PENGEMBANGAN INSTRUMEN PENILAIAN LITERASI MATEMATIKA MAHASISWA CALON GURU MATEMATIKA

Lusi Eka Afri[1], Nurrahmawati[2]
[1,2]Universitas Pasir Pengaraian, Indonesia
lusiekaafri13@gmail.com

Abstract Twenty-first-century learning requires teachers, particularly prospective mathematics teachers, to possess strong mathematical literacy in order to guide students in understanding, applying, and interpreting mathematics in various real-life contexts. This study aimed to develop a valid and reliable mathematical literacy assessment instrument for prospective mathematics teachers by adopting the 4-D development model. The instrument focused on counting techniques and probability concepts, aligned with the AKM/IEA mathematical literacy cognitive levels of *knowing*, *applying*, and *reasoning*. The trial participants consisted of 64 undergraduate students enrolled in the mathematics education program at the second, third, and fourth-year levels. Content validity analysis using Aiken's index indicated that all items were valid, with coefficients ranging from 0.80 to 0.95. Furthermore, construct validity testing through *Confirmatory Factor Analysis* (CFA) demonstrated adequate convergent validity, as indicated by an *Average Variance Extracted* (AVE) value of 0.581 and a composite reliability (ϱc) value of 0.815. These findings indicate that the developed mathematical literacy assessment instrument meets validity and reliability criteria and is therefore suitable for use as an item bank in basic statistics courses or related subjects.

Keywords: *mathematical literacy, assessment instrument, prospective mathematics teachers, instrument development*

Abstrak Pembelajaran abad ke-21 menuntut guru, khususnya calon guru matematika, memiliki kemampuan literasi matematika yang kuat agar mampu membimbing peserta didik dalam memahami, menggunakan, dan menafsirkan konsep matematika dalam berbagai konteks kehidupan. Penelitian ini bertujuan untuk mengembangkan instrumen penilaian literasi matematika yang valid dan reliabel bagi mahasiswa calon guru matematika dengan mengadopsi model pengembangan 4-D. Instrumen yang dikembangkan difokuskan pada materi teknik membilang dan peluang dengan mengacu pada level kognitif literasi matematika AKM/IEA, yaitu knowing, applying, dan reasoning. Subjek uji coba melibatkan 64 mahasiswa program studi pendidikan matematika tingkat dua, tiga, dan empat. Hasil uji validitas isi menggunakan indeks Aiken menunjukkan bahwa seluruh butir soal valid dengan rentang nilai 0,80–0,95. Selanjutnya, uji validitas konstruk melalui Confirmatory Factor Analysis (CFA) menunjukkan validitas konvergen yang memadai dengan nilai Average Variance Extracted (AVE) sebesar 0,581 serta reliabilitas konstruk (composite reliability/ϱc) sebesar 0,815. Berdasarkan hasil tersebut, dapat disimpulkan bahwa instrumen penilaian literasi matematika yang dikembangkan telah memenuhi kriteria valid dan reliabel, sehingga layak digunakan sebagai bank soal pada mata kuliah statistika dasar atau mata kuliah sejenis.

Kata-kata Kunci: literasi matematika, instrumen penilaian, calon guru matematika, pengembangan instrumen

PENDAHULUAN

Pembelajaran abad ke-21 mengacu kepada pendekatan pendidikan yang mempersiapkan peserta didik untuk mampu menghadapi tantangan dan peluang di dunia yang terus berkembang dengan pesat. Adanya perubahan besar dalam pasar kerja global, di mana keterampilan *analytical thinking* dan *complex problem-solving* menjadi kompetensi inti yang paling dibutuhkan di masa depan (World Economic Forum, 2020). Dalam konteks ini, kebutuhan akan guru yang memiliki kemampuan dalam memahami, menggunakan dan menafsirkan matematika dalam berbagai konteks kehidupan menjadi sangat penting (OECD, 2024) karena kemampuan tersebut memungkinkan guru memberikan dukungan pembelajaran yang bermakna, kontekstual, dan responsif terhadap kebutuhan siswa. Kemampuan ini dikenal sebagai kemampuan literasi matematika, yang mencakup kapasitas individu untuk melakukan penalaran secara matematis dalam memecahkan masalah nyata (Ojose, 2011).

Dalam ekosistem pendidikan, guru memegang peranan dalam mewujudkan pendidikan berkualitas dan berkelanjutan guna mencapai target *Education 2030 Framework for Action* (UNESCO, 2021). Literasi matematika yang dikenal sebagai numerasi di Indonesia, juga menjadi pilar utama dalam Asesmen Kompetensi Minimum (AKM) yang merujuk pada standar kognitif *International Association for the Evaluation of Educational Achievement* (IEA), yaitu: *knowing, applying,* dan *reasoning* (Mullis, I. V. S. & Martin, 2017). Guru yang memiliki kemampuan literasi matematika yang baik tidak hanya mampu menyampaikan konsep-konsep dasar matematika secara efektif, tetapi juga dapat membimbing peserta didik dalam mengembangkan kemampuan literasi matematikanya. Kemampuan literasi matematika yang unggul pada guru memberikan dampak positif pada kualitas pengajaran dan pembelajaran matematika di sekolah, membantu mengurangi kesenjangan pendidikan, serta membangun fondasi yang kokoh untuk masa depan yang inovatif dan produktif (Whitney-Smith et al., 2022).

Namun demikian, terdapat kesenjangan yang nyata dalam ketersediaan instrumen evaluasi. Meskipun literasi matematika sudah menjadi standar global bagi siswa (PISA/AKM), instrumen secara spesifik dirancang untuk mengukur kompetensi mahasiswa calon guru matematika masih sangat terbatas. Padahal, literasi matematika calon guru memiliki dimensi yang lebih kompleks, mereka tidak hanya sebagai pengguna, tetapi juga calon pengembang literasi matematika bagi siswa. Tanpa instrumen yang tepat, dosen akan kesulitan memetakan kesiapan profesional mahasiswa calon guru dalam mengaplikasikan matematika pada situsi nyata.

Untuk menjamin kualitas pengembangan kemampuan tersebut, dosen perlu menyelenggarakan evaluasi yang efektif melalui kegiatan pengukuran yang tepat. Sebagaimana yang ditegaskan oleh (Mardapi, 2012) bahwa keberhasilan suatu pengukuran sangat bergantung pada kualitas instrumen yang digunakan. Instrumen dikatakan berkualitas jika memenuhi kriteria validitas dan reliabilitas yang tinggi. Instrumen dengan validitas rendah hanya akan menghasilkan data yang tidak relevan dengan tujuan pengukuran, sementara reliabilitas menjamin konsistensi hasil pengukuran (Azwar, 2012). Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk mengembangkan instrumen penilaian literasi matematika bagi mahasiswa calon guru yang valid dan reliabel, guna mendukung terciptanya calon pendidik yang kompeten di era transformasi pendidikan.

METODE

Penelitian ini termasuk penelitian dan pengembangan yang mengadopsi model pengembangan 4-D yang diperkenalkan oleh Thiagarajan et. al (1974). Model pengembangan ini terdiri dari empat tahap yaitu mendefinisikan, merancang, mengembangkan, menyebarkan. Produk akhir yang dihasilkan melalui prosedur ini adalah instrumen penilaian kemampuan literasi matematika mahasiswa yang divalidasi secara teoritis dan empiris bagi mahasiswa calon guru matematika.

Penelitian dilaksanakan di Program Studi Pendidikan Matematika Universitas Pasir Pengaraian pada semester genap Tahun Akademik 2023/2024. Subjek uji coba dalam penelitian ini melibatkan 64 mahasiswa pendidikan matematika yang berada pada tingkat dua, tiga dan empat. Pelaksanaan penelitian dilaksanakan selama rentang waktu tiga bulan, terhitung bulan April hingga Juni 2024.

Teknik analisis data dalam penelitian ini direncang untuk menguji kualitas instrumen melalui tiga aspek utama, yaitu validitas isi, validitas konstruk dan reliabilitas. Teknik untuk validitas isi dilakukan melalui penilaian ahli (*expert judgment*). Lima orang dosen ahli di bidang matermatika bertindak sebagai validator untuk menilai kesesuaian instrumen berdasarkan aspek bahasa, kesesuaian dengan level kognitif literasi matematika (merujuk kepada *Internasional Association for the Evaluation of Educational Achievement* atau IEA), relevansi materi, konteks, serta ketepatan kemungkinan penyelesaian. Kriteria penilaian oleh ahli diukur menggunakan skala 1 hingga 5, yang kemudian dikategorikan dari "Tidak Relevan" hingga "Sangat Relevan". Data penilaian oleh ahli dianalisis dengan menggunaka indeks Aiken yang dinyatakan sebagai berikut:

$$V = \frac{\sum s}{N(c-1)} \text{ dengan } s = r - l$$

Keterangan:

r : nilai penilai

c : nilai kategori tertinggi

l : nilai kategori terendah

N : jumlah penilai

(Aiken, 1985)

Kriteria penilaian butir soal disajikan pada tabel 1.

Tabel 1. Kriteria Penialian Butir Instrumen oleh Ahli

Interval	Kategori
4,5 - 5	Sangat Relevan
3,5 - 4,49	Relevan
2,5 - 3,49	Cukup Relevan
1,5 - 2,49	Kurang Relevan
1 - 1,49	Tidak Relevan

Rentang indek Aiken (V) adalah antara 0 dan 1. Semakin tinggi indek Aiken (V), maka semakin tinggi validitas isi. Dengan kata lain, indeks V mendekati nilai 1, sebaliknya semakin rendah indeks V maka nilai kevalidan sebuah butir soal juga semakin rendah (Aiken, 1985).

Validitas konstruk diuji melalui analisis faktor dengan bantuan perangkat lunak Rstudio. Validasi konstruk dijelaskan dengan persentase keragaman/ varians dari nilai *Kaiser-Mayer-Olkin* (KMO) dan *Bartlett's Test of Sphericity*. Selanjutnya uji *Confirmatory Factor Analysis* (CFA) digunakan untuk mengonfirmasi kesesuaian struktur teoritis dengan data empiris di lapangan menggunakan esttimasi *Robust Maximum Likelihood* (MLR) gunakan mengatasi data yang tidak berdistribusi normal secara multivariat.

Sementara itu, uji reliabilitas dilakukan dengan menggunakan formula *Cronbach-Alpha* dinyatakan sebagai berikut:

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

Keterangan:

$\sum \sigma_i^2$: jumlah varians butir instrumen

σ_t^2: varians skor total

k: bnayak butir pertanyaan dalam instrumen

Selain itu, penelitian ini juga menghitung *construct reliability* dan *Average Variance Extracted* (AVE) pada model CFA untuk memperkuat bukti konsistensi internal dan validitas konvergen instrumen.

HASIL DAN PEMBAHASAN

Instrumen penialian literasi matematika mahasiswa sebagai calon guru matematika dikembangkan dengan menggunakan model 4-D, yaitu mendefinisikan, merancang, mengembangkan dan menyebarkan. Fokus instrumen adalah mengukur kemampuan literasi matematika mahasiswa pada topik teknik membilang dan penghitungan peluang.

Tahap Pendefinisian

Tahap ini bertujuan untuk menetapkan dan mendefinisikan syarat pembelajaran dan menganalisis kebutuhan instrumen. Berdasarkan analisis awal ditemukan bahwa instrumen penilaian yang ada selama ini masih bersifat umum dan belum spesifik mengukur kemampuan literasi matematika mahasiswa calon guru pada konsep teknik membilang dan peluang.

Kebutuhan akan instrumen ini didorong oleh tuntutan pembelajaran abad-21 yang mengharuskan guru memiliki kemampuan memahami, menggunakan, dan menafsirkan matematika dalam berbagai konteks kehidupan (OECD, 2023). Literasi matematika bukan sekedar penguasaan rumus, melainkan kapasitas individu untuk merumuskan, menerapkan, dan menafsirkan matematika dalam situasi dunia nyata untuk memecahkan masalah. Oleh karena itu, didefisikan instrumen yang memampu memetakan kemampuan kognitif mahasiswa dalam menghadapi fenomena yang memerlukan nalar matematika.

Tahap Desain

Tahap perancangan dalam penelitian ini adalah merancang instrumen penilaian untuk konsep teknik membilang yang berkaitan dengan penghitungan peluang. Instrumen penilaian yang

dirancangan mengacu kepada *Internasional Association for the Evaluation of Educational Achievement* atau IEA, bahwa level kognitif literasi matematika-numerasi pada Asesmen Kompetensi Minimum (AKM) terdiri dari tiga level, yaitu *knowing*, *applying* dan *reasoning*.

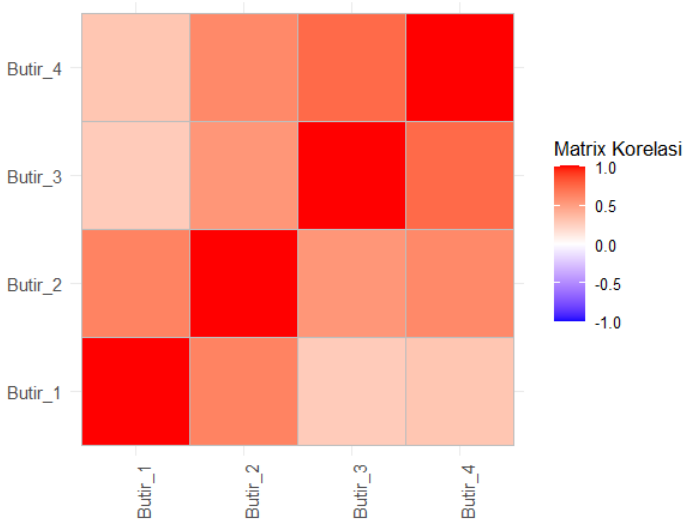
a. Level kognitif *knowing* merupakan level pertama dari literasi matematika pada AKM. Level ini berkaitan dengan aspek pengetahuan dan pemahaman, yaitu pemahaman tentang konsep dan prosedur matematika yang menjembatani penerapan matematika untuk menyelesaikan masalah dalam berbagai konteks. Soal yang dikembangkan berfokus pada kemampuan mahasiswa dalam mengambil informasi dari masalah, memahami konsep dasar, serta melakukan prosedur aljabar.
b. Level kognitif *applying* merupakan level kedua dari literasi matematika pada AKM. Level ini bekaitan dengan aspek penerapan, yaitu kemampuan mahasiswa dalam menerapkan fakta, konsep, dan prosedur matematika dalam menyelesaikan permasalahan sehari-hari. Beberapa aspek kemampuan yang termasuk pada level kognitif *applying* antara lain memilih strategi, membuat model, menerapkan fakta dan prosedur matematika untuk memecahkan masalah dunia nyata serta menafsirkan berbagai representasi matematis.
c. Level kognitif reasoning, yang mencakup kemampuan dalam menganalisis data dan informasi, membuat kesimpulan, dan memperluas pemahaman dengan kondisi baru merupakan bagian dari level *reasoning*. Beberap aspek kemampuan yang termasuk pada level kognitif ini antara lain kemampuan menganalisis informasi, mengevaluasi solusi alternatif, memberikan argumen matematis, hingga membuat kesimpulan yang valid berdasarkan fakta.

Tabel 2. Komponen dan Indikator Literasi Matematika

Level Kognitif	Indikator Pencapaian Level Kognitif
Knowing (Pengetahuan dan Pemahaman)	• Mampu mengambil atau memperoleh informasi dari masalah yang disajikan • Mampu memahami konsep dasar • Mampu melakukan prosedur aljabar yang efektif
Applying (Penerapan)	• Mampu mengetahui informasi yang terkandung dalam masalah yang disajikan • Mampu menentukan operasi, strategi, dan aturan yang sesuai dan efisien untuk memecahkan masalah dunia nyata yang dapat diselesaikan dengan menggunakan berbagai metode • Mahasiswa mampu menerapkan/melaksanakan strategi dan operasi untuk memecahkan masalah dunia nyata yang berkaitan dengan konsep dan prosedur matematika yang dikenal. • Mampu menyatakan hasil berdasarkan operasi pemecahan masalah
Reasoning (Penalaran)	• Mahasiswa mampu menganalisis informasi sebagai tahap awal dalam memecahkan masalah

Level Kognitif	Indikator Pencapaian Level Kognitif
	• Mahasiswa mampu berpikir, memilih strategi, dan solusi alternatif pemecahan masalah • Mahasiswa mampu membuat kesimpulan yang valid berdasarkan informasi dan fakta-fakta • Mahasiswa mampu memberikan argumen matematis untuk mendukung klaim

Draf instrumen terdiri dari 4 butir soal yang merepresentasikan level kognitif sebagai bank soal potensial untuk mata kuliah statistika matematika. Desain instrumen penilaian literasi matematika mahasiswa untuk konsep teknik membilang yang berkaitan dengan perhitungan peluang. Nilai korelasi antar butir berkisar antara 0.658 hingga 0.742, menunjukkan abwa setiap butir memiliki konstribusi yang cukup terhadap total skor. Visualisasi matrik korelasi terlihat bahwa derajat korelasi dalam kategori rendah dan sedang, serta tanpa indikasi butir yang menyimpang secara signifikan. Selain itu, nilai cronbach's alpha sebesar 0,811 (cut-off ≥ 0,7) dan tidak ada peningkatan signifikan nilai alpha ketika salah satu butir dihapus. Artinya semua butir memberikan kontribusi yang proporsional terhadap konsistensi internal instrumen, yaitu semua butir saling memperkuat dalam mengukur konstruk tunggal literasi matematika.



Gambar 1. Matriks Korelasi antar Butir

Tahap Pengembangan

Tahap ini bertujuan untuk menghasilkan instrumen final setelah melalui serangkaian uji coba dan revisi berdasarkan data empiris. Tahap pengembangan ini meliputi validitas isi, validitas konstruk dan reliabilitas.

Validitas Isi

Validitas dilakukan melalui penilaian ahli (*expert judgement*) oleh lima orang dosen bidang pendidikan matematika sebagai validator. Hasil dari telaah awal ini menunjukkan bahwa

instrumen yang telah dirancang belum baik. Berdasarkan masukan validator dilakukan perbaikan pada aspek kebahasaan agar tidak ambigu serta penambahan ilustrasi visual berupa gambar dan tabel untuk memperkuat konteks dunia nyata pada butir soal. Hasil komentar/saran validator dan revisinya disajikan pada tabel 3.

Tabel 3. Komentar, saran dan revisi instrumen

<table>
<tr><th>Sebelum Revisi</th><th>Setelah Revisi</th></tr>
<tr><td>Soal 1.
Dalam suatu pertemuan antar keluarga yang berjumlah 10 orang. Semua keluarga duduk pada suatu meja bundar. Ada berapa banyak cara duduk anggota keluarga tersebut dimana setiap anggota keluarga duduk berdekatan?</td><td>Dalam suatu pertemuan anggota keluarga yang berjumlah 10 orang, semua anggota keluarga duduk di sekitar meja bundar. Berapa banyak cara duduk anggota keluarga tersebut?</td></tr>
<tr><td colspan="2">Komentar/Saran: semua validator memberikan komentar dan saran terkait kesesuaian bahasa yang digunakan diantaranya:
1. Kalimat 1 berupa keterangan dan belum sebuah kalimat utuh. Kalimat 1 dan kalimat 2 dapat digabung menjadi satu kalimat utuh
2. Beberapa frase menimbulkan makna ambigu. Misalnya “semua keluarga” dan “setiap anggota keluarga duduk berdekatan” bermakna ada beberapa keluarga.
3. Soal no 1 dikategorikan baik dilihat dari segi kesesuaian soal dengan indikator literasi matematika untuk level kognitif knowing dan konteks yang digunakan sudah cukup relevan dengan situasi dunia nyata</td></tr>
<tr><td>Soal 2.
Dalam suatu pertemuan antar keluarga. Semua keluarga duduk pada suatu meja bundar. keluarga yang hadir sebanyak 4 keluarga dengan total 16 orang. Keluarga Bapak Ali berjumlah 4 orang, keluarga Bapak Edi berjumlah 3 orang, keluarga 4 orang dan keluarga Bapak Mustofa berjumlah 5 orang. Ada berapa banyak cara duduk anggota keluarga terseb dimana setiap anggota keluarga duduk berdekatan? Jelaskan jawabanmu dan bagaimana Anda menerapkan konsep permutasi dalam konteks ini.</td><td>
Dalam suatu pertemuan antar angg keluarga, semua keluarga duduk mengelil sebuah meja. Keluarga yang hadir sebanyak 16 orang dengan 4 kepala keluar
Adapun rincian jumlah masing-masing anggota keluarga disajikan pada tabel berik
<table><tr><td>Nama Kepala Keluarga</td><td>Jumlah Anggota Keluarga</td></tr><tr><td>Bapak Ali</td><td>4</td></tr><tr><td>Bapak Edi</td><td>3</td></tr><tr><td>Bapak Rahmad</td><td>4</td></tr><tr><td>Bapak Mustofa</td><td>5</td></tr></table>
Berapa banyak cara duduk anggota keluarga tersebut dengan ketentuan setiap angg keluarga dalam satu Kepala Keluarga duduk berdekatan?</td></tr>
<tr><td colspan="2">Komentar/saran:
1. Menambahkan gambar untuk mengilustrasikan situasi nyata atau masalah praktis yang relevan sehingga membuat soal lebih kontektual dan bermakna. Dalam hal ini gambar bahwa semua anggota keluarga duduk mengelilingi sebuah meja.
2. Menambahkan tabel untuk mempermudah mahasiswa konsteks soal. Awalnya rincian anggota dinarasikan dan kemudian direvisi menggunakan tabel.
3. Pertanyaan yang diberikan dapat digeneral. Dengan pertanyaan meminta mahasiswa menentukan banyak cara berarti secara tidak langsung sudah meminta mahasiswa untuk menentukan dan menerapkan strategi dan operasi untuk memecahkan masalah tersebut
4. Secara keseluruhan, soal ini sesuai dengan indikator literasi matematika untuk level kognitif apllying dan konteks yang relavan dengan kehidupan sehari-hari, serta</td></tr>
</table>

<table>
<tr><th>Sebelum Revisi</th><th>Setelah Revisi</th></tr>
<tr><td colspan="2">kemungkinan penyelesaiannya sudah dirancang bagaimana mahasiswa dapat menentukan dan menerapkan konsep dan prosedur matematika dengan tepat untuk menyelesaikan masalah dunia nyata.</td></tr>
<tr><td>Soal 3.
Seorang mahasiswa matematika harus mengerjakan 7 dari 10 soal yang diberikan dalam uj akhir semester. Dari 10 soal tersebut, 5 soal pertama berkaitan dengan topik aljabar linier, materi yang sangat dikuasai oleh mahasiswa tersebut. Sedangkan 5 soal lainnya berkaitan dengan beberapa topik, yaitu kalkulus, statistika dan geometri. Mahasiswa tersebut ingin memastikan bahwa dia mengerjakan paling sedikit 3 dari soal aljabar linier untuk memaksimalkan kemampuannya dalam topik tersebut. Hitunglah peluang bahwa mahasisw tersebut akan mengerjakan paling sedikit 3 dari 5 soal aljabar linier. Jelaskan langkah-langkah yang Anda lakukan untuk menyelesaikan masalah ini dan bagaimana anda menerapkan konsep kombinasi dan pelunag dalam konteks ini?</td><td>Seorang mahasiswa matematika harus mengerjakan 7 dari 10 soal yang diberikan dalam akhir semester. Dari 10 soal tersebut, 5 soal pertama berkaitan dengan topik aljabar linie materi yang sangat dikuasai oleh mahasiswa tersebut. Sedangkan 5 soal lainnya berkaita dengan beberapa topik, yaitu kalkulus, statistika dan geometri. Mahasiswa tersebut ingin memastikan bahwa dia mengerjakan paling sedikit 3 dari soal aljabar linier untuk memaksimalkan kemampuannya dalam topik tersebut. Berapa peluang bahwa mahasiswa tersebut akan mengerjakan paling sedikit 3 dari 5 soal aljabar linier?</td></tr>
<tr><td colspan="2">Komentar/saran:
1. Pertanyaan pada soal yang dirancang sudah memiliki solusi yang jelas bagaimana mahasiswa dapat memahami informasi yang ada, berpikir serta memilih strategi untuk memecahkan masalah dan menarik kesimpulan yang tepat berdasarkan informasi/fakta yang diperoleh. Jadi terkait langkah-langkah yang menjadi pertanyaan dalam soal dapat dibuang saja karena sudah terwakili dengan pertanyaan pertama.
2. Secara keseluruhan, soal ini sesuai dengan indikator literasi matematika untuk level kognitif reasoning yang mengajak mahasiswa untuk berpikir menganalisis, memadukan fakta, konsep, dan prosedur dalam memecahkan masalah hingga membuat kesimpulan yang valid.</td></tr>
<tr><td>Soal 4.
Di sebuah perguruan tinggi, terdapat kelompok mahasiswa yang merencanakan kegiatan wisata. Terdapat 15 mahasiswa yang awalnya direncanakan untuk mengikuti kegiatan tersebut dengan menggunakan 3 mobil. Mobil pertama memiliki kapasitas 5 orang, mobi kedua memiliki kapasitas 4 orang, dan mobil ketiga memiliki kapasitas 6 orang. Namun, mahasiswa berhalang ikut dalam kegiatan tersebut, sehingga hanya 14 mahasiswa yang a dialokasikan ke dalam 3 mobil. Tentukanlah banyak cara berbeda untuk mengalokasikan mahasiswa ke dalam 3 mobil dengan kapasitas mobil masing-masing yang telah ditentuk Berdasarkan perhitungan tersebut, apakah dengan pertimbangan mobil ketiga tidak terisi penuh akan lebih fleksibel dibandingkan mobil lainnya yang tidak terisi penuh? Berikan argument yang mendukung kesimpulan tersebut!</td><td>
Di sebuah perguruan tinggi, terdapat kelompok mahasis yang merencanakan kegiatan wisata. Awalnya terdapat mahasiswa yang berencana untuk mengikuti kegia tersebut dengan menggunakan 3 mobil.
Rincian kapasitas mobil disajikan pada tabel berikut:
<table><tr><th>Jenis Mobil</th><th>Kapasitas Mobil</th></tr><tr><td>Mobil 1</td><td>5</td></tr><tr><td>Mobil 2</td><td>4</td></tr><tr><td>Mobil 3</td><td>6</td></tr></table>Namun, satu mahasiswa berhalangan ikut dalam kegiatan tersebut, sehingga hanya mahasiswa yang akan ditempatkan ke dalam 3 mobil.
a. Tentukanlah banyak cara berbeda untuk menempatkan 14 mahasiswa ke dalam 3 mo dengan kapasitas mobil masing-masing yang telah ditentukan?
b. Berdasarkan perhitungan tersebut, apakah dengan pertimbangan mobil ketiga tic terisi penuh akan lebih fleksibel dibandingkan mobil lainnya yang tidak terisi penu Berikan argumen yang mendukung kesimpulan tersebut!</td></tr>
<tr><td colspan="2">Komentar/saran:
1. Menambahkan gambar dan tabel untuk mengilustrasikan masalah praktis yang relevan sehingga membuat soal lebih kontektual, bermakna dan mudah dipahami. Dalam hal ini gambar ditambahkan 3 buah mobil sesuai konteks soal dan tabel untuk merepresentasikan kapasitas mobil. Hal ini harapan membuka pemikiran mahasiswa bahwa bakal ada mobil yang tidak terisi penuh.
2. Dari segi kemungkinan penyelesaian dari soal yang dirancang sudah memandu mahasiswa untuk menyelesaikan masalah melalui berpikir untuk menganalisis, mengevaluasi hingga memberikan argumen dari hasil yang diperoleh.</td></tr>
</table>

Sebelum Revisi	Setelah Revisi
3. Secara keseluruhan, soal ini sesuai dengan indikator literasi matematika untuk level kognitif *reasoning* sangat baik untuk mengukur kemampuan mahasiswa untuk menganalisis, mengevaluasi, menyimpulkan dan membuat justifikasi.	

Setelah selesai diperbaiki, instrumen penilaian diberikan kembali kepada validator untuk dinilai masing-masing butir soalnya. Hasil penilaian dianalisis menggunakan rumus Aiken untuk mengetahui validasi setiap butir soal.

Tabel 4. Hasil Kevalidan Butir

No.	Nilai *V* Aiken	Validitas
1.	0,80	Sedang
2.	0,95	Tinggi
3.	0,85	Tinggi
4.	0,85	Tinggi

Berdasarkan tabel 4 terlihat bahwa hasil penilaian validator dengan nilai V Aiken mendekati 1 pada semua butir. Artinya instrumen penilaian literasi matematika sudah valid, dengan rincian soal 1 dikategori validitas medium dan soal 2,3 dan 4 dikategori validitas tinggi. Selanjutnya instrumen ini sudah siap untuk diujicobakan. Instrumen yang valid sudah layak untuk melanjutkan uji coba (Arifin & Retnawati, 2017).

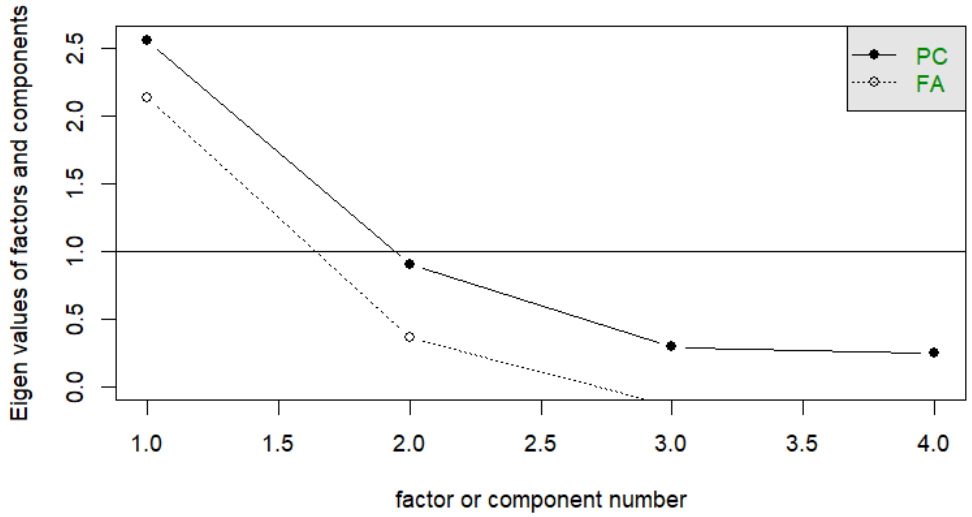
Validitas Konstruk melalui CFA

Instrumen diujicobakan kepada 64 mahasiswa untuk membuktikan validitas konstruk secara empiris. Data hasil ujicoba dilakukan analisis faktor untuk mengelompokkan variabel dengan karakteristik yang serupa ke dalam satu faktor. Dalam hal ini, pendekatan yang digunakan adalah *Confirmatory Factor Analysis* (CFA). Sebelumnya dilakukan pemeriksaan kecukupan sampel dan kelayakan matriks korelasi. Uji kelayakan Kaiser-Meyer-Olkin (KMO) menunjukkan nilai *Measure of Sampling Adequacy* (MSA) keseluruhan sebesar 0,67 (Kaiser, 1974). Nilai MSA masing-masing butir berkisar antara 0,61 hingga 0,67, menunjukkan kecukupan sampel yang sangat baik untuk analisis factor (Hair et al., 2020). Lebih jauh, uji Bartlett menunjukkan korelasi yang signifikan di antara variabel dan matriks korelasi yang terbentuk berbeda secara signifikan dari matriks identitas. Hal ini ditunjukkan, dengan statistik chi-kuadrat sebesar 188.5646 dan nilai signifikansi p ($=5.137 \times 10^{-38} < 0,01$. Oleh karena itu, data layak untuk lebih lanjut menggunakan CFA.

Uji normalitas multivariat dilakukan menggunakan *Mardia's Test*, yang terdiri dari dua komponen utama, yaitu Mardia *Skewness* dan Mardia *Kurtosis*. Hasil pengujian diperoleh nilai Mardia Skewness sebesar 59.044 dengan nilai p sebesar $1,000 \times 10^{-5}$, sedangkan nilai Mardia Kutosis sebesar 2.300 dengan nilai-p sebesar $2,14 \times 10^{-2}$. Hal ini menunjukkan bahwa adanya penyimpangan distirbusi, yaitu data tidak memenuhi asumsi normalitas multivariate (nilai p <

0,05). Oleh karena itu, *Confirmatory Factor Analysis* (CFA) menggunakan estimator *Robust Maximum Likelihood* (MLR).

Analisis pendahuluan terhadap struktur faktor dilakukan melalui pemeriksaan *scree plot* dan nilai eigen. Berdasarkan kriteria **Kaiser-Guttman** bahwa hanya faktor dengan eigenvalue lebih besar atau sama dengan satu yang dianggap signifikan. Gambar 3 *scree plot* menunjukkan bahwa baik pada pendekatan Principal Component Analysis (PCA) maupun Factor Analysis (FA), hanya terdapat satu faktor utama yang memiliki eigenvalue > 1. Untuk PCA, komponen pertama memiliki eigenvalue 2.56, sedangkan untuk FA, faktor pertama memiliki nilai 2.14. Sementara itu, faktor-faktor berikutnya memiliki eigenvalue jauh di bawah 1, bahkan negatif pada FA, yang menunjukkan bahwa kontribusinya terhadap varians sangat kecil dan tidak layak dipertahankan. Bentuk scree plot juga menunjukkan penurunan tajam setelah faktor pertama, membentuk pola "elbow" yang klasik, yang menguatkan kesimpulan bahwa struktur instrumen cenderung unidimensional.



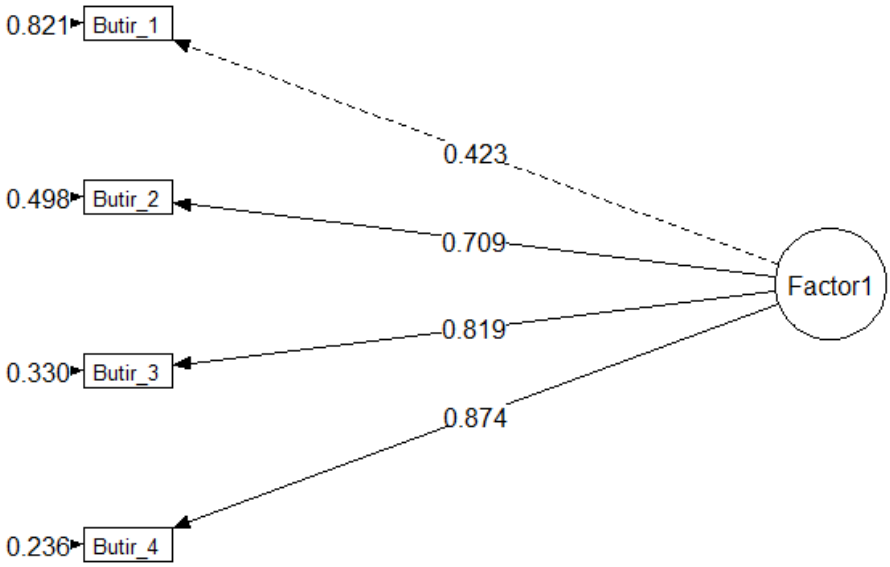
Gambar 2. Scree Plot

Validitas konstruk menggunakan CFA dengan estimator MLR. Dua model diuji untuk mengetahui kesesuaian struktur teoritis dengan data empiris., yaitu model 1 yang mengasumsikan seluruh indikator berpikir kritis membuat satu faktor, dan model 2 membagi indikator menjadi dua faktor sesuai dengan hasil eksplorasi dan kerangka konseptual. Perbandingan hasil CFA kedua model disajikan pada tabel 4.

Tabel 5. Hasil CFA

Indikator Kecocokan Model	Model 1 (1 Faktor)
Chi-Square (χ^2)	39.540 (df = 2, p < .001)
CFI (≥0,95)	0.797
TLI (≥0,95)	0.390

Indikator Kecocokan Model	Model 1 (1 Faktor)
RMSEA (≤ 0,05)	0.418 (CI: .314–.531)
SRMR (≤ 0,05)	0.111
AIC	1918.603
BIC	1940.060
R^2 Tertinggi	0.764 (Butir_4)



Gambar 3. Struktur Model Konfirmatori

Average Variance Extracted (AVE)* dan *construct reliability (rhoC)

Validitas dan reliabilitas konstruk dari model CFA, dilakukan penghitung *Average Variance Extracted* (AVE) untuk merepresentasikan proporsi varians indikator yang dijelaskan oleh konstruk laten, dan *construct reliability* (rhoC) untuk menunjukkan konsistensi internal dalam satu konstruk. Nilai AVE diperoleh sebesar 0.581 varians indikator dijelaskan oleh konstruk laten, melewati ambang batas minimum 0,5 sebagaimana dikemukakan(Fornell & Larcker, 1981). Hal ini mengindikasikan bahwa konstruk memiliki validitas konvergen yang memadai. Selain itu, nilai reliabilitas rhoC sebesar 0.815 menunjukkan konsistensi internal yang baik dan reliabel dalam mengukur konstruk yang dimaksud. Hasil estimasi reliabilitas yang diperoleh menurut (Ul Hassan & Miller, 2019) dan (Nitko, Anthony J & Brookhart, 2014), telah diterima untuk instrumen yang digunakan dalam penilaian skala kecil.

Tahap Penyebaran

Tahap penyebaran merupakan fase akhir dari model 4-D yang bertujuan untuk mengomunikasikan hasil pengembangan dan mengaplikasikan instrumen pada lingkup yang lebih luas. Setelah instrumen dinyatakan valid dan reliabel melalui uji coba lapangan, draf final

kemudian didistribusikan sebagai bank soal untuk mendukung evaluasi pada mata kuliah Statistika Dasar atau Statistika Matematika di Program Studi Pendidikan Matematika. Instrumen ini memiliki persentase varians yang dapat dijelaskan sebesar 78,6%. Hal ini mengindikasikan bahwa variabel-variabel yang diukur memiliki kontribusi yang sangat signifikan dalam merepresentasikan kemampuan literasi matematika mahasiswa.

SIMPULAN DAN SARAN

Penelitian ini berhasil menjawab kebutuhan mendesak akan instrumen evaluasi yang valid dan reliabel untuk mengukur literasi matematika mahasiswa calon guru. Berdasarkan hasil pengembangan dan uji coba, instrumen ini tidak hanya memenuhi standar teknis psikometri, tetapi juga memiliki substansi pemaknaan sebagai alat ukur yang komprehensif untuk memetakan level kognitif *knowing, applying*, hingga *reasoning*. Kesesuaian antara kerangka teoritis dengan data empiris di lapangan membuktikan bahwa instrumen ini mampu memberikan gambaran akurat mengenai kesiapan profesional mahasiswa dalam menafsirkan dan memecahkan masalah matematika dalam berbagai konteks kehidupan nyata.

Secara praktis, instrumen ini sangat relevan untuk diaplikasikan sebagai bank soal pada mata kuliah statistika dasar atau matematika lanjut guna membiasakan mahasiswa dengan aktivitas berpikir tingkat tinggi. Prospek pengembangan ke depan dapat difokuskan pada perluasan cakupan materi matematika lainnya di luar teknik membilang dan peluang, serta pengujian instrumen pada populasi mahasiswa calon guru yang lebih luas di berbagai wilayah. Dengan demikian, instrumen ini diharapkan dapat berkontribusi secara berkelanjutan dalam meningkatkan kualitas calon pendidik yang kompeten dalam menghadapi tantangan pendidikan global abad ke-21.

UCAPAN TERIMA KASIH

Penulis menyampaikan terima kasih yang sebesar-besarnya kepada Program Studi Pendidikan Matematika Universitas Pasir Pengaraian yang telah memfasilitasi pelaksanaan penelitian ini. Apresiasi setinggi-tingginya diberikan kepada para validator yang telah memberikan masukan berharga demi penyempurnaan instrumen ini. Selain itu, terima kasih juga disampaikan kepada seluruh mahasiswa pendidikan matematika tingkat dua, tiga, dan empat yang telah berpartisipasi aktif sebagai subjek uji coba dalam penelitian pengembangan ini.

DAFTAR PUSTAKA

Aiken, L. R. (1985). Three coefficients for analyzing the reliability and validity of ratings. *Educational and Psychological Measurement, 45*(1), 131–142.

Arifin, Z. (2017). Kriteria instrumen dalam suatu penelitian. *Jurnal Theorems (The Original Research of Mathematics), 2*(1), 28–36.

Arifin, Z., & Retnawati, H. (2017). Pengembangan instrumen pengukur higher order thinking skills matematika siswa SMA kelas X. *PYTHAGORAS: Jurnal Pendidikan Matematika, 12*(1), 98–108.

Azwar, S. (2012). *Penyusunan skala psikologis* (Edisi ke-2). Pustaka Pelajar.

Fornell, C., & Larcker, D. F. (1981). Structural equation models with unobservable variables and measurement error: Algebra and statistics. *Journal of Marketing Research, 18*(3), 382–388.

Hair, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research, 109*, 101–110. https://doi.org/10.1016/j.jbusres.2019.11.069

Kaiser, H. F. (1974). An index of factorial simplicity. *Psychometrika, 39*(1), 31–36. https://doi.org/10.1007/BF02291575

Mardapi, D. (2012). *Pengukuran, penilaian, dan evaluasi pendidikan*. Nuha Medika.

Mullis, I. V. S., & Martin, M. O. (2017). *TIMSS 2019 assessment framework*. International Association for the Evaluation of Educational Achievement (IEA).

Nitko, A. J., & Brookhart, S. M. (2014). *Educational assessment of students* (7th ed.). Pearson Education.

OECD. (2023). *PISA 2022 assessment and analytical framework*. OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-2022-assessment-and-analytical-framework_dfe0bf9c-en

OECD. (2024). *PISA 2022 results (Volume III): Creative minds, creative schools*. OECD Publishing. https://www.oecd-ilibrary.org/education/pisa-2022-results-volume-iii

Ojose, B. (2011). Mathematics literacy: Are we able to put the mathematics we learn into everyday use? *Journal of Mathematics Education, 4*(1), 89–100.

Retnawati, H. (2016). *Analisis kuantitatif instrumen penelitian*. Parama Publishing.

Ul Hassan, M., & Miller, F. (2019). Optimal item calibration for computerized achievement tests. *Psychometrika, 84*(4), 1101–1128. https://doi.org/10.1007/s11336-019-09673-6

UNESCO. (2021). *Reimagining our futures together: A new social contract for education*. UNESCO.

World Economic Forum. (2020). *The future of jobs report 2020*. World Economic Forum.