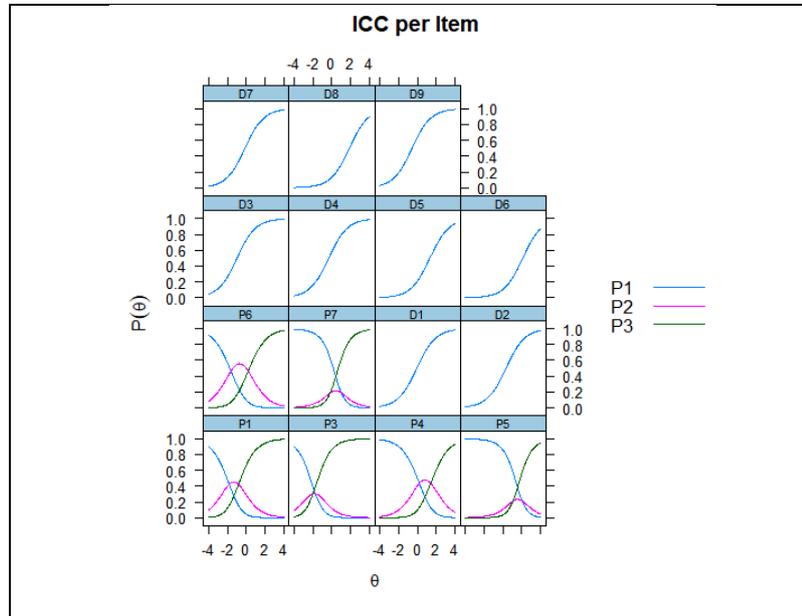
pada kisaran -1,809 hingga 2,459 dan tingkat kesukaran kedua ( $b_2$ ) berada pada kisaran -1,106 hingga 1,338.

Pada P3 (item soal nomor 6) tingkat kesukaran atau *threshold* pertama ( $b_1$ ) sebesar -1,118 bermakna bahwa untuk dapat menyelesaikan kategori 2 setelah kategori 1 peserta didik harus memiliki kemampuan minimal -1,118. Sedangkan tingkat kesukaran atau *threshold* kedua ( $b_2$ ) sebesar -1,094 bermakna bahwa untuk dapat menyelesaikan kategori 3 setelah kategori 2 peserta didik harus memiliki kemampuan minimal -1,094. Kesukaran item sebesar -1,118 berarti tingkat kesukaran item sedang. Kesukaran item sebesar -1,094 berarti tingkat kesukaran item sedang. *Location Indices (LI)* pada P3 sebesar -1,106 yang menunjukkan bahwa ringkasan tingkat kesukaran untuk P3 memiliki tingkat kesukaran sedang.

Pada D3 (item soal nomor 3) tingkat kesukaran atau *threshold (b)* sebesar 0,997 bermakna bahwa untuk menjawab benar item soal nomor 3 dengan probabilitas menjawab benar 50% diperlukan kemampuan minimal 0,997. Kesukaran item 0,997 berarti tingkat kesukaran item sedang. Untuk nilai tingkat kesukaran atau *threshold* masing-masing item di paket C dapat dilihat pada Tabel 25 di atas. Secara keseluruhan, tingkat kesukaran item pada paket C berada pada kategori sedang hingga sukar.

Estimasi karakteristik butir untuk paket D dapat dilihat pada Gambar 34 dan Tabel 26.

**Gambar 34.** *Item Characteristic Curve Paket D*



**Tabel 26.** *Nilai Parameter Butir Paket D*

Kode Item	No Item	$b_1$	$b_2$	$LI_{IRF}$	Kode Item	No Item	$b$
P1	4	-1,812	-0,808	-1,310	D1	1	-0,050
P3	6	-1,728	-1,999	-1,864	D2	2	0,309
P4	7	0,215	1,413	0,814	D3	3	-0,975
P5	8	2,057	1,046	1,551	D4	10	-0,336
P6	9	-1,622	0,223	-0,699	D5	11	1,323
P7	16	1,029	-0,190	0,420	D6	12	2,104
					D7	13	-0,122
					D8	14	1,877
					D9	15	-0,565

Berdasarkan hasil uji data data campuran paket D untuk tingkat kesukaran ( $b$ ) item dikotomus (D1 hingga D9) berada pada kisaran -0,975 hingga 2,104. Sedangkan untuk item-item

politomus (P1 hingga P7 tanpa P2 karena sudah dieliminasi akibat tidak valid secara konstruk) tingkat kesukaran pertama ( $b_1$ ) berada pada kisaran -1,812 hingga 2,057 dan tingkat kesukaran kedua ( $b_2$ ) berada pada kisaran -1,999 hingga 1,413.

Pada P4 (item soal nomor 7) tingkat kesukaran atau *threshold* pertama ( $b_1$ ) sebesar 0,215 bermakna bahwa untuk dapat menyelesaikan kategori 2 setelah kategori 1 peserta didik harus memiliki kemampuan minimal 0,215. Sedangkan tingkat kesukaran atau *threshold* kedua ( $b_2$ ) sebesar 1,413 bermakna bahwa untuk dapat menyelesaikan kategori 3 setelah kategori 2 peserta didik harus memiliki kemampuan minimal 1,413. Kesukaran item sebesar 0,215 berarti tingkat kesukaran item sedang. Kesukaran item sebesar 1,413 berarti tingkat kesukaran item sedang. *Location Indices (LI)* pada P4 sebesar 0,814 yang menunjukkan bahwa ringkasan tingkat kesukaran untuk P4 memiliki tingkat kesukaran sedang.

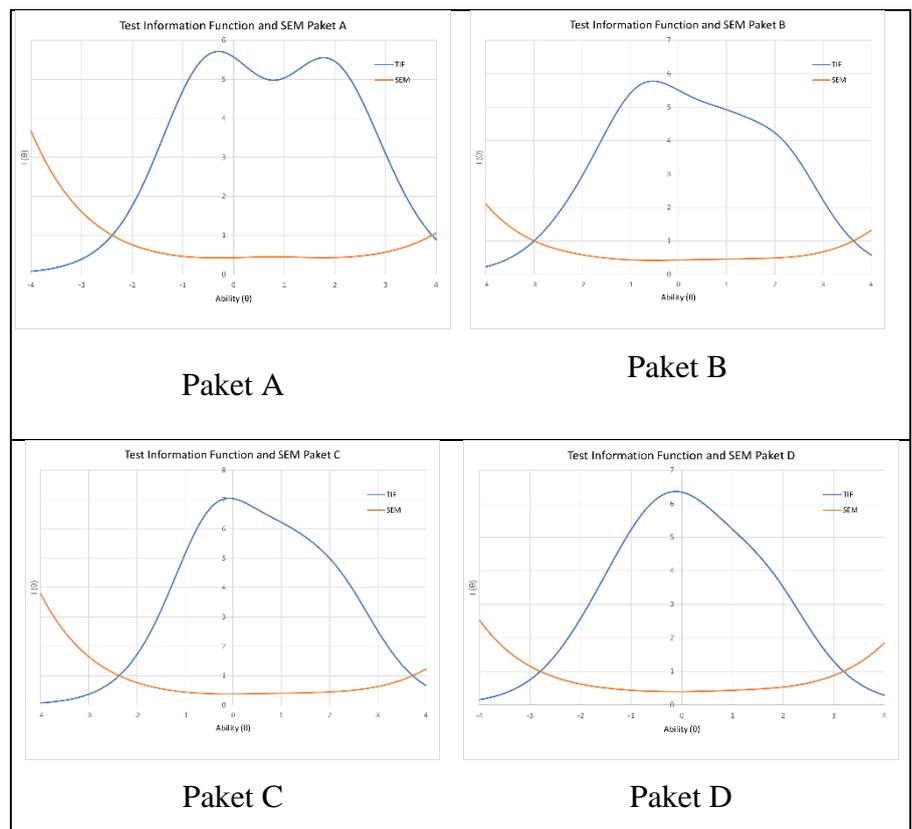
Pada D4 (item soal nomor 10) tingkat kesukaran atau *threshold (b)* sebesar -0,336 bermakna bahwa untuk menjawab benar item soal nomor 10 dengan probabilitas menjawab benar 50% diperlukan kemampuan minimal -0,336. Kesukaran item -0,336 berarti tingkat kesukaran item sedang. Untuk nilai tingkat kesukaran atau *threshold* masing-masing item di paket D dapat dilihat pada Tabel 26 di atas. Secara keseluruhan, tingkat kesukaran item pada paket D berada pada kategori sedang hingga sukar.

### c) Fungsi Informasi dan Kesalahan Pengukuran

Setiap butir soal memiliki fungsi informasinya sendiri sehingga dapat diketahui item mana yang cocok dengan model sehingga mudah untuk dipilih (Retnawati, 2014, p. 18). Fungsi informasi setiap butir berpengaruh pada nilai fungsi informasi tes, karena fungsi informasi tes merupakan jumlahan dari fungsi

butir penyusun tes (Hambleton & Swaminathan, 1985, p. 94). Sehingga apabila fungsi informasi dari setiap itemnya tinggi maka fungsi informasi dari tes akan tinggi, begitupun sebaliknya. Berikut adalah hasil fungsi informasi dan kesalahan pengukuran pada pengujian soal AKM numerasi.

**Gambar 35.** Hasil Fungsi Informasi dan Kesalahan Pengukuran



Hasil perhitungan TIF dan SEM pada soal AKM numerasi paket A menunjukkan bahwa nilai maksimum fungsi informasi perangkat soal paket A sebesar 5,71 yang dicapai oleh peserta didik dengan kemampuan  $\theta = -0,3$  dengan kesalahan pengukuran (SEM) sebesar 0,42. Hal tersebut bermakna bahwa informasi tes akan memberikan informasi terbesar dengan SEM terkecil apabila jika dikerjakan oleh peserta didik dengan kemampuan  $\theta = -0,3$ . Gambar 35 juga menunjukkan bahwa pada

paket A, titik perpotongan antara kurva TIF dan SEM berada pada  $\theta = -2,4$  dan  $\theta = 3,9$ . Maka dari itu soal AKM numerasi paket A reliabel apabila diujikan kepada peserta didik yang memiliki kemampuan pada rentan nilai  $-2,4 < \theta < 3,9$ .

Hasil perhitungan TIF dan SEM pada soal AKM numerasi paket B menunjukkan bahwa nilai maksimum fungsi informasi perangkat soal paket B sebesar 5,77 yang dicapai oleh peserta didik dengan kemampuan  $\theta = -0,5$  dengan kesalahan pengukuran (SEM) sebesar 0,42. Hal tersebut bermakna bahwa informasi tes akan memberikan informasi terbesar dengan SEM terkecil apabila jika dikerjakan oleh peserta didik dengan kemampuan  $\theta = -0,5$ . Gambar 35 juga menunjukkan bahwa pada paket B, titik perpotongan antara kurva TIF dan SEM berada pada  $\theta = -3$  dan  $\theta = 3,6$ . Maka dari itu soal AKM numerasi paket B reliabel apabila diujikan kepada peserta didik yang memiliki kemampuan pada rentan nilai  $-3,0 < \theta < 3,6$ .

Hasil perhitungan TIF dan SEM pada soal AKM numerasi paket C menunjukkan bahwa nilai maksimum fungsi informasi perangkat soal paket C sebesar 7,04 yang dicapai oleh peserta didik dengan kemampuan  $\theta = -0,1$  dengan kesalahan pengukuran (SEM) sebesar 0,38. Hal tersebut bermakna bahwa informasi tes akan memberikan informasi terbesar dengan SEM terkecil apabila jika dikerjakan oleh peserta didik dengan kemampuan  $\theta = -0,1$ . Gambar 35 juga menunjukkan bahwa pada paket C, titik perpotongan antara kurva TIF dan SEM berada pada  $\theta = -2,4$  dan  $\theta = 3,7$ . Maka dari itu soal AKM numerasi paket C reliabel apabila diujikan kepada peserta didik yang memiliki kemampuan pada rentan nilai  $-2,4 < \theta < 3,7$ .

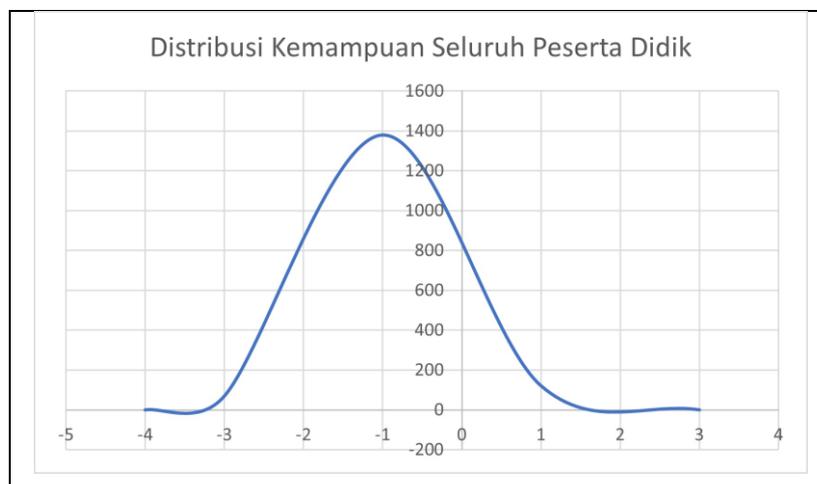
Hasil perhitungan TIF dan SEM pada soal AKM numerasi paket D menunjukkan bahwa nilai maksimum fungsi informasi perangkat soal paket D sebesar 6,36 yang dicapai oleh peserta

didik dengan kemampuan  $\theta = -0,1$  dengan kesalahan pengukuran (SEM) sebesar 0,40. Hal tersebut bermakna bahwa informasi tes akan memberikan informasi terbesar dengan SEM terkecil apabila jika dikerjakan oleh peserta didik dengan kemampuan  $\theta = -0,1$ . Gambar 35 juga menunjukkan bahwa pada paket D, titik perpotongan antara kurva TIF dan SEM berada pada  $\theta = -2,8$  dan  $\theta = 3,2$ . Maka dari itu soal AKM numerasi paket D reliabel apabila diujikan kepada peserta didik yang memiliki kemampuan pada rentan nilai  $-2,8 < \theta < 3,2$ .

#### d) Estimasi Parameter Kemampuan

Estimasi kemampuan peserta didik dari data hasil uji coba soal AKM numerasi kelas XI SMA dianalisis menggunakan program R yang kemudian dilanjutkan dengan *Ms. Excel*. Metode estimasi parameter kemampuan dalam penelitian ini menggunakan EAP karena menurut Uysal et al. (2022, p. 147) untuk struktural unidimensi EAP merupakan metode yang paling umum digunakan. Disamping itu diperkuat oleh penelitian Sass et al. (2008, p. 79) yang menyatakan bahwa EAP estimasi parameter kemampuannya lebih akurat dibanding metode MLE. Hasil analisis kemampuan peserta didik menggunakan IRT dapat dilihat pada Gambar 36.

**Gambar 36.** Grafik Distribusi Kemampuan Peserta Didik



Pada IRT, kemampuan seseorang yang terdefiniskan sebagai *ability* ( $\theta$ ) secara teoritis memiliki rentan nilai dari  $-\infty$  hingga  $\infty$  (Hambleton et al., 1991, p. 80). Namun secara praktis rentan nilai baku *ability* berada pada rentan nilai -4 hingga 4 (Naga, 1992, p. 207). Level kategori kemampuan (*ability*) oleh Hikamudin (2017, p. 11) dibagi menjadi lima kategori yakni sangat rendah untuk  $\theta \leq -3$ , rendah untuk  $-3 < \theta \leq -1$ , sedang untuk  $-1 < \theta \leq 1$ , tinggi untuk  $1 < \theta \leq 3$ , dan sangat tinggi untuk  $\theta > 3$ . Berdasarkan pengkategorian tersebut dan terilustrasikan pada Gambar 36, maka kemampuan (*ability*) peserta didik dalam menyelesaikan soal AKM numerasi pada rentan nilai -1,52 hingga 2,83. Kemampuan (*ability*) peserta didik berada pada kategori rendah hingga tinggi.

Hasil nilai *ability* kemudian dikonversi ke dalam skala 0 hingga 100 lazimnya pada penilaian di sekolah. Rumus *T-Score* digunakan untuk mengkonversi dari nilai *ability* ke nilai dengan skala 0 hingga 100. Setelah itu dilanjutkan pengkategorian berdasarkan tingkat kompetensi dan indeks numerasi peserta didik yang menjadi acuan pada rapor pendidikan AKM oleh Kemendikbudristek,. Disamping itu juga yang hasilnya dapat dilihat pada Tabel berikut.

**Tabel 27.** *Indeks Numerasi Peserta Didik SMA Kelas XI*

<b>Interval Kemampuan</b>	<b>Tingkat Kompetensi</b>	<b>Frekuensi</b>	<b>Persentase (%)</b>	<b>Bobot</b>	<b>Indeks</b>
			<b>b</b>	<b>c</b>	<b>b*c</b>
$\theta > 67,4$	Mahir	72	4,6	3	0,138
$56,5 < \theta \leq 67,4$	Cakap	169	10,8	2	0,216
$45,6 < \theta \leq 56,5$	Dasar	844	53,7	1,5	0,8055
$\theta \leq 45,6$	Perlu intervensi khusus	485	30,9	1	0,309
<b>Jumlah</b>		<b>1570</b>	<b>100</b>		<b>1,4685</b>

Tabel 27 menjelaskan bahwa indeks numerasi secara keseluruhan untuk peserta didik SMA kelas XI sebesar 1,4685 masuk pada kategori di bawah kompetensi minimum, dimana kurang dari 50% peserta didik telah mencapai kompetensi minimum untuk numerasi. Namun jika dipetakan tingkat kompetensinya, peserta didik dominan pada tingkat kompetensi “Dasar” sebesar 53,7%. Sedangkan persentase tingkat kompetensi yang lain sebesar 30.9% untuk tingkat kompetensi “Memerlukan intervensi khusus”, sebesar 10,8% untuk tingkat kompetensi “Cakap”, dan sebesar 4,6% untuk tingkat kompetensi “Mahir”.

**d. Penyetaraan**

Uji tingkat kesetaraan paket soal sama halnya dengan menguji apakah antar paket soal paralel atau tidak paralel. Apabila dua paket soal paralel, maka tidak perlu dilakukan proses penyetaraan. Sedangkan apabila dua paket soal tidak paralel, maka proses penyetaraan perlu dilakukan agar tidak merugikan kelompok tertentu dalam penentuan keputusan (Hambleton & Swaminathan, 1985).

Menurut Allen & Yen (1979, p. 59) dua paket soal dikatakan paralel jika mempunyai rata-rata dan varians yang sama. Uji beda rata-rata dilakukan dengan analisis varians, uji *Tukey* untuk uji pasang dan uji homogenitas varians dengan uji *Levene* (Hinkle et al., 1979, pp. 245–272).

Uji beda rerata menunjukkan nilai  $F = 9,794$  dengan taraf signifikansi 0,000. Hal tersebut menunjukkan bahwa keempat paket soal berbeda secara signifikan. Dengan hasil tersebut maka dilakukan uji berikutnya yakni uji pasang untuk mengetahui pasangan paket soal mana yang berbeda secara signifikan. Berikut adalah hasil uji pasang dari keempat paket soal.

**Tabel 28.** Hasil Uji Pasang (*Uji Tukey*)

	Beda Rerata	Signifikansi
Paket A – Paket B	-0,86739	0,024
Paket A – Paket C	-1,53047	0,000
Paket A – Paket D	-1,31746	0,000
Paket B – paket C	-0,66307	0,134
Paket B – Paket D	-0,45007	0,451
Paket C – Paket D	0,21300	0,899

Tabel 28 menunjukkan bahwa nilai signifikansi pada pasangan Paket A – Paket B, Paket A – Paket C, dan Paket A – Paket D kurang dari 0,05. Hal tersebut menunjukkan bahwa ketiga pasang tersebut terjadi perbedaan rerata secara signifikan. Sedangkan pasangan Paket B – Paket C, Paket B – Paket D, dan Paket C – Paket D nilai signifikansinya lebih dari 0,05 yang menunjukkan tidak terjadi perbedaan rerata secara signifikan.

Uji *Levene* dilakukan untuk menguji homogenitas varians menunjukkan nilai statistik dengan taraf signifikansi 0,003 kurang dari 0,05. Hal tersebut menunjukkan varians tidak homogen. Secara lengkap hasil dapat dilihat di Lampiran 2C. Untuk tiga pasangan paket

yang tidak terjadi perbedaan rerata secara signifikan dilakukan uji lanjut untuk mengetahui homogenitas varians. Hasil uji *Levene* menunjukkan bahwa pasangan paket B – Paket C dan Paket C – Paket D berturut-turut nilai signifikansinya kurang dari 0,05 yang menunjukkan bahwa varians tidak homogen. Sedangkan pasangan paket B – Paket D nilai signifikansinya lebih dari 0,05 yang menunjukkan varians yang homogen. Hasil uji beda rerata dan uji homogenitas kesetaraan paket soal dapat dilihat pada Tabel 29.

**Tabel 29.** Hasil Uji Kesetaraan Keempat Paket Soal

Aspek Uji	A - B	A - C	A - D	B - C	B - D	C - D	A - B - C - D
Rerata	berbeda	berbeda	berbeda	tidak berbeda	tidak berbeda	tidak berbeda	berbeda
Varians	tidak sama	tidak sama	tidak sama	tidak sama	sama	tidak sama	tidak sama
	tidak paralel	tidak paralel	tidak paralel	paralel	paralel	paralel	tidak paralel

Berdasarkan Tabel 29, pasangan paket soal yang tidak paralel dilakukan proses lanjut yakni penyetaraan. Desain penyetaraan yang digunakan adalah desain tes dengan butir bersama (*anchor*). Berikut adalah sebaran *anchor*, jumlah butir soal, dan jumlah responden pada masing-masing paket soal.

**Tabel 30.** Sebaran Anchor dan Jumlah Responden

Paket Soal	Anchor	Jumlah butir soal	Jumlah Responden
A	10, 11, 12, 13	16	392
B	10, 11, 12, 13	16	395
C	10, 11, 12, 13	17	386
D	10, 11, 12, 13	15	397

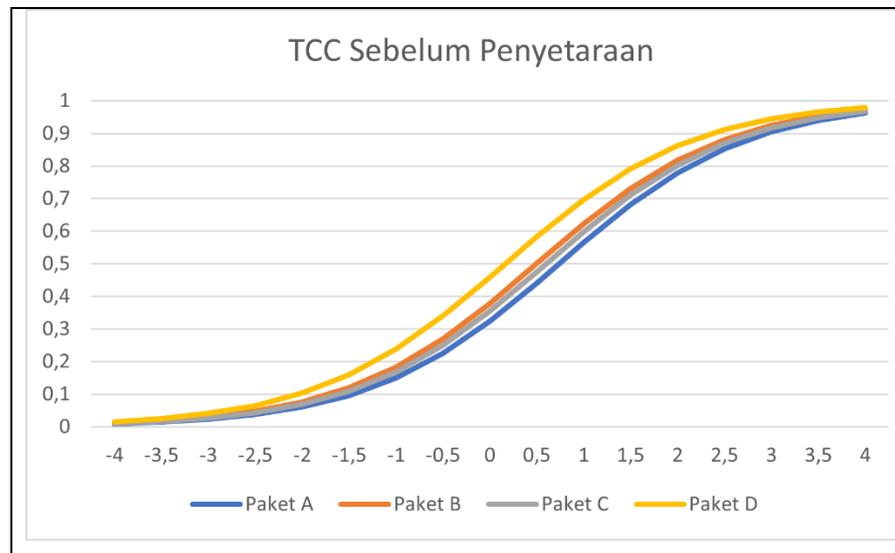
Analisis pertama yang dilakukan dalam proses penyetaraan adalah analisis karakteristik tes berdasarkan teori respon butir. Analisis ini sudah dilakukan sebelumnya menggunakan model *Rasch*. Berikut adalah statistik deskriptif tingkat kesukaran dari keempat paket soal.

**Tabel 31.** *Statistik Deskriptif Tingkat Kesukaran Empat Paket Soal*

	<b>Paket A</b>	<b>Paket B</b>	<b>Paket C</b>	<b>Paket D</b>
Rata-rata	0,740	0,497	0,599	0,165
Simpangan baku	1,342	1,154	1,176	1,187
Maksimum	2,825	2,417	2,538	2,104
Minimum	-1,010	-1,087	-1,106	-1,864

Tabel 31 menunjukkan bahwa dari keempat paket soal, paket A memiliki rata-rata tingkat kesukaran yang paling tinggi dibandingkan dengan paket lainnya dengan tingkat kesukaran maksimum sebesar 2,825. Sedangkan paket D memiliki rata-rata tingkat kesukaran yang paling rendah dibandingkan dengan paket lainnya dengan tingkat kesukaran minimum sebesar -1,864. Meskipun *Test Characteristics Curve* (TCC) dari keempat paket memiliki karakteristik yang relatif setara, namun perbedaan parameter butir berpotensi menghasilkan skor kemampuan (*ability*) yang berbeda di seluruh paket soal. Berikut adalah gambaran TCC dari keempat paket soal sebelum dilakukan penyetaraan.

**Gambar 37.** TCC Keempat Paket Soal Sebelum Penyetaraan



Analisis kedua adalah mengetahui konstanta hasil penyetaraan menggunakan salah satu metode penyetaraan yakni metode Rerata & Rerata. Penyetaraan dilakukan hanya pada paket B, paket C, dan paket D yang disetarakan ke paket A. Berikut adalah hasil konstanta penyetaraan yang telah diestimasi menggunakan metode Rerata & Rerata.

**Tabel 32.** Hasil Konstanta Penyetaraan Metode Rerata & Rerata

<i>Link Test</i>	<b>Konstanta</b>	<b>Nilai Estimasi</b>
Paket B – A	$\alpha$	1,000
	$\beta$	0,244
Paket C – A	$\alpha$	1,000
	$\beta$	0,141
Paket D – A	$\alpha$	1,000
	$\beta$	0,575

Tabel 32 menunjukkan bahwa konstanta *alpha* seluruhnya bernilai 1,000 sebagai akibat dari model *Rasch* yang digunakan pada analisis parameter butir dimana nilai daya beda sebesar 1,000. Kemudian untuk nilai konstanta *betha* bervariasi dengan interval nilai

dari 0,141 hingga 0,575. Konstanta penyetaraan yang telah terestimasi kemudian digunakan untuk mentransformasi skor awal menjadi skor baru setelah tes disetarakan. Berikut adalah persamaan kesetaraan skor ( $\theta$ ) metode Rerata & Rerata.

**Tabel 33.** *Persamaan Kesetaraan Skor Metode Rerata & Rerata*

<i>Link Test</i>	<b>Persamaan Kesetaraan Skor</b>
Paket B – A	$\theta_{ba}^* = \theta_b + 0,244$
Paket C – A	$\theta_{ca}^* = \theta_c + 0,141$
Paket D – A	$\theta_{da}^* = \theta_d + 0,575$

Persamaan kesetaraan skor pada Tabel 33 digunakan untuk mentransformasi *ability* ( $\theta$ ) peserta tes dengan paket B, paket C, dan paket D ke *ability* baru setelah disetarakan ke paket A. Berikut adalah estimasi *ability* ( $\theta$ ) setelah ketiga paket disetarakan ke paket A.

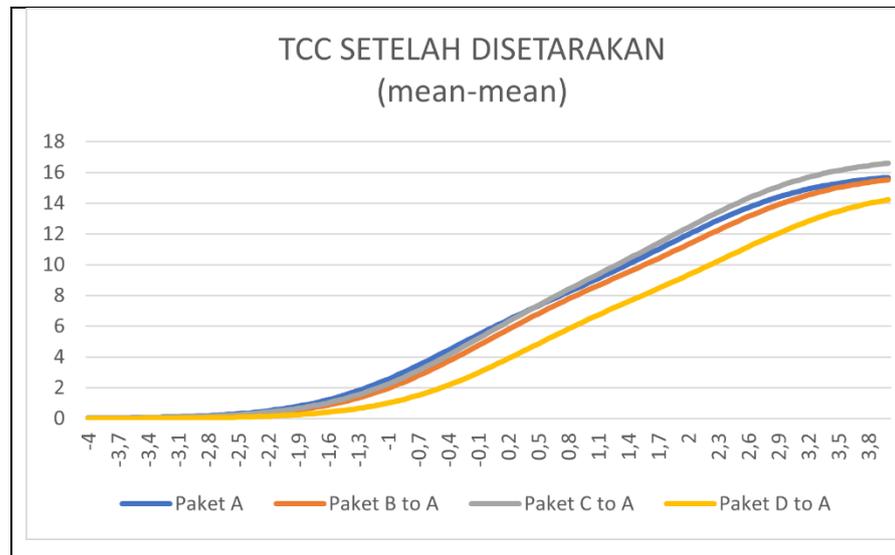
**Tabel 34.** *Hasil Skor Penyetaraan untuk Beberapa Ability*

<b>Ability (<math>\theta</math>)</b>	<b>Paket B to A</b>	<b>Paket C to A</b>	<b>Paket D to A</b>
-4	0,001074	0,001212	0,000499
-3	0,005807	0,006535	0,002712
-2	0,02977	0,033133	0,014356
-1	0,123192	0,13234	0,067113
0	0,319	0,327785	0,218374
1	0,522698	0,532874	0,429362
2	0,708092	0,731203	0,622164
3	0,883186	0,899919	0,821599
4	0,970437	0,975731	0,948357

Tabel 34 menunjukkan bahwa dengan metode Rerata & Rerata, ketika peserta didik berada pada  $\theta = -4$  pada paket B, itu setara dengan  $\theta = 0,001074$  setelah paket B disetarakan dengan paket A. hal tersebut berlaku ketika paket C dan paket D disetarakan dengan paket A  $\theta = -4$  akan setara dengan 0,001212 dan 0,000499. Interpretasi tersebut

berlaku untuk nilai  $\theta$  lainnya. Hasil tersebut kemudian diperkuat dengan TCC dari ketiga paket yang telah disetarakan ke paket A menggunakan metode Rerata & Rerata.

**Gambar 38.** TCC Setelah Penyetaraan

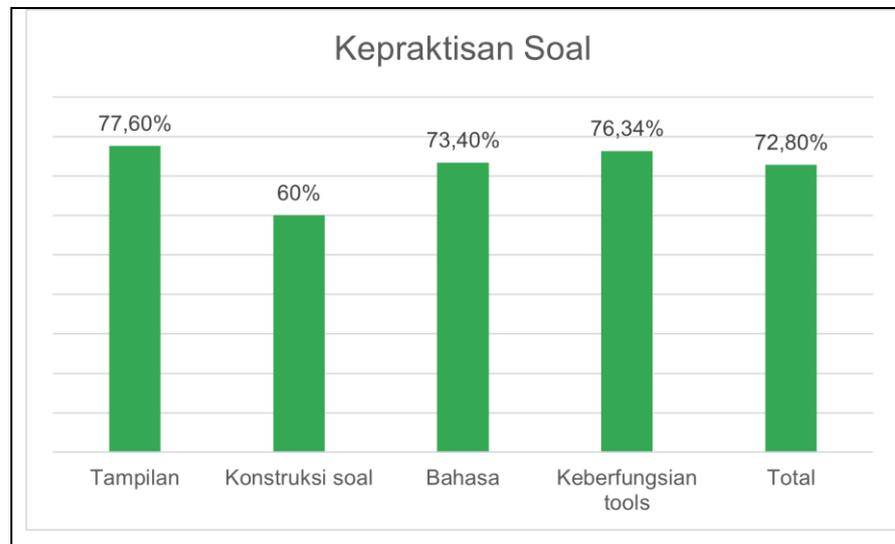


Gambar 38 menunjukkan bahwa penyetaraan dengan metode Rerata & Rerata menghasilkan kurva yang saling berdekatan untuk paket B dan C, namun tidak untuk paket D. Hal tersebut mengindikasikan bahwa untuk paket B dan paket C hampir memiliki karakteristik tingkat kesukaran yang setara satu sama lain, namun untuk paket D tidak memiliki tingkat kesukaran yang setara dengan paket lainnya.

**e. Keefektifan Produk**

Keefektifan produk yang dimaksud dalam bagian ini merupakan keefektifan dari soal AKM numerasi berbasis komputer yang telah dikembangkan. Efektifitas yang diharapkan mencakup dua hal yakni kepraktisan soal dan kemanfaatan soal. Analisis kepraktisan soal diperoleh dari respon peserta didik dari angket yang telah dibagikan bersamaan dengan penyelesaian soal AKM numerasi. Sebanyak 1446 peserta didik mengisi angket yang telah dibagikan. Hasil respon peserta didik terhadap kepraktisan soal dapat dilihat pada Gambar 38.

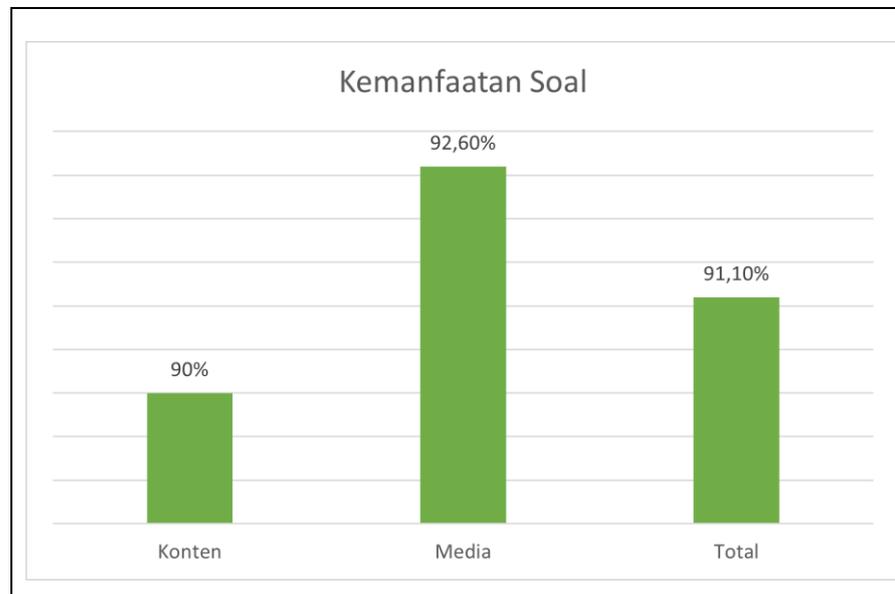
**Gambar 39.** Respon Peserta Didik Terhadap Kepraktisan Soal



Berdasarkan Gambar 39 diperoleh persentase kepraktisan soal masing-masing aspek dimana 77,60% aspek tampilan berkategori praktis digunakan, pada aspek konstruksi soal persentase sebesar 60% berkategori cukup praktis untuk digunakan, pada aspek bahasa persentase sebesar 73,40% berkategori praktis untuk digunakan, pada aspek keberfungsian *tools* sebesar 76,34% berkategori praktis untuk digunakan. Secara keseluruhan dari respon peserta didik terhadap kepraktisan soal diperoleh persentase sebesar 72,80% dengan kategori praktis untuk digunakan.

Selanjutnya analisis kemanfaatan soal diperoleh dari respon pendidik pengampu Matematika kelas XI dari angket yang telah disebarkan bersamaan dengan pengambilan data pada peserta didik. hasil respon pendidik terhadap kemanfaatan soal dapat dilihat pada Gambar 40.

**Gambar 40.** *Respon Pendidik Terhadap Kemanfaatan Soal*



Berdasarkan Gambar 40 diperoleh persentase kemanfaatan soal masing-masing aspek dimana 90% aspek tampilan berkategori sangat manfaat untuk dipergunakan, dan pada aspek media persentase sebesar 92,60% berkategori sangat manfaat untuk dipergunakan. Secara keseluruhan dari respon pendidik terhadap kemanfaatan soal diperoleh persentase sebesar 91,10% dengan kategori sangat manfaat untuk digunakan.

### **3. Revisi dan Produk Akhir**

#### **a. Revisi Instrumen Berdasarkan Penilaian Validator**

Hasil validasi soal AKM numerasi dan media CBT yang telah dikembangkan melalui enam orang ahli untuk validasi soal dan dua orang ahli untuk validasi media dinyatakan valid keduanya namun perlu dilakukan revisi. Dalam proses validasi, seluruh validator memberik saran dan masukan terhadap soal numerasi dan media CBT yang telah dikembangkan. Maka dari itu peneliti melakukan revisi terhadap soal numerasi dan media CBT. Secara umum saran dan masukan yang diberikan oleh validator adalah sebagai berikut.

1) Saran dan masukan validasi soal AKM numerasi, meliputi:

- a) Penggunaan dan konsistensi istilah yang kurang lazim digunakan pada perangkat soal;
  - b) Kurangnya ilustrasi tambahan dan kebermaknaan stimulus soal;
  - c) Beberapa soal yang petunjuk atau intruksinya kurang jelas;
  - d) Tata kalimat yang kurang sesuai EYD;
  - e) Perlu peninjauan ulang kesesuaian item soal dengan konteks soal dan kesukaran yang ditetapkan.
- 2) Saran dan masukan validasi media CBT, meliputi:
- a) Penyesuaian kembali pemilihan *background* dengan tema;
  - b) Pengadaan petunjuk media tambahan untuk pendidik dan administrator;
  - c) Penegasan tujuan tes dan keterangan petunjuk pengisian Quiz dan Kuesioner yang ada pada media;
  - d) Perbaikan resolusi gambar yang tertera dalam soal.

Secara keseluruhan item-item soal direvisi untuk diperbaiki stimulus soal dan beberapa dilakukan perbaikan sesuai saran dan masukan validator. Adapun beberapa item soal sebelum dan sesudah direvisi tersajikan pada bagian berikut.

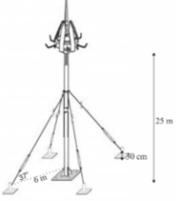
### Item soal nomor 1 untuk paket A

Sebelum diperbaiki:

**Tiang Penyangga Penangkal Petir**

Tiang penyangga penangkal petir merupakan pipa penegak yang berfungsi untuk menaikkan posisi *head* terminal penangkal petir elektrostatis agar mendapatkan level yang tinggi, karena semakin tinggi level penangkal petir maka sangat berarti untuk mendapatkan radius perlindungan.  
 Sumber: [pakarpetir.co.id](http://pakarpetir.co.id)

Pada umumnya tiang penyangga penangkal petir diletakkan pada lantai tertinggi suatu bangunan.  
 Suatu perusahaan memproduksi tiang penyangga penangkal petir dengan 4 kawat penyangga yang dipasang simetris pada tiang, tampak pada gambar berikut.



Total kawat yang diperlukan untuk menyangga tiang adalah ....  
 (Keterangan:  $\sin 37^\circ = \frac{3}{5}$ ;  $\cos 37^\circ = \frac{4}{5}$ ; dan  $\tan 37^\circ = \frac{3}{4}$ )

- A. 14,4 m
- B. 19,2 m
- C. 30 m
- D. 32 m
- E. 40 m

Setelah diperbaiki:

**Tiang Penyangga Penangkal Petir**

Tiang penyangga penangkal petir merupakan pipa penegak yang berfungsi untuk menaikkan posisi *head* terminal penangkal petir elektrostatik agar mendapatkan level yang tinggi, karena semakin tinggi level penangkal petir maka sangat berarti untuk mendapatkan radius perlindungan (Sumber: pakarpetir.co.id). Tiang penyangga penangkal petir memiliki 4 kawat penyangga yang dipasang simetris pada tiang utamanya, dan tampak pada gambar berikut.

Panjang kawat yang diperlukan untuk menyangga tiang adalah ....  
(Keterangan:  $\sin 37^\circ = \frac{3}{5}$ ;  $\cos 37^\circ = \frac{4}{5}$ ; dan  $\tan 37^\circ = \frac{3}{4}$ )

A. 14,4 m  
B. 19,2 m  
C. 30 m  
D. 32 m  
E. 40 m

Item soal nomor 4 untuk paket B

Sebelum diperbaiki:

**Boneka Salju**

Swiss salah satu negara yang menjadi incaran wisatawan untuk dikunjungi saat musim salju, karena 60% wilayah Swiss merupakan bagian dari Pegunungan Alpen.

Niko dan keluarga akan mengunjungi Swiss saat musim salju tiba. Salah satu keputusan akan mengunjungi negara tersebut adalah kedua anaknya ingin membuat bola salju secara langsung. Sebegitu antusiasnya dengan liburan tersebut, jauh-jauh hari sebelum keberangkatan kedua anak Niko membuat rancangan ukuran boneka salju yang akan mereka buat.

Apabila mereka akan membuat boneka salju yang terdiri atas dua bola dengan perbandingan diameternya 1 : 3 dan tinggi boneka adalah 112 cm, maka tentukan kebenaran dari pernyataan-pernyataan di bawah ini dengan memberi tanda centang pada kolom **Benar** atau **Salah** untuk setiap pernyataan!

Pernyataan	Benar	Salah
Isian kepala boneka salju adalah $\frac{34496}{3} \text{ cm}^3$		
Isian badan boneka salju adalah $174636 \text{ cm}^3$		
Kedua tangan boneka salju terbuat dari ranting sepanjang 18 cm jika ukuran per masing-masing tangan $\frac{3}{7}$ dari jari-jari badan boneka salju		
Jika topi boneka salju berbentuk kerucut dengan tinggi 14 cm, maka luas kain yang diperlukan adalah $616\sqrt{2} \text{ cm}^2$		

Setelah diperbaiki:

**Boneka Salju**

Swiss salah satu negara yang menjadi incaran wisatawan untuk dikunjungi saat musim salju, karena 60% wilayah Swiss merupakan bagian dari Pegunungan Alpen. Umumnya wisatawan akan membuat berbagai bentuk dan ukuran boneka salju saat mengunjungi Swiss. Apabila suatu boneka salju dibuat dari dua bola dengan perbandingan diameternya 1 : 3 dan tinggi boneka 112 cm, maka tentukan kebenaran dari pernyataan-pernyataan di bawah ini dengan memberi tanda centang (✓) pada kolom **Benar** atau **Salah** untuk setiap pernyataan!

No	Pernyataan	Benar	Salah
1	Volume kepala boneka salju adalah $\frac{34496}{3} \text{ cm}^3$		
2	Volume badan boneka salju adalah $174636 \text{ cm}^3$		
3	Kedua tangan boneka salju terbuat dari ranting sepanjang 18 cm jika ukuran per masing-masing tangan $\frac{3}{7}$ dari jari-jari badan boneka salju		
4	Jika topi boneka salju berbentuk kerucut dengan tinggi 14 cm, maka luas kain yang diperlukan adalah $616\sqrt{2} \text{ cm}^2$		

### Item soal nomor 8 untuk paket C

Sebelum diperbaiki:

**Stand Microphone**

Seorang penyanyi tentu tidak asing lagi dengan *stand microphone*. Benda tersebut kerap kali dipakai para penyanyi yang sedang ingin bergerak bebas tanpa memegang *mic*. *Stand microphone* bisa digunakan siapa saja tanpa terkecuali, karena benda tersebut dapat diatur ketinggiannya mengikuti tinggi seseorang yang menggunakannya.

Sebagai seorang *content creator*, Fina sering membuat konten bernyanyi dengan bantuan *stand microphone* yang dimilikinya, baik itu sendirian atau mengajak kolaborasi dengan teman-temannya. Oleh karena itu Fina memiliki lebih dari satu *stand microphone*.

Berikut adalah gambaran *stand microphone* yang dimiliki Fina, serta posisi ideal ia dan teman-temannya saat bernyanyi.



Apabila posisi *mic* selalu di depan mulut para penyanyi dan diasumsikan bahwa mulut hingga ujung kepala berjarak 15 cm, maka jodohkanlah pernyataan yang berkaitan dengan tinggi teman-teman Fina yang diajak kolaborasi bernyanyi terhadap sudut kemiringan ( $\theta$ ) *stand microphone* agar posisi *mic* tepat di depan mulut penyanyi!

Tinggi Bono 179 cm (a) ●	● (i) $\theta = \tan^{-1} \left( \frac{4}{30} \right)$
Tinggi Rubi 149 cm (b) ●	● (ii) $\theta = \tan^{-1} \left( \frac{15}{30} \right)$
Tinggi Joko 185 cm (c) ●	● (iii) $\theta = \tan^{-1} \left( \frac{23}{30} \right)$
	● (iv) $\theta = \tan^{-1} \left( \frac{40}{30} \right)$
	● (v) $\theta = \tan^{-1} \left( \frac{55}{30} \right)$

Setelah diperbaiki:

**Stand Microphone**

Fina adalah seorang *content creator* yang sering membuat konten bernyanyi dengan bantuan *stand microphone* yang dimilikinya, baik itu sendirian atau mengajak kolaborasi dengan teman-temannya. Oleh karena itu Fina memiliki lebih dari satu *stand microphone*. Berikut adalah ilustrasi gambar dalam dua dimensi *stand microphone* yang dimiliki Fina, serta posisi ideal ia dan teman-temannya saat bernyanyi.



Apabila posisi *mic* selalu di depan mulut para penyanyi dan diasumsikan bahwa mulut hingga ujung kepala berjarak 15 cm, maka pasangkanlah pernyataan-pernyataan yang berkaitan dengan tinggi teman-teman Fina yang diajak kolaborasi bernyanyi terhadap sudut kemiringan ( $\theta$ ) *stand microphone* agar posisi *mic* tepat di depan mulut penyanyi!

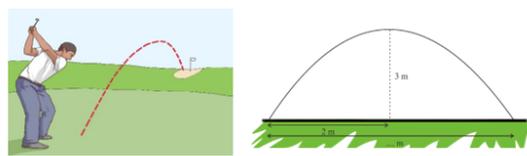
Tinggi Bono 179 cm (a) ●	● (i) $\theta = \tan^{-1} \left( \frac{4}{30} \right)$
Tinggi Rubi 149 cm (b) ●	● (ii) $\theta = \tan^{-1} \left( \frac{15}{30} \right)$
Tinggi Sinta 168 cm (c) ●	● (iii) $\theta = \tan^{-1} \left( \frac{23}{30} \right)$
Tinggi Joko 185 cm (d) ●	● (iv) $\theta = \tan^{-1} \left( \frac{34}{30} \right)$
Tinggi Joko 185 cm (d) ●	● (v) $\theta = \tan^{-1} \left( \frac{40}{30} \right)$
Tinggi Dewi 160 cm (e) ●	● (vi) $\theta = \tan^{-1} \left( \frac{55}{30} \right)$

### Item soal nomor 14 untuk paket D

Sebelum diperbaiki:

**Olahraga Golf**

Olahraga golf merupakan salah satu jenis olahraga di luar ruangan yang dimainkan secara perorangan atau tim yang saling berlomba-lomba memasukkan bola ke dalam lubang-lubang yang ada di lapangan dengan jumlah pukulan tersedikit mungkin. Olahraga Golf merupakan olahraga favorit dari Pak Hary. Setiap hari minggu pagi, dia meluangkan waktu untuk menyalurkan hobinya tersebut. Pada hari Minggu itu, pukulan pertama Pak Hary tidak sesuai target, atau tidak masuk ke lubang yang ada. Pukulan Pak Hary tersebut membentuk suatu fungsi kuadrat seperti gambar berikut.



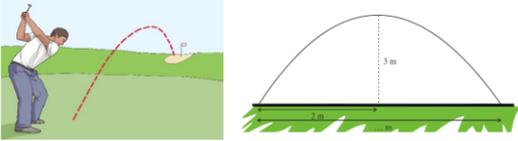
*Sumber: alatgolf.com*

Apabila lapangan golf terasumsikan sebagai suatu garis lurus, jarak Pak Hary memukul bola hingga bola terlempar dan mengenai rumput adalah ... m.

Setelah diperbaiki:

**Olahraga Golf**

Olahraga Golf merupakan olahraga favorit dari Pak Hary. Pada hari Minggu itu, pukulan pertama Pak Hary tidak sesuai target, atau tidak masuk ke lubang yang ada. Pukulan Pak Hary tersebut membentuk suatu pantulan seperti gambar berikut.



Sumber: alatgolf.com

Apabila lapangan golf terasumsikan sebagai suatu garis lurus, jarak Pak Hary memukul bola hingga bola terlempar dan mengenai rumput adalah ... m.

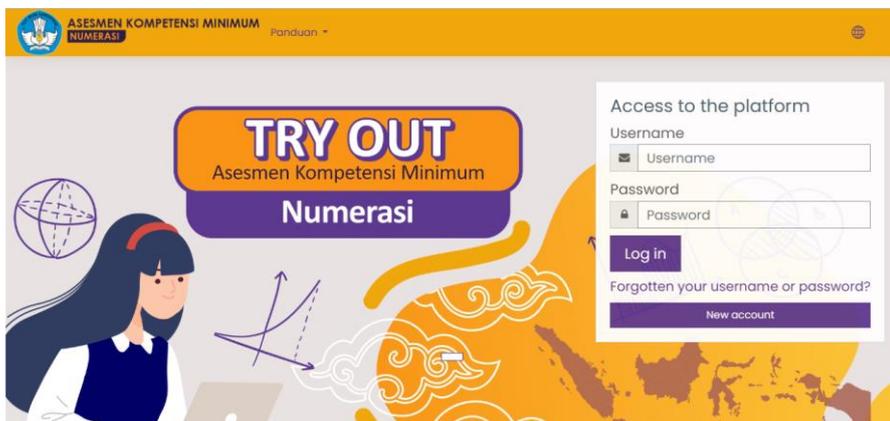
Sedangkan tampilan media CBT sebelum dan sesudah direvisi juga tersajikan pada bagian berikut, namun untuk lebih lengkapnya disajikan pada bagian berikut.

Perubahan pemilihan *background* yang tidak sesuai dengan tema

Sebelum diperbaiki:



Setelah diperbaiki:



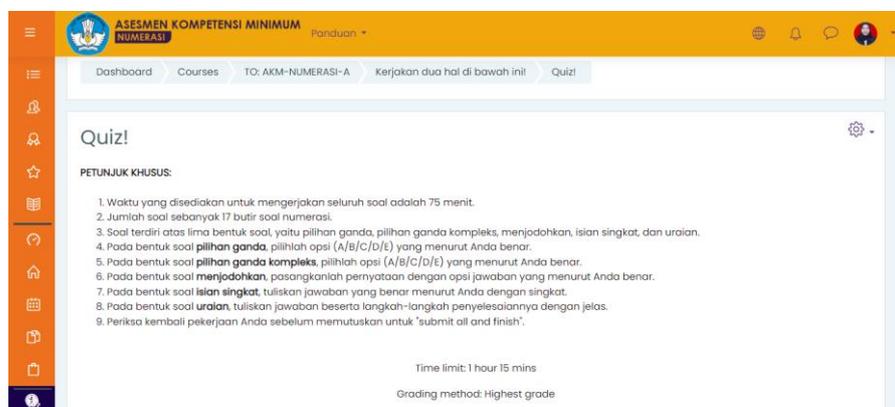
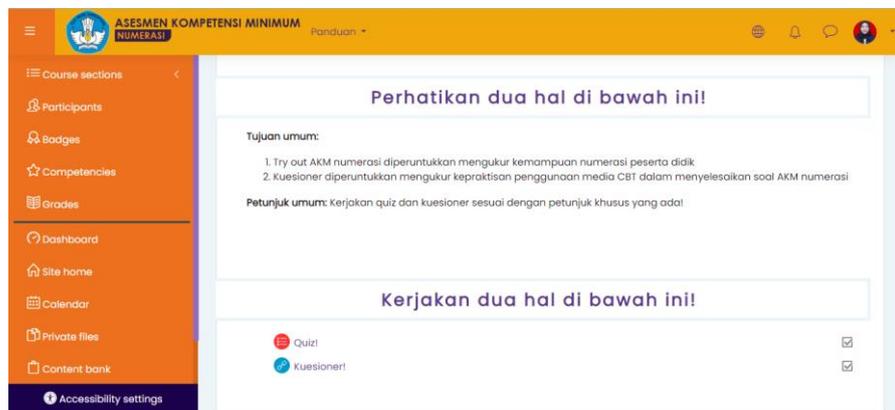
## Panambahan petunjuk penggunaan media

Awalnya petunjuk penggunaan hanya diperuntukkan untuk peserta didik saja, kemudian diperbaiki dengan menambah petunjuk bagi pendidik agar mudah menggunakan media ini.



## Penambahan tujuan tes dan keterangan petunjuk pengisian Quiz dan Kuesioner

Awalnya petunjuk pengerjaan soal dibuat secara umum dan ditampilkan pada gambar di bawah ini. Namun kemudian petunjuk diperinci sehingga secara umum terdapat di luar menu Quiz dan Kuesioner, dan secara khusus ada di dalam masing-masing menu.



### Perbaikan resolusi gambar pada soal

Awalnya ukuran gambar dibuat sama dengan ukuran gambar di *draft* instrumen soal, kemudian setelah mendapat masukan resolusi gambarnya diperbesar sehingga salah satunya nampak seperti gambar di bawah.



#### **b. Revisi Berdasarkan Uji Coba**

Produk soal AKM numerasi berbasis komputer yang dikembangkan sebanyak 54 item soal yang terdiri dari 14 soal aljabar, 18 soal geometri dan pengukuran, serta 22 soal data dan ketidakpastian. Keseluruhan item soal tersebut diujikan kepada peserta didik dengan membaginya menjadi 4 paket soal dengan terdapat 4 item *anchor* pada masing-masing paket, sehingga untuk paket A dan C terdapat 17 item soal serta paket B dan D terdapat 16 item soal. Hasil analisis uji coba menunjukkan bahwa ukuran sampel yang ditetapkan sudah mencukupi dengan nilai KMO dari masing-masing paket di atas 0,50. Setiap paket juga memiliki hanya satu nilai *eigen* yang dominan. Hal tersebut menunjukkan bahwa soal yang telah dikembangkan tepat untuk mengukur numerasi.

Berdasarkan hasil uji kecocokan item (*item fit*) dengan IRT, pada paket A terdapat dua item yang berkriteria *mix* (*fit* pada nilai *infit* namun *overfit* pada nilai *outfit*). Pada paket B terdapat satu item yang berkriteria *overfit* dan empat item yang berkriteria *mix* (*fit* pada nilai *infit* namun *overfit* pada nilai *outfit*). Pada paket C terdapat dua item

yang berkriteria *mix* (*fit* pada nilai *infit* namun *overfit* pada nilai *outfit*). Pada paket D terdapat satu item yang berkriteria *mix* (*fit* pada nilai *infit* namun *overfit* pada nilai *outfit*). Item-item tersebut dipertimbangkan untuk dipertahankan namun dilakukan revisi sesuai keperluan untuk menjaga konstruksi instrumen. Apabila item tersebut digugurkan, terdapat domain numerasi yang tidak terwakilkan karena item ada yang menjadi *anchor*. Keputusan tersebut diperkuat dengan oleh Susac et al. (2018, p. 3) yang menyatakan bahwa item yang memiliki nilai *infit* dan *outfit* pada rentan nilai 0,5 hingga 1,5 masih produktif dalam pengukuran sehingga bisa untuk dipertahankan. Untuk yang berkategori *mix* semuanya pada nilai *infit* nya di atas 0,75 sesuai dengan kategori *fit*. Item *mix* tersebut juga layak dipertahankan karena menurut Zi Yan & Heene (2021, p. 56) jika nilai *outfit* lebih sensitif terhadap *outlying* sehingga pertimbangan utama ada pada nilai *infit* daripada nilai *outfit*.

## **B. Kajian Produk Akhir**

### **1. Konstruksi Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA**

Instrumen soal AKM numerasi berbasis komputer yang dikembangkan mengadopsi model pengembangan Retnawati (2016, p. 3). Produk akhir berupa seperangkat tes yang dirancang khusus untuk melatih dan membiasakan bernumerasi melalui penyelesaian soal berbasis AKM. Dimana AKM sendiri merupakan salah satu pengujian yang ada pada Asesmen Nasional (AN) sekarang ini. Instrumen soal AKM numerasi disesuaikan dengan domain dan kompetensi yang diujikan pada AKM numerasi pada AN. Domain yang digunakan untuk pengembangan soal ini mencakup tiga domain yakni aljabar, geometri dan pengukuran, serta data dan ketidakpastian. Meninggalkan domain bilangan karena pada pengujian AKM numerasi sesuai AN untuk kelas XI domain bilangan tidak diujikan. Ketiga domain tersebut mencakup beberapa subdomain yakni, 1) persamaan dan pertaksamaan, 2) relasi dan fungsi (termasuk pola

bilangan), 3) bangun geometri, 4) data dan representasinya, dan 5) ketidakpastian dan peluang. Tes yang dikembangkan terdiri atas lima bentuk soal yakni bentuk pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian. Tes yang dihasilkan terdiri atas 4 paket soal dimana setiap paket soal terdiri atas 16-17 soal serta mewakili ketiga domain yang diuji. Soal yang dikembangkan merupakan soal dengan data nilai campuran yakni skala dikotomis dan politomis.

Instrumen soal yang dikembangkan telah melewati serangkaian tahap pengembangan sehingga diperoleh seperangkat instrumen yang valid, reliabel, dan memiliki kualitas butir yang baik untuk penggunaan lebih lanjut. Penelitian Faizah et al. (2019, p. 115) menyatakan bahwa pengembangan produk akan layak digunakan apabila mengikuti prosedur pengembangan secara sistematis. Hal tersebut diperkuat Gall et al. (2003, p. 569) yang menegaskan bahwa harus melewati langkah prosedur yang kompleks untuk mengembangkan suatu produk. Pada proses pengembangan soal AKM numerasi berbasis komputer ini lebih menekankan agar produk yang diperoleh sesuai dengan kerangka kerja AKM sehingga hasil produknya dapat digunakan, terlebih sebagai bahan referensi peserta didik untuk menghadapi AKM numerasi disetiap tahunnya.

## **2. Kualitas Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA**

Kualitas instrumen soal yang dikembangkan dalam penelitian ini dilihat dari lima aspek yakni validitas isi, validitas konstruk, dan kualitas butir. Berikut adalah pembahasan dari kelima aspek kualitas soal tersebut.

Validitas isi berkaitan dengan analisis rasional domain yang akan diukur untuk mengetahui representasi dengan kemampuan yang akan diukur (Abidin & Retnawati, 2019, p. 148). Pembuktian validitas isi dilakukan berdasarkan hasil penilaian para ahli (*expert judgment*). *Expert judgment* merupakan orang-orang yang memiliki keahlian pada bidangnya masing-masing, dalam hal ini bidang yang dimaksud adalah bidang

pengukuran, matematika, serta media. Penilaian validitas isi dilakukan secara kuantitatif dan kualitatif. Telaah secara kuantitatif menghasilkan nilai indeks *Aiken's V*, sedangkan telaah secara kualitatif menghasilkan masukan dan saran terhadap instrumen yang telah dikembangkan. Secara keseluruhan item soal yang dikembangkan memiliki nilai indeks *Aiken's V* > 0,78 yang menunjukkan bahwa item tersebut valid. Item-item soal yang valid menunjukkan kesesuaiannya dengan indikator, konten, dan konteks yang hendak diukur. Hal tersebut diperkuat dengan penelitian Degroote et al. (2020, p. 6) yang menyatakan bahwa instrumen soal valid secara isi apabila item-item soal merefleksikan apa yang hendak diukur, dengan konstruk soal yang jelas. Meskipun seluruh butir dikatakan valid namun tetap dilakukan perbaikan berdasarkan penilaian validasi isi secara kualitatif dari para *expert judgment*.

Validitas konstruk merupakan validitas yang berkaitan dengan tingkat skala yang mencerminkan dan bertindak sebagai konsep yang diukur (Faizah et al., 2019, p. 113). Dengan kata lain validitas konstruk dapat mengidentifikasi sejauh mana item soal dapat mengukur apa yang ingin benar-benar diukur. Pada penelitian ini validitas konstruk menggunakan *Confirmatory Factor Analysis* (CFA). CFA bersifat mengkonfirmasi data empiris agar sesuai model pada data. Dari hasil CFA seluruh item pada keempat paket, untuk nilai *loading factor* terdapat dua item yang kurang dari 0,3. Item soal tersebut adalah item nomor 3 di paket A dan nomor 5 di paket D. Kedua item tersebut tidak lolos dalam analisis validitas konstruk karena tidak mencerminkan kemampuan item soal dalam mengukur numerasi. Hal tersebut sesuai dengan penelitian Apriandi et al. (2022, p. 1495) dan Wijono & Mardapi (2016, p. 237) yang menyatakan bahwa nilai *loading factor* kurang dari 0,3 menunjukkan adanya ketidakcocokan butir dengan variabel yang diukur. Analisis ini memberikan bukti bahwa untuk paket B dan paket C seluruh item valid secara konstruk. Namun untuk paket A hanya 16 item yang valid secara konstruk, dan paket D hanya 15 item yang valid secara konstruk. Dengan

kata lain, item-item soal pada paket A yang lolos untuk dianalisis IRT sebanyak 16 item dan paket D sebanyak 15 item.

Uji kualitas butir diuji dengan analisis data campuran dengan estimasi model *Rasch*. Ditetapkannya model *Rasch* karena telah melewati uji kecocokan butir terhadap model PCM dan GPCM untuk data politomus, dan model *Rasch*, 2PL, dan 3PL untuk data dikotomus. Pertimbangan keputusan pada uji kecocokan butir adalah nilai *infit* dan *outfit*. Zi Yan & Heene (2021, p. 56) menjelaskan bahwa nilai *outfit* lebih sensitif terhadap *outlying* sehingga pertimbangan utama ada pada nilai *infit* daripada nilai *outfit*. Selain itu item akan tidak diprioritaskan apabila nilai *infit* atau *outfit* tidak pada rentang nilai 0,5 hingga 1,5 karena diluar rentang nilai tersebut item tidak produktif dalam pengukuran dan tidak bisa untuk dipertahankan (Susac et al., 2018, p. 3). Dengan pertimbangan tersebut dalam analisis kecocokan butir telah ditetapkan bahwa seluruh paket soal analisis lanjut IRT menggunakan model *Rasch*.

Pada penskoran dikotomus model *Rasch* merupakan kasus khusus dari model IRT 1-PL dimana daya beda seluruh item adalah 1 (Djidu et al., 2022, p. 107). Sedangkan untuk penskoran politomus model *Rasch* yang digunakan adalah PCM (*Partial Credit Model*) yang merupakan model respon item keluarga *Rasch* yang mampu menangani respon item lebih dari dua kategori (Djidu et al., 2022, pp. 143–144). Karena PCM merupakan perpanjangan langsung dari model *Rasch* sederhana dikotomus, maka daya beda seluruh item juga bernilai 1.

Pada penelitian ini untuk penskoran politomus dengan PCM memiliki tiga kategori respons (0, 1, dan 2), sehingga PCM akan memperkirakan dua kategori *threshold* ( $b_1, b_2$ ) yang berdekatan. Karena setiap item memiliki dua tingkat kesukaran item ( $b_1$  &  $b_2$ ), untuk penyimpulan tingkat kesukaran suatu item maka diperlukan *generalized item difficulty summaries* menggunakan estimasi *Location Indices* (LI) dengan pendekatan nilai *Item Response Function* (IRF). Ali et al. (2015, p. 10) menyatakan bahwa LI IRF lebih merepresentasikan tingkat kesukaran

untuk data politomus dibanding pendekatan nilai yang lain. Penggunaan model PCM dalam mengukur kualitas butir tes ini terlebih untuk data politomus didasari oleh penelitian yang dilakukan oleh Santoso et al. (2022, p. 11) dimana estimasi model fit menggunakan model PCM lebih bagus hasilnya dibandingkan model lain untuk pengujian instrumen politomus dengan tiga kategori respon (0, 1, dan 2) yang mana sama dengan penelitian ini. Sedangkan penggunaan model *Rasch* terlebih untuk data dikotomus didasari oleh penelitian yang dilakukan oleh Lia et al. (2020, p. 43) dan Sabekti & Khoirunnisa (2018, p. 69) yang mengkonfirmasi bahwa model *Rasch* lebih direkomendasikan untuk analisis kualitas butir pada pengembangan instrumen tes, yang mana sesuai dengan penelitian ini.

### **3. Kesetaraan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA**

Penyetaraan merupakan proses statistik yang digunakan untuk mengatur skor pada dua perangkat tes atau lebih sehingga skor-skornya dapat saling tukar (Kolen & Brennan, 2014). Definisi tersebut sejalan dengan tujuan yang dinyatakan oleh Retnawati (2014, p. 93) bahwa dari penyetaraan akan dihasilkan skor yang dapat saling menggantikan dari satu tes ke tes lainnya. Hal tersebut dapat terjadi karena adanya skor transformasi yang bertujuan agar skor pengukuran dari dua tes berada dalam skala yang sama. Pada penelitian ini desain penyetaraan yang digunakan adalah desain tes dengan butir bersama (*anchor*). Hal tersebut dikarenakan pada setiap paket soal terdapat *anchor* yang berada pada nomor soal yang sama. Menurut Retnawati (2014, p. 97) desain tes dengan butir bersama merupakan dua tes atau lebih yang diberikan pada dua grup atau lebih yang berbeda, namun instrumen tes memiliki *common item* atau butir bersama (kemudian diistilahkan dengan *anchor*). Penelitian Yusron et al. (2020, pp. 3–4) menegaskan bahwa instrumen tes yang didalamnya terdapat *anchor* tidak mencirikan adanya kesetaraan, oleh karena itu masih diperlukan proses penyetaraan.

Proses penyetaraan dilakukan secara IRT dengan metode yang digunakan adalah metode Rerata & Rerata. Pada IRT, penyetaraan merupakan proses menempatkan skor dari dua atau lebih bentuk tes paralel ke skala skor umum (Hambleton et al., 1991; Yusron et al., 2020, p. 3). Sedangkan metode Rerata & Rerata memiliki tujuan untuk menentukan konstanta yang melibatkan dua parameter yakni parameter data pembeda dan tingkat kesukaran item (Loyd & Hoover, 1980, p. 180). Penyetaraan konstanta  $\alpha$  dan  $\beta$  dapat dihitung dengan rerata dari parameter yang terlibat. Karena pada penelitian ini estimasi parameter dilakukan dengan model *Rasch*, maka untuk parameter daya beda nilainya sebesar 1,00.

Pada penelitian ini penyetaraan yang dilakukan dengan menyetarakan masing-masing paket B, paket C, dan paket D ke paket A sehingga persamaan kesetaraan skor antar paket yakni  $\theta_{ba}^* = \theta_b + 0,244$ ;  $\theta_{ca}^* = \theta_c + 0,141$ ;  $\theta_{da}^* = \theta_d + 0,575$ . Dari ketiga persamaan kesetaraan tersebut kemudian diperoleh *Test Characteristic Curve* (TCC) yang menginformasikan bahwa pada penyetaraan instrumen soal AKM numerasi ini kurva yang saling berdekatan adalah antara paket B yang telah disetarakan ke paket A dengan paket C yang telah disetarakan ke paket A. Kedua kurva yang berdekatan tersebut mengindikasikan hampir memiliki karakteristik tingkat kesukaran yang setara satu sama lain. Hal tersebut diperkuat dengan penelitian (Yusron et al., 2020, p. 3) yang menjelaskan bahwa semakin rapat gambar TCC setiap paket soal maka tingkat kesetaraan semakin baik. Meskipun demikian tidak berlaku pada paket D yang telah disetarakan ke paket A yang mana kurva tersebut berdekatan dengan kurva yang lain hanya pada rentang nilai *ability* ( $\theta$ ) dari -4 hingga -1,9. Sehingga untuk kurva ini tidak sepenuhnya mengindikasikan hampir memiliki karakteristik tingkat kesukaran yang setara dengan paket B dan paket C.

Hasil proses penyetaraan dalam penelitian ini kemudian menjadi bahan refleksi dan sesuai dengan penelitian (Yusron et al., 2020, p. 10) yang menjelaskan bahwa terdapat banyak hal yang harus diperhatikan

dalam pengembangan instrumen tes agar memperoleh paket-paket soal yang hampir memiliki kesetaraan karakteristik tes. Sedangkan penelitian Retnawati (2016b, p. 177) juga menjelaskan bahwa terdapat dua hal yang perlu diperhatikan dalam proses penyetaraan yakni (1) estimasi parameter, yang meliputi banyak responden, banyak butir, serta metode estimasi, dan (2) proses estimasi penyetaraan, yang meliputi distribusi parameter butir dan kemampuan, metode penyetaraan, banyak butir, serta pemilihan penggunaan *software*.

#### **4. Profil Peserta Didik dalam Menyelesaikan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA**

Soal AKM numerasi berbasis komputer yang dikembangkan mampu mengukur kemampuan peserta didik untuk menyelesaikan soal dengan baik. Hasil uji coba diperoleh indeks numerasi dan estimasi tingkat kompetensi yang telah dicapai peserta didik. Indeks numerasi yang diperoleh untuk instrumen AKM numerasi yang telah diuji cobakan sebesar 1,4685 yang menyatakan bahwa hasil “Di bawah kompetensi minimum”. Hasil tersebut memperlihatkan bahwa kemampuan bernumerasi peserta didik masing sangat rendah. Kurang ada pembiasaan penyelesaian soal-soal numerasi terlebih sejenis AKM baik atas inisiatif peserta didik sendiri atau dalam proses pembelajaran dengan dampingan pendidik. Pada pelaksanaan AKM numerasi yang telah berjalan, hanya peserta didik yang terpilih menjadi peserta ANBK yang mendapat fasilitas tambahan berupa bimbingan belajar guna sukses dalam pelaksanaan AKM numerasi. Hal tersebut menjadi salah satu pemicu bahwa numerasi peserta didik disetiap sekolah kurang terlatih dengan baik. Hasil penelitian ini sejalan dengan penelitian Syaifuddin (2022) dan Safitri & Retnawati (2020) yang menunjukkan hasil tes AKM numerasi peserta didik berada pada kategori rendah. Rendahnya numerasi peserta didik diharapkan meningkatkan rasa kepedulian pendidik dalam membiasakan peserta didik menyelesaikan soal sejenis AKM numerasi, sehingga kemampuan numerasi peserta didik dapat meningkat (Aufa & Manoy, 2022, p. 227).

Estimasi tingkat kompetensi peserta didik dalam menyelesaikan soal AKM numerasi dominan pada tingkat kompetensi “Dasar” sebesar 53,7%. Hal tersebut bermakna bahwa peserta didik telah memiliki penguasaan konsep dasar dalam menyelesaikan masalah. Konsep dasar yang dimaksud adalah konsep dasar pada domain aljabar, geometri dan pengukuran, serta data dan ketidakpastian. Namun hal yang menghambat peserta didik adalah masih memiliki kesulitan dalam menerapkan konsep dasar tersebut kedalam situasi yang relevan dalam hal ini konteks masalah yang tersajikan. Sehingga tindak lanjut yang perlu dilakukan pendidik adalah dengan memberikan contoh-contoh soal yang berkaitan dengan penerapan konsep dasar dalam penyelesaian soal, serta mengupayakan adanya diskusi bersama untuk menginterpretasi serta menarik kesimpulan dari penyelesaian masalah pada soal.

Disamping dominannya tingkat kompetensi “Dasar” yang dicapai oleh peserta didik, beberapa tingkat kompetensi lain juga telah dicapai peserta didik, yakni sebesar 30,9% peserta didik telah mencapai kompetensi “perlu intervensi khusus”. Pada tingkat kompetensi ini peserta didik memiliki pengetahuan matematika dan bernumerasi yang terbatas. Penguasaan konsep yang ditunjukkan sangat parsial dan keterampilan komputasi yang juga sangat terbatas. Tindak lanjut yang perlu dilakukan pendidik adalah melakukan pendampingan, mengadakan diskusi, serta rutin melakukan validasi terhadap serangkaian proses pencarian pengetahuan yang dilakukan oleh peserta didik. Selanjutnya sebesar 10,8% peserta didik telah mencapai kompetensi “cakap”. Pada tingkat kompetensi ini peserta didik telah mampu menerapkan pengetahuan matematika dalam bernumerasi, atau dengan kata lain peserta didik mampu menerapkan ke dalam konteks-konteks soal yang lebih beragam. Namun kekurangan yang dimiliki peserta didik pada tingkat kompetensi ini adalah kemampuan bernalarnya dalam menganalisis adanya kesalahan. Sehingga tindak lanjut yang perlu dilakukan adalah pendidik memberi penugasan yang sifatnya membandingkan antar penyelesaian masalah sehingga peserta didik

mampu mengevaluasi temuannya dan menarik kesimpulan. Disamping itu pendidik juga melakukan pendampingan untuk peserta didik melakukan justifikasi temuannya. Terakhir, sebesar 4,6% peserta didik telah mencapai kompetensi “mahir”. Pada tingkat kompetensi ini peserta didik telah mampu bernalar dalam menyelesaikan masalah kompleks serta non rutin berdasarkan penguasaan konsep matematika yang dimiliki. Meskipun tingkat kompetensi ini merupakan tingkat kompetensi yang tertinggi, peserta didik juga perlu diberikan tindak lanjut agar kemampuannya tetap terasah dan meningkat. Salah satu yang masih perlu dibiasakan kepada peserta didik dalam tingkat kompetensi ini adalah memberikan penugasan yang sifatnya membandingkan temuan-temuannya dengan temuan peserta didik lain atau kelompok lain atau sumber lain, hal ini dilakukan agar peserta didik mampu menganalisis guna membuat generalisasi hasil.

#### **5. Keefektifan Soal AKM Numerasi Berbasis Komputer untuk Kelas XI SMA**

Mengetahui keefektifan soal AKM numerasi yang telah dikembangkan merupakan tahap akhir dalam penelitian ini. Hal tersebut dilakukan dengan menganalisis hasil angket respon peserta didik dan pendidik setelah menggunakan produk yang telah peneliti kembangkan. Berdasarkan analisis diperoleh bahwa seluruh aspek yang diuji pada kepraktisan soal menunjukkan kategori cukup praktis hingga praktis digunakan. Sedangkan seluruh aspek yang diuji pada kemanfaatan soal menunjukkan kategori sangat manfaat untuk digunakan. Kedua hal tersebut menunjukkan bahwa produk yang telah dikembangkan efektif untuk dipergunakan lebih lanjut. Hasil keefektifan tersebut dapat menjadi solusi buat para pendidik untuk menjadikan produk ini sebagai bahan referensi dalam meningkatkan kemampuan numerasi. Hal ini diperkuat dengan penelitian Syaifuddin (2022, p. 322) yang menjelaskan bahwa dalam mendukung peningkatan kemampuan numerasi, dalam proses pembelajaran pendidik mendorong peserta didik untuk memecahkan

masalah melalui referensi soal-soal AKM numerasi yang tersedia dan bisa diakses secara umum.

### **C. Keterbatasan Penelitian**

Penelitian ini telah dilaksanakan dengan sebaik-baiknya, namun tak bisa dipungkiri masih terdapat keterbatasan dalam pelaksanaannya. Berikut adalah keterbatasan penelitian pada pengembangan soal AKM numerasi berbasis komputer untuk kelas XI SMA.

1. Waktu penelitian di dua sekolah berdekatan dengan pelaksanaan Ujian Kenaikan Kelas (UKK), menyebabkan waktu penelitian dipersingkat serta beberapa rombel kelas tidak diizinkan untuk dilaksanakan penelitian karena dipergunakan guru untuk *review* materi dan perbaikan nilai, sehingga mempengaruhi jumlah subjek penelitian;
2. Pelaksanaan penelitian terbatas sampai uji coba, tidak mengukur kemampuan numerasi lebih lanjut berdasarkan hasil perakitan setelah uji coba;

Meskipun dalam penelitian ini masih terdapat banyak keterbatasan, namun berdasarkan hasil analisis secara empiris menggunakan data respon yang dihasilkan, instrumen soal AKM numerasi ini tetap layak digunakan dan dimanfaatkan lebih lanjut.

## BAB V

### SIMPULAN DAN SARAN

#### A. Simpulan tentang Produk

Berdasarkan hasil penelitian pengembangan ini, kesimpulan yang dapat ditarik adalah sebagai berikut.

1. Konstruksi instrumen soal AKM numerasi yang dikembangkan terdiri atas domain aljabar, geometri dan pengukuran, serta data dan ketidakpastian. Konteks yang dilibatkan meliputi personal, sosial budaya, dan saintifik dengan level kognitif mencakup pemahaman (*knowing*), penerapan (*applying*), dan penalaran (*reasoning*). Kompetensi numerasi yang kemudian diturunkan menjadi indikator disesuaikan dengan kerangka kerja AKM numerasi. Soal yang dikembangkan sebanyak 54 soal, terbagi dalam 4 paket soal dengan 4 butir *anchor* dan 5 jenis bentuk soal yakni soal pilihan ganda, pilihan ganda kompleks, menjodohkan, isian singkat, dan uraian. Alokasi waktu yang ditetapkan dalam pelaksanaan penelitian adalah 90 menit.
2. Kualitas soal AKM numerasi berbasis komputer untuk kelas XI SMA yang dikembangkan telah memiliki kualitas soal tes yang baik. Hal tersebut ditunjukkan dengan instrumen yang telah memenuhi kriteria valid secara isi dan konstruk, reliabel digunakan dalam mengukur kemampuan numerasi peserta didik (baik yang memiliki tingkat kompetensi numerasi mahir, cakap, dasar, dan perlu intervensi khusus), serta memiliki tingkat kesukaran yang cukup baik.
3. Kesetaraan soal AKM numerasi yang dikembangkan secara bersama-sama untuk paket A, paket B, paket C, dan paket D tidak paralel. Setelah paket B, C, dan D disetarakan ke paket A dengan menggunakan metode rerata & rerata persamaan kesetaraan skor yang diperoleh yakni  $\theta_{ba}^* = \theta_b + 0,244$  untuk persamaan kesetaraan skor paket B ke paket A,  $\theta_{ca}^* = \theta_c + 0,141$  untuk persamaan kesetaraan skor paket C ke paket A, dan  $\theta_{da}^* = \theta_d + 0,575$  untuk persamaan kesetaraan skor paket D ke paket A. Disamping itu hasil kesetaraan dari kurva karakteristik tes menunjukkan bahwa paket B dan paket C yang

disetarakan ke paket A cenderung setara satu sama lain, namun tidak dengan paket D.

4. Profil peserta didik dalam menyelesaikan soal AKM numerasi berbasis komputer yang dikembangkan menunjukkan bahwa kemampuan numerasi peserta didik berada di bawah kompetensi minimum dengan tingkat kompetensi yang dominan adalah tingkat kompetensi “Dasar” sebesar 53,7%, tingkat “Perlu intervensi khusus” sebesar 30,9%, tingkat “Cakap” sebesar 10,8%, dan tingkat “Mahir” sebesar 4,6%.
5. Produk soal AKM numerasi berbasis komputer yang telah dikembangkan memiliki keefektifan dimana ditunjukkan dengan tingkat persentase kepraktisan soal sebesar 72,8% yang menunjukkan bahwa produk soal yang dikembangkan praktis digunakan atau memiliki kemudahan dalam penggunaannya, serta tingkat persentase kemanfaatan soal sebesar 91,1% yang menunjukkan produk soal yang dikembangkan sangat menfaat, terlebih jika dipergunakan lebih lanjut.

#### **B. Saran Pemanfaatan Produk**

Berdasarkan temuan dari hasil penelitian yang telah dilakukan, maka saran pemanfaatan produk soal AKM numerasi berbasis komputer untuk kelas XI SMA yang telah dikembangkan yakni sebagai berikut.

1. Instrumen soal AKM numerasi ini bermanfaat dalam menganalisis kemampuan numerasi peserta didik, sehingga hasil tes ini dapat dijadikan acuan untuk mengestimasi tingkat kompetensi numerasi peserta didik. Terlebih dapat dijadikan pembandingan dengan hasil AKM numerasi 2021 yang resmi dari Kemdikbud;
2. Instrumen soal AKM numerasi ini dikembangkan sesuai kerangka kerja AKM dari Kemdikbud, sehingga produk ini dapat dijadikan bahan referensi buat guru, sekolah atau praktisi pendidikan lainnya untuk mempersiapkan peserta didik yang terpilih menjadi peserta AKM di setiap tahunnya.

#### **C. Diseminasi dan Pengembangan Produk Lebih Lanjut**

Diseminasi hasil penelitian dapat dilakukan dengan berbagai cara. Hal yang dilakukan peneliti untuk penyebaran informasi hasil penelitian berupa penerbitan

artikel ilmiah pada jurnah ilmiah terindeks scopus dan atau sinta, serta mengikuti konferensi ilmiah yang selaras dengan ruang lingkup penelitian ini.

Pengembangan instrumen soal AKM numerasi ini memiliki implikasi untuk praktik dan penelitian. Sejak UN dihapus tahun 2021, asesmen secara nasional diubah menjadi Asesmen Nasional (AN) yang mana salah satu topik ujinya adalah AKM numerasi. Dengan pelaksanaan AKM yang masih terbilang baru tersebut, instrumen ini dapat menjadi gambaran untuk pengujian AKM kepada seluruh peserta didik. Mengingat pada pelaksanaan AN hanya beberapa peserta didik yang dipilih secara random untuk menjadi peserta AN. Selain itu, masih kurangnya ketersediaan bahan referensi AKM numerasi yang sesuai dengan kerangka kerja AKM oleh Kemdikbud. Penyajian instrumen soal AKM numerasi dengan metode CBT berbantuan LMS Moodle diharapkan dapat mengefisienkan pelaksanaan serta mendukung era digitalisasi ini.

Riset lebih lanjut dapat dilakukan pula pada penelitian eksperimen, penelitian eksploratif, penelitian korelasi, dan jenis penelitian lainnya menggunakan instrumen serupa, sehingga akan memberikan sumbangsih khusus, terlebih pada pendidikan matematika di Indonesia. Disamping itu, untuk riset lanjutan yang berkaitan dengan melakukan tes numerasi harus memahami bahwa tantangan kedepan butir tes yang dikembangkan harus bermanfaat serta prediktif dalam mengukur kemampuan peserta didik pada saat akan memasuki dunia kerja. Riset lanjutan yang berkaitan dengan pengembangan soal matematika disarankan untuk melibatkan juru gambar sehingga konten soal memberikan nuansa matematika yang sebenarnya sesuai dengan cerita. Terakhir, untuk riset lanjutan yang sifatnya berbasis media CBT diupayakan agar media CBT bukan sekedar dipergunakan untuk meletakkan soal, namun media CBT interaktif yang profesional.

## DAFTAR PUSTAKA

- Abidin, M., & Retnawati, H. (2019). A diagnosis of difficulties in answering questions of circle material on junior high school students. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 23(2), 144–155. <https://doi.org/10.21831/pep.v23i2.16454>
- Adelia, B. D., & Deta, U. A. (2022). Analisis perspektif peserta didik, guru dan calon guru fisika tentang asesmen kompetensi minimum (AKM). *Inovasi Pendidikan Fisika*, 11(1), 1–10. <https://doi.org/10.26740/ipf.v11n1.p1-10>
- Aida, N., Kusaeri, K., & Hamdani, S. (2017). Karakteristik instrumen penilaian hasil belajar matematika ranah kognitif yang dikembangkan mengacu pada model PISA. *Suska Journal of Mathematics Education*, 3(2), 130–139. <https://doi.org/10.24014/sjme.v3i2.3897>
- Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement*, 45(1), 131–142. <https://doi.org/10.1177/0013164485451012>
- Al-Hamad, N. Q. (2022). Moodle As a Learning Management System: Perceived Efficacy and Actual Use. *Journal of Educators Online*, 19(3), 1–19. <https://doi.org/10.9743/JEO.2022.19.3.2>
- Ali, U. S., Chang, H.-H., & Anderson, C. J. (2015). Location Indices for Ordinal Polytomous Items Based on Item Response Theory. *ETS Research Report Series*, 2015(2), 1–13. <https://doi.org/10.1002/ets2.12065>
- Alkhateeb, M. A., & Abdalla, R. A. (2021). Factors influencing student satisfaction towards using learning management system moodle. *International Journal of Information and Communication Technology Education*, 17(1), 138–153. <https://doi.org/10.4018/IJICTE.2021010109>
- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*.

Brooks/Cole Publishing Company.

- Amnie, E., Rosidin, U., Herlina, K., & Abdurrahman, A. (2021). Developing assessment in improving students' digital literacy skills. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 25(1), 1–15. <https://doi.org/10.21831/pep.v25i1.33600>
- Anam, F., Suteja, J. R., Septianto, A., Purnomo, A., & Utami, Y. P. (2019). Improving the numeracy mathematics ability: The role of abacus learning model. *Journal of Physics: Conference Series*, 1594(1), 1–5. <https://doi.org/10.1088/1742-6596/1594/1/012041>
- Anderha, R. R., & Maskar, S. (2021). Pengaruh kemampuan numerasi dalam menyelesaikan masalah matematika terhadap prestasi belajar mahasiswa pendidikan matematika. *Jurnal Ilmiah Matematika Realistik*, 2(1), 1–10. <http://jim.teknokrat.ac.id/index.php/pendidikanmatematika/article/view/774>
- Andiani, D., Hajizah, M. N., & Dahlan, J. A. (2020). Analisis rancangan assesmen kompetensi minimum (AKM) numerasi program merdeka belajar. *Majamath: Jurnal Matematika Dan Pendidikan Matematika*, 4(1), 80–90. <http://ejurnal.unim.ac.id/index.php/majamath/article/view/1010/544>
- Andrian, Y., & Rusman, R. (2019). Implementasi pembelajaran abad 21 dalam kurikulum 2013. *Jurnal Penelitian Ilmu Pendidikan*, 12(1), 14–23. <https://doi.org/10.21831/jpipfip.v12i1.20116>
- Anggraena, Y., Ginanto, D., Felicia, N., Andiarti, A., Herutami, I., Alhapip, L., Iswoyo, S., Hartini, Y., & Mahardika, R. L. (2022). *Panduan Pembelajaran dan Asesmen Pendidikan Anak Usia Dini, Pendidikan Dasar, dan Menengah*. Badan Standar, Kurikulum, dan Asesmen Pendidikan Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi Republik Indonesia.
- Apino, E., & Retnawati, H. (2017). *Perangkat pembelajaran matematika SMA semester genap menggunakan model pembelajaran creative problem solving berorientasi pada higher order thinking skills siswa*. Parama Publishing.

- Apriandi, D., Retnawati, H., & Abadi, A. M. (2022). Construct Validity and Reliability of the Learning Motivation Questionnaire. *TEM Journal*, *11*(4), 1494–1499. <https://doi.org/10.18421/TEM114-09>
- Arifin, W. N., Yusoff, M. S. B., & Naing, N. N. (2012). Confirmatory factor analysis (CFA) of USM Emotional Quotient Inventory (USMEQ-i) among medical degree program applicants in Universiti Sains Malaysia (USM). *Education in Medicine Journal*, *4*(2), 26–44. <https://doi.org/10.5959/eimj.v4i2.33>
- Atmazaki, Ali, N. B. V., Muldian, W., Miftahussururi, Hanifah, N., Nento, M. N., & Akbari, Q. S. (2017). *Panduan Gerakan Literasi Nasional*. Kementerian Pendidikan dan Kebudayaan.
- Aufa, N. I., & Manoy, J. T. (2022). Student's mathematical literacy in solving asesmen kompetensi minimum in terms of gender. *MATHEdunesa: Jurnal Ilmiah Pendidikan Matematika*, *11*(1), 219–229. <https://doi.org/10.26740/mathedunesa.v11n1.p219-229>
- Balt, M., Ehlert, A., & Fritz, A. (2019). Assessment in inclusive mathematics education: approaches to designing progress assessments for numeracy learning. *Inclusive Mathematics Education*, *x*(x), 197–216. <https://doi.org/10.1007/978-3-030-11518-0>
- Barham, A. I., Ihmeideh, F., Al-Falasi, M., & Alabdallah, A. (2019). Assessment of first-grade students' literacy and numeracy levels and the influence of key factors. *International Journal of Learning, Teaching and Educational Research*, *18*(12), 174–195. <https://doi.org/10.26803/ijlter.18.12.11>
- Bellini, D., Crescentini, A., Zanolla, G., Cubico, S., Favretto, G., Faccincani, L., Ardolino, P., & Giancesini, G. (2019). Mathematical Competence Scale (MCS) for primary school: The psychometric properties and the validation of an instrument to enhance the sustainability of talents development through the

- numeracy skills assessment. *Sustainability (Switzerland)*, *11*(9), 1–13. <https://doi.org/10.3390/su11092569>
- Butler, A. C. (2018). Multiple-choice testing in education: are the best practices for assessment also good for learning? *Journal of Applied Research in Memory and Cognition*, *7*(3), 323–331. <https://doi.org/10.1016/j.jarmac.2018.07.002>
- Cahyanovianty, A. D., & Wahidin. (2021). Analisis kemampuan numerasi peserta didik kelas VIII dalam menyelesaikan soal asesmen kompetensi minimum. *Jurnal Cendekia: Jurnal Pendidikan Matematika*, *05*(02), 1439–1448. <https://doi.org/10.31004/cendekia.v5i2.651>
- Cardona-Zea, V. P., & Rojas-Galeano, S. A. (2012). Recognising irregular answers in automatic assessment of fill-in-the-blank tests. *IEEE Workshop on Engineering Applications*, *x*(x), 1–4. <https://doi.org/10.1109/WEA.2012.6220096>
- Cavus, N. (2015). Distance Learning and Learning Management Systems. *Procedia - Social and Behavioral Sciences*, *191*, 872–877. <https://doi.org/10.1016/j.sbspro.2015.04.611>
- Charmila, N., Zulkardi, & Darmawijoyo. (2016). Pengembangan soal matematika model PISA menggunakan konteks Jambi. *Jurnal Penelitian Dan Evaluasi Pendidikan*, *20*(2), 198–207. <https://doi.org/10.21831/pep.v20i2.7444>
- Degroote, L., Desmet, A., De Bourdeaudhuij, I., Van Dyck, D., & Crombez, G. (2020). Content validity and methodological considerations in ecological momentary assessment studies on physical activity and sedentary behaviour: A systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, *17*(1), 1–13. <https://doi.org/10.1186/s12966-020-00932-9>
- Djidu, H., Ismail, R., RachmaningtyA, N. A., Sumin, Imawan, O. R., Suhariyono, Aviory, K., Prihono, E. W., Kurniawan, D. D., Syahbrudin, J., Nurdin, Marinding, Y., Firmansyah, Retnawati, H., & Hadi, S. (2022). *Analisis*

*Instrumen Penelitian dengan Teori Tes Klasik dan Modern Menggunakan Program R.* UNY Press.

DOĞRUÖZ, E., & AKIN ARIKAN, Ç. (2020). Comparison of different ability estimation methods based on 3 and 4PL item response theory. *Pamukkale University Journal of Education*, 50, 50–69. <https://doi.org/10.9779/pauefd.585774>

Draaijer, S., Jordan, S., & Ogden, H. (2018). Calculating the random guess score of multiple-response and matching test items. *Communications in Computer and Information Science*, 829(x), 210–222. [https://doi.org/10.1007/978-3-319-97807-9\\_16](https://doi.org/10.1007/978-3-319-97807-9_16)

Eduka, T. M. (2020). *Strategi Sukses UTBK SBMPTN SAINTEK 2021*. Genta Smart. [https://books.google.co.id/books?hl=id&lr=&id=fMTyDwAAQBAJ&oi=fnd&pg=PP2&dq=matematika+saintek+sbmptn&ots=4xnW4LbQAF&sig=-Zefd3Qgvb7F3DGzRkCobhT\\_J2k&redir\\_esc=y#v=onepage&q=matematika+saintek+sbmptn&f=false](https://books.google.co.id/books?hl=id&lr=&id=fMTyDwAAQBAJ&oi=fnd&pg=PP2&dq=matematika+saintek+sbmptn&ots=4xnW4LbQAF&sig=-Zefd3Qgvb7F3DGzRkCobhT_J2k&redir_esc=y#v=onepage&q=matematika+saintek+sbmptn&f=false)

Efendi, R., Lesmana, L. S., Putra, F., Yandani, E., & Wulandari, R. A. (2021). Design and implementation of computer based test (CBT) in vocational education. *Journal of Physics: Conference Series*, 1764(1), 1–12. <https://doi.org/10.1088/1742-6596/1764/1/012068>

Faizah, U., Zuchdi, D., & Alsamiri, Y. (2019). An authentic assessment model to assess kindergarten students' character. *Research and Evaluation in Education*, 5(2), 103–119. <https://doi.org/10.21831/reid.v5i2.24588>

Fakhrunisa, F., & Prabawanto, S. (2020). Online learning in COVID-19 pandemic: an investigation of mathematics teachers' perception. *ICEEL 2020: 2020 The 4th International Conference on Education and E-Learning*, 207–213. <https://doi.org/10.1145/3439147.3439179>

- Fong, W. L., & Kaur, B. (2015). A study of mathematics written assessment in singapore secondary schools. *The Mathematics Educator*, 16(1), 19–44. [http://math.nie.edu.sg/ame/matheduc/tme/tmeV16\\_1/TME16\\_2.pdf](http://math.nie.edu.sg/ame/matheduc/tme/tmeV16_1/TME16_2.pdf)
- Gal, I., & Tout, D. (2014). *Comparison of PIAAC and PISA frameworks for numeracy and mathematical literacy*. 102, 57. <https://doi.org/10.1787/5jz3wl63cs6f-en>
- Gall, M. D., Gall, J. P., & Borg, W. R. (2003). *Educational research: An introduction* (7th ed.). Allyn & Bacon.
- Geiger, V., Goos, M., & Forgasz, H. (2015). A rich interpretation of numeracy for the 21st century: a survey of the state of the field. *ZDM - International Journal on Mathematics Education*, 47(4), 531–548. <https://doi.org/10.1007/s11858-015-0708-1>
- Getenet, S. T. (2022). Teachers' knowledge framework for designing numeracy rich tasks across non-mathematics curriculum areas. *International Journal of Education in Mathematics, Science and Technology*, 10(3), 663–680. <https://doi.org/10.46328/ijemst.2137>
- Gierl, M. J., Bulut, O., Guo, Q., & Zhang, X. (2017). Developing, Analyzing, and Using Distractors for Multiple-Choice Tests in Education: A Comprehensive Review. *Review of Educational Research*, 87(6), 1082–1116. <https://doi.org/10.3102/0034654317726529>
- Goshima, R., Shimakawa, H., Harada, F., & Phuong, D. D. (2019). Extraction of poor learning items with automatic labeling in fill-in-the-blank test. *14th International Conference on Computer Science and Education, ICCSE 2019, Iccse*, 907–912. <https://doi.org/10.1109/ICCSE.2019.8845523>
- Griffin, P., McGaw, B., & Care, E. (2012). *Assessment and teaching of 21st century skills*. Springer.

- Guilford, J. P. (1956). *Fundamental statistics in psychology and education*. McGraw-Hill.
- Haberkorn, K., Pohl, S., Carstensen, C., & Wiegand, E. (2016). Scoring of complex multiple choice items in NEPS competence tests. *Methodological Issues of Longitudinal Surveys*, 523–540. <https://doi.org/10.1007/978-3-658-11994-2>
- Hair, J. F., Black, W. C., Babin, B. J., & Anseron, R. E. (2013). *Multivariate Data Analysis* (7th ed.). Pearson Education.
- Haladyna, T. M., Rodriguez, M. C., & Stevens, C. (2019). Are Multiple-choice Items Too Fat? *Applied Measurement in Education*, 32(4), 350–364. <https://doi.org/10.1080/08957347.2019.1660348>
- Hall, J., & Zmood, S. (2019). Australia’s literacy and numeracy test for initial teacher education students: Trends in numeracy for low- and high-achieving students. *Australian Journal of Teacher Education*, 44(10), 1–17. <https://doi.org/10.14221/ajte.2019v44n10.1>
- Hambleton, R. K., & Swaminathan, H. (1985). *Item response theory*. MA: Kluwer Inc.
- Hambleton, R. K., Swaminathan, H., & Rogers, J. H. (1991). *Fundamental of Item Response Theory*. SAGE Publications.
- Haręzlak, K., & Werner, A. (2012). Database access and management with the use of the MOODLE platform. *Advances in Intelligent and Soft Computing*, 98, 49–65. [https://doi.org/10.1007/978-3-642-23187-2\\_4](https://doi.org/10.1007/978-3-642-23187-2_4)
- Hayati, N., & Mardapi, D. (2014). Pengembangan Butir Soal Matematika SD Di Kabupaten Lombok Timur Sebagai Upaya dalam Pengadaan Bank Soal. *Jurnal Kependidikan*, 44(1), 26–38. <https://doi.org/10.21831/jk.v44i1.2189>
- Hendrowati, T. Y., & Faelasofi, R. (2021). Numeracy and literacy skill in elementary school students: the utilization of kampus mengajar perintis

program. *Desimal: Jurnal Matematika*, 4(3), 365–372.  
<https://doi.org/10.24042/djm>

Herman, Shara, A. M., Silalahi, T. F., Sherly, & Julyanthry. (2022). Teachers' attitude towards minimum competency assessment at sultan agung senior high school in pematangsiantar, indonesia. *Journal of Curriculum and Teaching*, 11(1), 1–14. <https://doi.org/10.5430/jct.v11n2p1>

Heru, Yuliana, R. E., & Zulpah, I. (2021). Design of supplementary mathematics module for preparation of minimum competency assessment for fifth grade elementary school students. *Jurnal Math Educator Nusantara*, 8(1), 75–90. <https://doi.org/10.29407/jmen.v8i1.17682>

Hewi, L., & Shaleh, M. (2020). Refleksi hasil PISA (The Programme For International Student Assesment): upaya perbaikan bertumpu pada pendidikan anak usia dini). *Jurnal Golden Age*, 4(01), 30–41. <https://doi.org/10.29408/jga.v4i01.2018>

Hikamudin, E. (2017). Estimasi Kemampuan Siswa Dalam Ujian Nasional Menggunakan Metode Bayes. *Jurnal Penelitian Kebijakan Pendidikan*, 10(2), 1–14. <https://doi.org/10.24832/jpkp.v10i2.171>

Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1979). *Applied Statistics for The Behavioral Sciences*. Houghton Mifflin Company.

Kemdikbud. (2017). *Materi pendukung literasi numerasi*. <https://gln.kemdikbud.go.id/glnsite/wp-content/uploads/2017/10/literasi-numerasi.pdf>

Kemdikbud. (2019). *Pendidikan di Indonesia belajar dari hasil PISA 2018*. Pusat Penilaian Pendidikan, Badan Penelitian dan Pengembangan, Kementerian Pendidikan dan Kebudayaan.

Kemendikbudristek. (2022). *Buku Panduan Capaian Hasil Asesmen Nasional*.

Pusat Asesmen Pendidikan, BSKAP. merdekabelajar.kemdikbud.go.id

- Khoshsima, H., Toroujeni, S. M. H., Thompson, N., & Ebrahimi, M. R. (2019). Computer-based (CBT) vs. paper-based (PBT) testing: Mode effect, relationship between computer familiarity, attitudes, aversion and mode preference with CBT test scores in an asian private EFL context. *Teaching English with Technology*, 19(1), 86–101. [https://1549007223wpdm\\_article6b.pdf](https://1549007223wpdm_article6b.pdf)
- Knowles, A., Linsell, C., Baeumer, B., & Anakin, M. (2021). The development and efficacy of an undergraduate numeracy assessment tool. *The Annual Meeting of the Mathematics Education Research Group of Australasia (MERGA) (43rd, Singapore)*, 243–250. <https://eric.ed.gov/?id=ED616235>
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3rd ed., pp. 487–536). Springer. <https://doi.org/10.1007/978-1-4939-0317-7>
- Komariah, P. D., Ediyanto, A., Harsasi, M., Safari, Nurcholis, H., & Budhi, W. S. (2020). *Fokus Asesmen Kompetensi Minimum SMK/MAK*. Erlangga.
- Kurpius, S. E. R., & Stafford, M. E. (2006). *Testing and Measurement: A Used-Friendly Guide*. SAGE Publications, Ltd.
- Kusainun, N. (2020). Analisis standar penilaian pendidikan di Indonesia. *Jurnal Pendidikan*, 5(1), 1–7. <https://doi.org/10.26740/jp.v5n1.p%25p>
- Kustantina, V. A., Nuryadi, & Marhaeni, N. H. (2022). Improving the students' numerical literacy skills by using interactive mathematical comics on pythagorean theorem. *JIPMuktj: Jurnal Ilmu Pendidikan Muhammadiyah Kramat Jati*, 3(1), 10–16. <https://jurnal.pcmkramatjati.or.id/index.php/JIPMUKJT/index%0Ahttps://jurnal.pcmkramatjati.or.id/index.php/JIPMUKJT/index%0A>

- Lee, W. S. (2013). Propensity score matching and variations on the balancing test. *Empirical Economics*, 44(1), 47–80. <https://doi.org/10.1007/s00181-011-0481-0>
- Lia, R. M., Rusilowati, A., & Isnaeni, W. (2020). NGSS-oriented chemistry test instruments: Validity and reliability analysis with the Rasch model. *Research and Evaluation in Education*, 6(1), 41–50. <https://doi.org/10.21831/reid.v6i1.30112>
- Looney, A., Cumming, J., van Der Kleij, F., & Harris, K. (2018). Reconceptualising the role of teachers as assessors: teacher assessment identity. *Assessment in Education: Principles, Policy and Practice*, 25(5), 442–467. <https://doi.org/10.1080/0969594X.2016.1268090>
- Loyd, B. H., & Hoover, H. D. (1980). Vertical equating using the rasch model. *Journal of Educational Measurement*, 17(3), 179–193. <https://doi.org/10.1111/j.1745-3984.1980.tb00825.x>
- Luna, J. M., Castro, C., & Romero, C. (2017). MDM tool: A data mining framework integrated into Moodle. *Computer Applications in Engineering Education*, 25(1), 90–102. <https://doi.org/10.1002/cae.21782>
- Machromah, I. U., Utami, N. S., Setyaningsih, R., Mardhiyana, D., & Fatmawati, L. W. S. (2021). Minimum competency assessment: Designing tasks to support students' numeracy. *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, 12(14), 3268–3277. <https://turcomat.org/index.php/turkbilmater/article/view/10898/8186>
- Manurung, A. B., & Rajagukguk, J. (2019). Desain evaluasi hasil belajar fisika berbasis computer based test (CBT) pada materi pokok usaha dan energi. *INPAFI (Inovasi Pembelajaran Fisika)*, 6(4), 7–13. <https://doi.org/10.24114/inpafi.v6i4.12479>
- Mardapi, D. (2008). *Teknik penyusunan instrumen tes dan non tes*. Mitra Cendika.

- Mardapi, D. (2016). *Pengukuran, Penilaian dan Evaluasi Pendidikan*. Parama Publishing.
- Megawati, L. A., & Sutarto, H. (2021). Analysis numeracy literacy skills in terms of standardized math problem on a minimum competency assessment. *Unnes Journal of Mathematics Education*, 10(2), 155–165. <https://doi.org/10.15294/ujme.v10i2.49540>
- Meriana, T., Murniarti, E., & Dasar Kanaan, S. (2021). Analisis pelatihan asesmen kompetensi minimum. *Jurnal Dinamika Pendidikan*, 14(2), 110–116. <https://doi.org/10.51212/jdp.v14i2.7>
- Naga, D. (1992). *Pengantar Teori Sekor pada Pengukuran Pendidikan*. Gunadarma.
- Nortvedt, G. A., & Wiese, E. (2020). Numeracy and migrant students: a case study of secondary level mathematics education in Norway. *ZDM - Mathematics Education*, 52(3), 527–539. <https://doi.org/10.1007/s11858-020-01143-z>
- Nuzulia, N., & Gafur, A. (2022). Pengembangan buku latihan berbasis asesmen kompetensi minimum (AKM) untuk meningkatkan kemampuan literasi dan numerasi siswa di SDN Janti 02 Sidoarjo. *Madrosatuna: Journal of Islamic Elementary School*, 6(1), 1–10. <https://doi.org/10.21070/madrosatuna.v6i1.1564>
- OECD. (2000). *Measuring student knowledge and skills: the PISA 2000 assessment of reading, mathematical and scientific literacy*. [Www.oecd.org/education/school/programme-for-international-student-assessment-pisa/33692793.pdf](http://www.oecd.org/education/school/programme-for-international-student-assessment-pisa/33692793.pdf)
- OECD. (2009). *Assessment framework: key competencies in reading, mathematics and science*. OECD.
- OECD. (2013). *PISA 2012 assessment and analytical framework: mathematics*,

- reading, science, problem solving and financial literacy*. OECD Publishing.  
<https://doi.org/10.1787/9789264190511-en>
- OECD. (2018). *PISA Mathematics Scale: Overall Mathematics*.  
<https://pisadataexplorer.oecd.org/ide/idepisa/report.aspx>
- OECD. (2019). *PISA 2018 Assessment and Analytical Framework*. OECD Publishing. <https://doi.org/10.1787/b25efab8-en>
- Ozkale, A., & Ozdemir Erdogan, E. (2020). An analysis of the interaction between mathematical literacy and financial literacy in PISA\*. *International Journal of Mathematical Education in Science and Technology*, 53(8), 1–21.  
<https://doi.org/10.1080/0020739X.2020.1842526>
- Papenberg, M., & Musch, J. (2017). Of Small Beauties and Large Beasts: The Quality of Distractors on Multiple-Choice Tests Is More Important Than Their Quantity. *Applied Measurement in Education*, 30(4), 273–286.  
<https://doi.org/10.1080/08957347.2017.1353987>
- PIAAC Numeracy Expert Group. (2009). PIAAC Numeracy: A Conceptual Framework. In *OECD Education Working Papers* (Issue 35). OECD Publishing. <https://doi.org/10.1787/220337421165>
- Pinastiti, L. A., Sihombing, R. M., & Irfansyah, I. (2022). Design activities as a literation for introduction to horticulture agriculture for children aged 5 - 12 years old. *The 2nd International Conference on Art, Craft, Culture and Design (ICON-ARCCADE 2021)*, 625, 201–206.  
<https://doi.org/10.2991/assehr.k.211228.026>
- Pramadya, W., Riyadi, & Indriati, D. (2019). Self-assessment profile on statistics using computer-based mathematical summative test. *Journal of Physics: Conference Series*, 1188(1), 1–6. <https://doi.org/10.1088/1742-6596/1188/1/012053>

- Purnomo, H., Sa'dijah, C., Hidayanto, E., Sisworo, S., Permadi, H., & Anwar, L. (2022). Development of instrument numeracy skills test of minimum competency assessment (MCA) in Indonesia. *International Journal of Instruction, 15*(3), 635–648. <https://doi.org/10.29333/iji.2022.15335a>
- Pusmenjar. (2020). Desain Pengembangan Soal Asesmen Kompetensi Minimum. In *Desain Pengembangan AKM*. Badan Penelitian dan Pengembangan dan Perbukuan, Kementerian Pendidikan dan Kebudayaan.
- Pusmenjar. (2021a). *Framework asesmen kompetensi minimum (AKM)*. Badan Penelitian dan Pengembangan dan Perbukuan, Kementerian Pendidikan dan Kebudayaan.
- Pusmenjar. (2021b). *Pembelajaran paradigma baru*. Badan Penelitian dan Pengembangan dan Perbukuan, Kementerian Pendidikan, Kebudayaan, Riset, dan Teknologi.
- Queensland, C. of T. (2015). *Numeracy teaching across the curriculum in Queensland: Resources for teachers*. Queensland College of Teachers.
- Rahmawati, W. A., Usodo, B., & Fitriana, D. L. (2021). Mathematical literacy skills students of the junior high school in solving PISA-like mathematical problems. *IOP Conference Series: Earth and Environmental Science, 1808*(1), 1–10. <https://doi.org/10.1088/1742-6596/1808/1/012045>
- Retnawati, H. (2014). *Teori respons butir dan penerapannya*. Nuha Medika.
- Retnawati, H. (2016a). *Analisis kuantitatif instrumen penelitian*. Parama Publishing.
- Retnawati, H. (2016b). Perbandingan metode penyetaraan skor tes menggunakan butir bersama dan tanpa butir bersama. *Jurnal Kependidikan, 46*(2), 164–178. <https://doi.org/10.21831/jk.v46i2.10383>
- Retnawati, H., Kartowagiran, B., Arlinwibowo, J., & Sulistyarningsih, E. (2017).

Why are the mathematics national examination items difficult and what is teachers' strategy to overcome it? *International Journal of Instruction*, 10(3), 257–276. <https://doi.org/10.12973/iji.2017.10317a>

Riduwan. (2010). *Skala Pengukuran Variabel-Variabel Penelitian*. Alfabeta.

Rohim, D. C. (2021). Konsep asesmen kompetensi minimum untuk meningkatkan kemampuan literasi numerasi siswa sekolah dasar. *Jurnal VARIDIKA*, 33(1), 54–62. <https://doi.org/10.23917/varidika.v33i1.14993>

Romm, A. T., Schoer, V., & Kika, J. C. (2019). A test taker's gamble: The effect of average grade to date on guessing behaviour in a multiple choice test with a negative marking rule. *South African Journal of Economic and Management Sciences*, 22(1), 1–13. <https://doi.org/10.4102/sajems.v22i1.2542>

Sabekti, A. W., & Khoirunnisa, F. (2018). Penggunaan Rasch model untuk mengembangkan instrumen pengukuran kemampuan berpikir kritis siswa pada topik ikatan kimia. *Jurnal Zarah*, 6(2), 68–75. <https://doi.org/10.31629/zarah.v6i2.724>

Safitri, A., & Retnawati, H. (2020). The estimation of mathematics literacy ability of junior high school students with partial credit model (pcm) scoring on quantity. *Journal of Physics: Conference Series*, 1581(1), 1–8. <https://doi.org/10.1088/1742-6596/1581/1/012030>

Şahin, A., & Anıl, D. (2017). The effects of test length and sample size on item parameters in item response theory. *Kuram ve Uygulamada Egitim Bilimleri*, 17(1), 321–335. <https://doi.org/10.12738/estp.2017.1.0270>

Sajjad, M., Iltaf, S., & Khan, R. A. (2020). Nonfunctional distractor analysis: An indicator for quality of multiple choice questions. *Pakistan Journal of Medical Sciences*, 36(5), 982–986. <https://doi.org/10.12669/pjms.36.5.2439>

Santoso, A., Kartianom, K., & Kassymova, G. K. (2019). Kualitas butir bank soal

- statistika (Studi kasus: Instrumen ujian akhir mata kuliah statistika Universitas Terbuka). *Jurnal Riset Pendidikan Matematika*, 6(2), 165–176. <https://doi.org/10.21831/jrpm.v6i2.28900>
- Santoso, A., Pardede, T., Apino, E., Djidu, H., Rafi, I., Rosyada, M. N., Retnawati, H., & Kassymova, G. K. (2022). Polytomous scoring correction and its effect on the model fit: A case of item response theory analysis utilizing R. *Psychology, Evaluation, and Technology in Educational Research*, 5(1), 1–13. <https://doi.org/10.33292/petier.v5i1.148>
- Sass, D. A., Schmitt, T. A., & Walker, C. M. (2008). Estimating non-normal latent trait distributions within item response theory using true and estimated item parameters. *Applied Measurement in Education*, 21(1), 65–88. <https://doi.org/10.1080/08957340701796415>
- Simanullang, N. H. S., & Rajagukguk, J. (2020). Learning Management System (LMS) Based on Moodle to Improve Students Learning Activity. In *Journal of Physics: Conference Series* (Vol. 1462, Issue 1, pp. 1–7). <https://doi.org/10.1088/1742-6596/1462/1/012067>
- Sulistiyani, N., & Kusumawardana, A. S. (2022). Pendampingan penyusunan modul numerasi berorientasi asesmen kompetensi minimum di sekolah dasar. *Jurnal Masyarakat Mandiri*, 6(1), 464–474. <https://doi.org/10.31764/jmm.v6i1.6431>
- Supahar, & Prasetyo, Z. K. (2015). Developing Assessment Instrument of The Performance of The Inquiry Abilities in Physics Subject of High School Students. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 19(1), 104–107. <https://doi.org/10.21831/pep.v19i1.4560>
- Supriyati, Y., Iriyadi, D., & Falani, I. (2021). The development of equating application for computer based test in physics hots category. *Journal of Technology and Science Education*, 11(1), 117–128. <https://doi.org/10.3926/jotse.1135>

- Susac, A., Planinic, M., Klemencic, D., & Milin Sipus, Z. (2018). Using the Rasch model to analyze the test of understanding of vectors. *Physical Review Physics Education Research*, *14*(2), 1–6. <https://doi.org/10.1103/PhysRevPhysEducRes.14.023101>
- Syaifuddin, M. (2022). Minimum competency assessment to measure mathematical literacy of junior high school students. *Journal of Education Research and Evaluation*, *6*(2), 316–326. <https://doi.org/10.23887/jere.v6i2.46263>
- Tangianu, F., Mazzone, A., Berti, F., Pinna, G., Bortolotti, I., Colombo, F., Nozzoli, C., Regina, M. La, Greco, A., Filannino, C., Silingardi, M., & Nardi, R. (2018). Are multiple-choice questions a good tool for the assessment of clinical competence in Internal Medicine? *Italian Journal of Medicine*, *12*, 88–96. <https://doi.org/10.4081/itjm.2018.980>
- Tesio, L., Caronni, A., Simone, A., Kumbhare, D., & Scarano, S. (2023). Interpreting results from Rasch analysis 2. Advanced model applications and the data-model fit assessment. *Disability and Rehabilitation*, *x*(x), 1–14. <https://doi.org/10.1080/09638288.2023.2169772>
- Thai, T., Hartup, K., Colbourn, A., & Yeung, A. (2021). Using an online numeracy practice test to support education students for the numeracy component of the LANTITE. *Australian Journal of Teacher Education*, *46*(9), 73–90. <https://doi.org/10.14221/ajte.2021v46n9.5>
- Tosuncuoglu, I. (2018). Importance of assessment in ELT. *Journal of Education and Training Studies*, *6*(9), 163–167. <https://doi.org/10.11114/jets.v6i9.3443>
- Uto, M., & Okano, M. (2021). Learning automated essay scoring models using item-response-theory-based scores to decrease effects of rater biases. *IEEE Transactions on Learning Technologies*, *14*(6), 763–776. <https://doi.org/10.1109/TLT.2022.3145352>
- Uysal, İ., Şahin-Kürşad, M., & Kılıç, A. F. (2022). Effect of item parameter drift in

- mixed format common items on test equating. *Participatory Educational Research*, 9(5), 143–160. <https://doi.org/10.17275/per.22.108.9.5>
- Westwood, P. (2008). *What Teachers Need to Know About Numeracy*. ACER Press.
- Widarti, H. R., Rokhim, D. A., Septiani, M. O., & Dzikrulloh, M. H. A. (2022). Identification of science teacher practices and barriers in preparation of minimum competency assessment in the Covid-19 pandemic era. *Orbital*, 14(1), 63–67. <https://doi.org/10.17807/orbital.v14i1.1695>
- Widiana, I. W., Jampel, I. N., & Prawini, I. P. (2018). The effectiveness of traditional game-based communication learning activity for cognitive process dimension learning achievement. *Cakrawala Pendidikan*, 37(2), 260–269. <https://doi.org/10.21608/pshj.2022.250026>
- Widianti, W., & Hidayati, N. (2021). Analisis kemampuan literasi matematis siswa SMP pada materi segitiga dan segiempat. *Jurnal Pembelajaran Matematika Inovatif*, 4(1), 27–38. <https://doi.org/10.22460/jpmi.v4i1.27-38>
- Widyaningsih, S. W., Yusuf, I., Prasetyo, Z. K., & Istiyono, E. (2021). The development of the hots test of physics based on modern test theory: Question modeling through e-learning of moodle lms. *International Journal of Instruction*, 14(4), 51–68. <https://doi.org/10.29333/iji.2021.1444a>
- Wijono, S., & Mardapi, D. (2016). Model evaluasi ujian nasional kompetensi keahlian teknik pemesinan SMK. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 20(2), 234–243. <https://doi.org/10.21831/pep.v20i2.8731>
- Winarni, E. W., Purwandari, E. P., & Hafiza, S. (2022). Automatic Essay Assessment for Blended Learning in Elementary School. *International Journal on Advanced Science, Engineering and Information Technology*, 12(1), 85–91. <https://doi.org/10.18517/ijaseit.12.1.11835>
- Wiradika, I. N. I., & Retnawati, H. (2022). Item characteristics analysis using the

rasch model in the development of reading literacy instruments for elementary schools students. *Progres Pendidikan*, 3(3), 142–150.  
<https://doi.org/10.29303/prospek.v3i3.263>

Xin, O. K., & Singh, D. (2021). Development of Learning Analytics Dashboard based on Moodle Learning Management System. *International Journal of Advanced Computer Science and Applications*, 12(7), 838–843.  
<https://doi.org/10.14569/IJACSA.2021.0120793>

Yusron, E., Retnawati, H., & Rafi, I. (2020). Bagaimana hasil penyetaraan paket tes USBN pada mata pelajaran matematika dengan teori respon butir? *Jurnal Riset Pendidikan Matematika*, 7(1), 1–12.  
<https://doi.org/10.21831/jrpm.v7i1.31221>

Zainudin, M., Subali, B., & Jailani. (2019). Construct validity of mathematical creativity instrument: First-order and second-order confirmatory factor analysis. *International Journal of Instruction*, 12(3), 595–614.  
<https://doi.org/10.29333/iji.2019.12336a>

Zi Yan, T. B., & Heene, M. (2021). *Applying The Rasch Model Fundamental Measurement in The Human Sciences* (4 th). Routledge.