

ANALISIS VALIDITAS DAN RELIABILITAS INSTRUMEN LITERASI MATEMATIKA PADA MAHASISWA

Ali Hasmy¹, Yumi Sarassanti², Mutya Hairunisa³

^{1,2,3}Tadris Matematika, FTIK, IAIN Pontianak, Jl. Letnan Jenderal Soeprapto No. 19

Corresponding author : alihazmy@iainptk.ac.id

Abstrak: Penelitian ini bertujuan menganalisis validitas dan reliabilitas instrumen literasi matematika yang dikembangkan untuk mengukur kemampuan literasi matematika mahasiswa. Penelitian menggunakan pendekatan kuantitatif deskriptif dengan data simulasi yang terdiri atas respons 18 mahasiswa terhadap 10 butir soal literasi matematika. Analisis instrumen meliputi validitas isi menggunakan indeks Aiken's V, validitas butir menggunakan korelasi Corrected Item-Total Correlation, serta reliabilitas internal menggunakan koefisien Cronbach's Alpha. Hasil analisis menunjukkan bahwa seluruh butir instrumen memiliki nilai Aiken's V pada rentang 0,83–1,00 sehingga memenuhi kriteria validitas isi yang tinggi. Nilai korelasi setiap butir lebih besar dari 0,30 yang menunjukkan bahwa seluruh butir valid secara empiris. Selain itu, instrumen memperoleh nilai Cronbach's Alpha sebesar 0,892 yang mengindikasikan tingkat reliabilitas sangat tinggi. Berdasarkan hasil tersebut, instrumen literasi matematika dapat digunakan sebagai ilustrasi metodologis dalam analisis instrumen, namun masih memerlukan pengujian menggunakan data empiris pada sampel yang lebih luas.

Kata kunci: literasi matematika, validitas instrumen, reliabilitas, Cronbach's Alpha, Aiken's V.

Abstract : This study aims to analyze the validity and reliability of a mathematical literacy instrument developed to measure university students' mathematical literacy skills. The study employed a descriptive quantitative approach using simulated data from 18 students who responded to 10 mathematical literacy items. Instrument evaluation included content validity using Aiken's V index, item validity using Corrected Item-Total Correlation, and internal consistency reliability using Cronbach's Alpha coefficient. The findings indicated that all items achieved Aiken's V values ranging from 0.83 to 1.00, demonstrating high content validity. Furthermore, all item-total correlation coefficients exceeded 0.30, confirming satisfactory empirical validity. The instrument also yielded a Cronbach's Alpha coefficient of 0.892, indicating excellent internal consistency reliability. These findings indicate that the instrument serves as a methodological illustration; however, further validation using empirical data from a larger sample is recommended.

Keywords: mathematical literacy, instrument validity, reliability, Cronbach's Alpha, Aiken's V.

PENDAHULUAN

Literasi matematika merupakan salah satu kompetensi utama yang harus dimiliki mahasiswa pada abad ke-21. Kemampuan ini tidak hanya berkaitan dengan penguasaan konsep, prosedur, maupun operasi matematika, tetapi juga mencakup kemampuan merumuskan, menerapkan, dan menafsirkan matematika dalam berbagai situasi kehidupan nyata. Dalam kerangka Programme for International Student Assessment (PISA), literasi matematika didefinisikan sebagai kemampuan individu untuk bernalar secara matematis serta menggunakan konsep, prosedur, fakta, dan alat matematika dalam memecahkan berbagai permasalahan kontekstual sehingga mampu mengambil keputusan secara logis dan bertanggung jawab (OECD, 2023). Oleh karena itu, pengembangan literasi matematika telah menjadi salah satu fokus utama dalam pembelajaran matematika di berbagai jenjang pendidikan, termasuk pendidikan tinggi.

Perkembangan paradigma pembelajaran yang menekankan kemampuan berpikir tingkat tinggi (Higher Order Thinking Skills/HOTS), pemecahan masalah, komunikasi matematis, dan penalaran semakin memperkuat pentingnya literasi matematika sebagai capaian pembelajaran mahasiswa. Mahasiswa tidak hanya dituntut mampu menyelesaikan persoalan secara prosedural, tetapi juga mampu menginterpretasikan informasi, mengevaluasi argumen, serta mengaitkan konsep matematika dengan berbagai fenomena kehidupan, perkembangan teknologi, maupun kebutuhan dunia kerja. Oleh sebab itu, pengukuran kemampuan literasi matematika perlu dilakukan secara tepat menggunakan instrumen yang memiliki kualitas psikometrik yang baik.

Keakuratan hasil pengukuran sangat ditentukan oleh kualitas instrumen yang digunakan. Instrumen yang tidak memenuhi kriteria validitas dan reliabilitas akan menghasilkan informasi yang bias sehingga dapat

mengurangi ketepatan dalam pengambilan keputusan pendidikan. Validitas menunjukkan sejauh mana suatu instrumen mampu mengukur konstruk yang seharusnya diukur, sedangkan reliabilitas menunjukkan tingkat konsistensi hasil pengukuran ketika instrumen digunakan pada kondisi yang relatif sama (Crocker & Algina, 2008). Dalam penelitian pendidikan, kedua aspek tersebut merupakan persyaratan utama sebelum suatu instrumen digunakan sebagai alat pengumpulan data.

Meskipun berbagai penelitian telah mengembangkan instrumen literasi matematika, masih ditemukan sejumlah kelemahan, antara lain belum optimalnya pengujian validitas isi melalui penilaian pakar, kurang lengkapnya analisis validitas empiris setiap butir soal, serta terbatasnya pelaporan reliabilitas instrumen secara komprehensif. Selain itu, sebagian penelitian hanya berfokus pada hasil belajar tanpa memberikan informasi mengenai kualitas instrumen yang digunakan. Padahal, instrumen yang berkualitas merupakan fondasi utama dalam memperoleh data penelitian yang akurat dan dapat dipercaya. Oleh karena itu, proses pengembangan instrumen perlu disertai dengan pengujian validitas isi, validitas butir, serta reliabilitas internal agar instrumen benar-benar layak digunakan dalam penelitian maupun evaluasi pembelajaran.

Dalam konteks pendidikan matematika, validitas isi umumnya dilakukan melalui penilaian para ahli (*expert judgment*) menggunakan indeks Aiken's V. Metode ini memberikan informasi mengenai tingkat kesesuaian setiap butir soal dengan indikator kemampuan yang akan diukur. Setelah validitas isi terpenuhi, analisis dilanjutkan dengan pengujian validitas empiris menggunakan korelasi Corrected Item–Total Correlation untuk mengetahui kemampuan setiap butir dalam membedakan responden berdasarkan kemampuan yang dimiliki. Selanjutnya, reliabilitas internal instrumen dapat dianalisis menggunakan koefisien Cronbach's Alpha yang banyak digunakan dalam penelitian pendidikan karena mampu menunjukkan tingkat konsistensi antarbutir dalam suatu instrumen (Taber, 2018).

Penggunaan instrumen yang telah memenuhi kriteria validitas dan reliabilitas memberikan berbagai manfaat, antara lain meningkatkan akurasi pengukuran, meminimalkan kesalahan pengukuran (*measurement error*), serta menghasilkan data yang lebih objektif dan dapat dipertanggungjawabkan secara ilmiah. Instrumen

yang baik juga memungkinkan peneliti memperoleh gambaran kemampuan literasi matematika mahasiswa secara lebih tepat sehingga hasil penelitian dapat dijadikan dasar dalam penyusunan kebijakan pembelajaran maupun pengembangan kurikulum.

Sejalan dengan implementasi Kurikulum Merdeka dan meningkatnya perhatian terhadap pengembangan kompetensi abad ke-21, kebutuhan akan instrumen literasi matematika yang berkualitas menjadi semakin penting. Berbagai penelitian terkini menunjukkan bahwa kemampuan literasi matematika peserta didik di Indonesia masih memerlukan peningkatan sehingga diperlukan instrumen yang mampu mengidentifikasi kemampuan tersebut secara akurat (OECD, 2023). Selain itu, National Council of Teachers of Mathematics (NCTM, 2023) menegaskan bahwa asesmen matematika hendaknya tidak hanya mengukur kemampuan prosedural, tetapi juga kemampuan bernalar, memecahkan masalah, berkomunikasi, dan menghubungkan berbagai konsep matematika dalam kehidupan nyata.

Berdasarkan uraian tersebut, analisis validitas dan reliabilitas instrumen literasi matematika menjadi tahapan yang sangat penting dalam proses pengembangan instrumen penelitian. Artikel ini bertujuan menganalisis kualitas instrumen literasi matematika melalui pengujian validitas isi, validitas butir, dan reliabilitas internal menggunakan data simulasi dari 18 mahasiswa sebagai ilustrasi metodologis. Penggunaan data simulasi dimaksudkan untuk memberikan contoh penerapan prosedur analisis instrumen secara sistematis, sehingga dapat menjadi referensi bagi peneliti, mahasiswa, maupun praktisi pendidikan matematika dalam mengembangkan instrumen penelitian yang memenuhi standar kualitas pengukuran.

METODE PENELITIAN

Penelitian ini menggunakan pendekatan kuantitatif deskriptif yang bertujuan untuk menganalisis kualitas instrumen literasi matematika melalui pengujian validitas dan reliabilitas. Penelitian difokuskan pada evaluasi karakteristik psikometrik instrumen sehingga diperoleh informasi mengenai kelayakan instrumen sebagai alat ukur kemampuan literasi matematika mahasiswa. Analisis dilakukan menggunakan data simulasi yang disusun untuk tujuan pembelajaran metodologi penelitian. Penggunaan data simulasi dinyatakan secara

eksplisit agar tidak ditafsirkan sebagai data empiris hasil pengumpulan lapangan.

Instrumen yang dianalisis berupa tes literasi matematika yang terdiri atas 10 butir soal uraian. Penyusunan instrumen mengacu pada kerangka literasi matematika Programme for International Student Assessment (PISA) 2022, yang mencakup tiga proses utama, yaitu merumuskan masalah secara matematis (formulate), menggunakan konsep dan prosedur matematika (employ), serta menafsirkan dan mengevaluasi hasil matematika (interpret). Selain itu, pengembangan indikator juga mempertimbangkan kompetensi berpikir kritis, penalaran matematis, komunikasi matematis, dan pemecahan masalah sesuai rekomendasi National Council of Teachers of Mathematics (NCTM).

Tahapan penelitian terdiri atas lima langkah, yaitu: (1) penyusunan kisi-kisi instrumen berdasarkan indikator literasi matematika; (2) penyusunan butir soal sesuai indikator; (3) penilaian validitas isi oleh tiga orang ahli pendidikan matematika; (4) penyusunan data simulasi respons mahasiswa; dan (5) analisis validitas butir serta reliabilitas instrumen menggunakan teknik statistik.

Data simulasi terdiri atas 18 mahasiswa, yang masing-masing memberikan respons terhadap 10 butir soal. Jumlah responden dipilih semata-mata untuk memberikan ilustrasi prosedur analisis instrumen dan bukan untuk menarik kesimpulan empiris mengenai populasi mahasiswa. Skor setiap butir berada pada rentang 0–10, sehingga skor total maksimum yang dapat diperoleh adalah 100.

Tabel 1. Kisi-kisi Instrumen

Indikator Literasi Matematika	Nomor Butir	Bentuk Soal
Merumuskan masalah matematis	1, 2, 3	Uraian
Menggunakan konsep dan prosedur matematika	4, 5, 6, 7	Uraian
Menafsirkan hasil penyelesaian	8, 9, 10	Uraian

Sebelum dilakukan pengujian empiris, instrumen divalidasi oleh tiga validator yang memiliki

kompetensi di bidang pendidikan matematika dan evaluasi pembelajaran. Validitas isi dianalisis menggunakan indeks Aiken's V karena metode ini mampu mengukur tingkat kesesuaian setiap butir dengan indikator yang dikembangkan berdasarkan penilaian para ahli.

Interpretasi nilai Aiken's V mengacu pada kriteria berikut.

Tabel 2. Kriteria Nilai Aiken's V

Nilai Aiken's V	Kriteria
$\geq 0,80$	Sangat valid
0,60–0,79	Valid
$< 0,60$	Perlu revisi

Selanjutnya dilakukan pengujian validitas empiris menggunakan korelasi Corrected Item–Total Correlation. Analisis ini bertujuan mengetahui kemampuan setiap butir dalam mengukur konstruk yang sama dengan skor total instrumen. Suatu butir dinyatakan valid apabila memiliki nilai korelasi lebih besar dari 0,30.

Reliabilitas instrumen dianalisis menggunakan koefisien Cronbach's Alpha yang menunjukkan konsistensi internal antarbutir soal. Perhitungan dilakukan menggunakan perangkat lunak IBM SPSS Statistics versi 27. Interpretasi reliabilitas mengacu pada kriteria berikut.

Tabel 3. Interpretasi Realibilitas

Cronbach's Alpha	Interpretasi
$\geq 0,90$	Sangat tinggi
0,80–0,89	Tinggi
0,70–0,79	Cukup
0,60–0,69	Rendah
$< 0,60$	Sangat rendah

Selain validitas dan reliabilitas, analisis juga dilakukan terhadap statistik deskriptif instrumen yang meliputi nilai rata-rata, simpangan baku, skor minimum, dan skor maksimum. Statistik deskriptif digunakan untuk memberikan gambaran umum mengenai distribusi skor responden pada data simulasi.

Hasil analisis validitas isi, validitas butir, dan reliabilitas selanjutnya diinterpretasikan secara deskriptif dengan membandingkan nilai yang diperoleh terhadap kriteria yang telah ditetapkan. Instrumen dinyatakan layak digunakan apabila seluruh butir memenuhi kriteria validitas serta memiliki nilai Cronbach's Alpha $\geq 0,70$. Karena penelitian ini menggunakan data simulasi, seluruh hasil dimaksudkan sebagai ilustrasi penerapan prosedur analisis instrumen dan bukan sebagai representasi kondisi empiris mahasiswa pada suatu perguruan tinggi.

HASIL DAN PEMBAHASAN

Penelitian ini bertujuan menganalisis kualitas instrumen literasi matematika yang dikembangkan untuk mengukur kemampuan literasi matematika mahasiswa. Analisis difokuskan pada tiga aspek utama, yaitu validitas isi, validitas butir, dan reliabilitas instrumen. Hasil analisis diperoleh berdasarkan data simulasi yang disusun sebagai ilustrasi penerapan prosedur pengujian instrumen dalam penelitian pendidikan matematika.

Validitas Isi Instrumen

Tahap pertama yang dilakukan adalah pengujian validitas isi melalui penilaian tiga orang ahli pendidikan matematika. Penilaian difokuskan pada kesesuaian butir soal terhadap indikator literasi matematika, kejelasan bahasa, serta kesesuaian konteks soal dengan tujuan pengukuran. Hasil analisis menunjukkan bahwa seluruh butir memperoleh nilai Aiken's V di atas batas minimum yang direkomendasikan.

Tabel 4. Hasil Validitas Isi Instrumen Menggunakan Aiken's V

No. Butir	Nilai Aiken's V	Kriteria
1	0,92	Sangat Valid
2	0,88	Sangat Valid
3	0,96	Sangat Valid
4	0,92	Sangat Valid
5	0,83	Sangat Valid
6	0,92	Sangat Valid
7	0,88	Sangat Valid
8	0,96	Sangat Valid
9	0,92	Sangat Valid
10	0,88	Sangat Valid

Berdasarkan Tabel 4, seluruh butir memperoleh nilai Aiken's V antara 0,83 hingga 0,96. Nilai tersebut menunjukkan bahwa setiap butir telah memiliki tingkat kesesuaian yang tinggi dengan indikator literasi matematika yang diukur. Dengan demikian, instrumen telah memenuhi persyaratan validitas isi dan layak untuk diuji secara empiris.

Hasil tersebut sejalan dengan pendapat Aiken (1985) yang menyatakan bahwa semakin tinggi nilai Aiken's V, semakin tinggi pula tingkat kesepakatan para ahli mengenai relevansi suatu butir terhadap konstruk yang diukur. Temuan ini juga memperlihatkan bahwa proses penyusunan kisi-kisi berdasarkan kerangka literasi matematika PISA mampu menghasilkan butir soal yang representatif.

Validitas Butir Instrumen

Setelah memenuhi validitas isi, instrumen dianalisis menggunakan korelasi Corrected Item–Total Correlation. Analisis ini bertujuan mengetahui kemampuan masing-masing butir dalam mengukur konstruk yang sama dengan skor total instrumen.

Tabel 5. Hasil Validitas Butir Instrumen

No. Butir	Corrected Item-Total Correlation	Kriteria
1	0,624	Valid
2	0,587	Valid
3	0,731	Valid
4	0,648	Valid
5	0,541	Valid
6	0,683	Valid
7	0,612	Valid
8	0,756	Valid
9	0,669	Valid
10	0,598	Valid

Hasil pada Tabel 5 menunjukkan bahwa seluruh butir memiliki nilai korelasi di atas 0,30. Hal tersebut mengindikasikan bahwa setiap butir memiliki hubungan yang baik dengan skor total sehingga mampu mengukur konstruk literasi matematika secara konsisten.

Butir kedelapan memiliki nilai korelasi tertinggi (0,756), sedangkan butir kelima memiliki nilai terendah (0,541). Walaupun demikian, seluruh nilai masih berada pada kategori valid sehingga tidak terdapat butir yang perlu dieliminasi.

Temuan ini menunjukkan bahwa seluruh indikator yang dikembangkan mampu merepresentasikan kemampuan literasi matematika mahasiswa. Menurut Crocker dan Algina (2008), suatu butir dinyatakan baik apabila memiliki korelasi positif terhadap skor total karena menunjukkan adanya konsistensi pengukuran antarbutir.

Reliabilitas Instrumen

Selanjutnya dilakukan pengujian reliabilitas menggunakan koefisien Cronbach's Alpha untuk mengetahui tingkat konsistensi internal instrumen.

Tabel 6. Hasil Uji Reliabilitas Instrumen

Statistik	Nilai
Jumlah Butir	10
Jumlah Responden	18
Cronbach's Alpha	0,892
Kategori	Reliabilitas Tinggi

Berdasarkan Tabel 6, nilai Cronbach's Alpha sebesar 0,892 menunjukkan bahwa instrumen

memiliki reliabilitas tinggi. Dengan demikian, instrumen mampu memberikan hasil pengukuran yang konsisten apabila digunakan pada kelompok responden dengan karakteristik yang relatif sama.

Nilai tersebut memenuhi kriteria yang dikemukakan oleh Taber (2018) bahwa instrumen dengan nilai Cronbach's Alpha di atas 0,80 telah memiliki konsistensi internal yang baik dan layak digunakan dalam penelitian pendidikan.

Statistik Deskriptif Skor Literasi Matematika

Selain menganalisis kualitas instrumen, dilakukan pula analisis deskriptif terhadap skor responden untuk memberikan gambaran umum mengenai hasil pengukuran.

Tabel 7. Statistik Deskriptif Skor Literasi Matematika

Statistik	Nilai
Jumlah Mahasiswa	18
Skor Minimum	66
Skor Maksimum	83
Rata-rata	74,50
Simpangan Baku	5,27

Berdasarkan Tabel 7, rata-rata skor literasi matematika mahasiswa adalah 74,50 dengan skor minimum 66 dan maksimum 83. Sebaran skor yang relatif merata menunjukkan bahwa instrumen mampu membedakan tingkat kemampuan mahasiswa tanpa menghasilkan variasi yang terlalu ekstrem.

Pembahasan

Hasil penelitian menunjukkan bahwa instrumen literasi matematika memenuhi seluruh persyaratan kualitas instrumen, yaitu validitas isi yang tinggi, validitas butir yang memadai, dan reliabilitas internal yang tinggi. Temuan tersebut mengindikasikan bahwa proses pengembangan instrumen yang diawali dengan penyusunan kisi-kisi berdasarkan kerangka literasi matematika PISA serta validasi oleh para ahli memberikan kontribusi terhadap meningkatnya kualitas butir soal.

Validitas isi yang tinggi menunjukkan bahwa indikator yang dikembangkan telah sesuai dengan konstruk literasi matematika. Hal ini penting karena literasi matematika tidak hanya mengukur kemampuan menghitung, tetapi juga kemampuan merumuskan masalah, menggunakan konsep matematika, serta menginterpretasikan solusi dalam berbagai konteks kehidupan nyata (OECD, 2023). Oleh sebab itu, penyusunan instrumen harus memperhatikan keseimbangan ketiga aspek tersebut agar pengukuran menjadi lebih komprehensif.

Hasil validitas butir memperlihatkan bahwa seluruh soal mampu membedakan mahasiswa yang memiliki kemampuan literasi matematika tinggi dan rendah. Kondisi ini menunjukkan bahwa setiap butir memberikan kontribusi terhadap pengukuran konstruk secara keseluruhan. Temuan tersebut sejalan dengan penelitian-penelitian pengembangan instrumen yang menyatakan bahwa korelasi butir yang berada di atas 0,30 merupakan indikator bahwa butir layak dipertahankan dalam instrumen.

Reliabilitas instrumen yang mencapai 0,892 menunjukkan tingkat konsistensi internal yang tinggi. Hal ini mengindikasikan bahwa seluruh butir memiliki keterkaitan dalam mengukur konstruk yang sama, yaitu literasi matematika. Menurut Taber (2018), nilai reliabilitas yang tinggi merupakan salah satu indikator penting bahwa instrumen dapat digunakan secara konsisten dalam kegiatan penelitian maupun evaluasi pembelajaran.

Secara keseluruhan, hasil penelitian ini memperkuat temuan OECD (2023) dan NCTM (2023) yang menegaskan bahwa pengukuran literasi matematika memerlukan instrumen yang dikembangkan berdasarkan indikator yang jelas serta melalui proses validasi yang sistematis. Dengan demikian, instrumen yang dihasilkan memiliki potensi untuk digunakan dalam penelitian pendidikan matematika maupun evaluasi pembelajaran setelah melalui pengujian empiris pada jumlah responden yang lebih besar.

SIMPULAN DAN SARAN

Berdasarkan hasil analisis validitas dan reliabilitas instrumen literasi matematika, dapat disimpulkan bahwa instrumen yang dikembangkan telah memenuhi kriteria sebagai instrumen yang berkualitas untuk mengukur kemampuan literasi matematika mahasiswa. Hasil uji validitas isi menunjukkan bahwa seluruh butir soal memiliki nilai Aiken's V pada rentang 0,83–0,96, sehingga setiap butir dinyatakan memiliki tingkat kesesuaian yang sangat tinggi dengan indikator literasi matematika yang diukur. Temuan ini menunjukkan bahwa proses penyusunan instrumen telah sesuai dengan kerangka literasi matematika yang mengacu pada indikator Programme for International Student Assessment (PISA) 2022.

Hasil uji validitas butir menunjukkan bahwa seluruh butir memiliki nilai Corrected Item–Total Correlation lebih besar dari 0,30, sehingga seluruh butir dinyatakan valid secara empiris dan mampu mengukur konstruk literasi matematika secara

konsisten. Tidak terdapat butir yang perlu dieliminasi maupun direvisi berdasarkan hasil analisis validitas empiris.

Selanjutnya, hasil uji reliabilitas menunjukkan bahwa instrumen memperoleh nilai Cronbach's Alpha sebesar 0,892, yang termasuk dalam kategori reliabilitas tinggi. Nilai tersebut mengindikasikan bahwa instrumen memiliki konsistensi internal yang sangat baik sehingga mampu memberikan hasil pengukuran yang stabil apabila digunakan pada kelompok responden dengan karakteristik yang sejenis.

Secara keseluruhan, hasil penelitian menunjukkan bahwa instrumen literasi matematika yang dikembangkan telah memenuhi aspek validitas isi, validitas empiris, dan reliabilitas internal. Oleh karena itu, instrumen ini layak digunakan sebagai contoh metodologis dalam pengembangan instrumen pendidikan matematika. Namun demikian, karena penelitian ini menggunakan data simulasi, hasil yang diperoleh belum dapat digeneralisasikan sebagai representasi kondisi empiris mahasiswa dan masih memerlukan pengujian lebih lanjut menggunakan data lapangan.

Berdasarkan hasil penelitian, terdapat beberapa saran yang dapat dipertimbangkan untuk penelitian selanjutnya.

Penelitian berikutnya disarankan menggunakan data empiris yang diperoleh melalui pengumpulan data secara langsung pada mahasiswa sehingga hasil analisis validitas dan reliabilitas dapat menggambarkan kondisi nyata serta memiliki daya generalisasi yang lebih tinggi. Jumlah responden perlu diperluas dengan melibatkan mahasiswa dari berbagai program studi maupun perguruan tinggi agar karakteristik instrumen dapat diuji pada populasi yang lebih beragam serta menghasilkan kualitas instrumen yang lebih representatif. Analisis instrumen tidak hanya menggunakan pendekatan Classical Test Theory (CTT), tetapi juga dapat dikembangkan menggunakan Model Rasch atau Item Response Theory (IRT) untuk memperoleh informasi yang lebih mendalam mengenai karakteristik setiap butir soal, tingkat kesukaran, kemampuan pembeda, serta kesesuaian model pengukuran. Penelitian selanjutnya juga dapat mengembangkan instrumen literasi matematika yang mengintegrasikan konteks etnomatematika, teknologi digital, atau Asesmen Kompetensi Minimum (AKM) sehingga instrumen yang dihasilkan lebih kontekstual, relevan dengan perkembangan kurikulum, dan sesuai dengan kebutuhan pembelajaran matematika abad ke-21.

Selain menguji validitas dan reliabilitas, penelitian lanjutan dapat melakukan analisis terhadap validitas konstruk menggunakan Exploratory Factor Analysis (EFA) atau Confirmatory Factor Analysis (CFA) untuk memperoleh bukti yang lebih kuat mengenai struktur konstruk literasi matematika yang diukur.

Dengan adanya pengembangan dan pengujian instrumen yang lebih komprehensif, diharapkan tersedia instrumen literasi matematika yang memiliki kualitas psikometrik tinggi sehingga dapat digunakan secara luas dalam penelitian maupun evaluasi pembelajaran matematika di perguruan tinggi.

DAFTAR RUJUKAN

- 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>
- Allen, M. J., & Yen, W. M. (2002). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.
- Azwar, S. (2021). *Reliabilitas dan Validitas* (Edisi ke-5). Yogyakarta: Pustaka Pelajar.
- Crocker, L., & Algina, J. (2008). *Introduction to Classical and Modern Test Theory*. Mason, OH: Cengage Learning.
- Hair, J. F., Black, W. C., Babin, B. J., & Anderson, R. E. (2022). *Multivariate Data Analysis* (9th ed.). Boston: Cengage Learning.
- National Council of Teachers of Mathematics. (2023). *Catalyzing Change in Mathematics Education*. Reston, VA: National Council of Teachers of Mathematics.
- OECD. (2023). *PISA 2022 Assessment and Analytical Framework*. Paris: OECD Publishing. <https://doi.org/10.1787/dfe0bf9c-en>
- OECD. (2024). *PISA 2022 Results (Volume I): The State of Learning and Equity in Education*. Paris: OECD Publishing.
- Sugiyono. (2023). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (Edisi ke-3). Bandung: Alfabeta.
- Taber, K. S. (2018). The use of Cronbach's alpha when developing and reporting research instruments in science education. *Research in Science Education*, 48(6), 1273–1296. <https://doi.org/10.1007/s11165-016-9602-2>
- Yusup, F. (2018). Uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Jurnal*

- Tarbiyah: Jurnal Ilmiah Kependidikan*, 7(1), 17–23.
- Widhiarso, W. (2019). *Analisis Butir Soal dengan Pendekatan Teori Tes Klasik*. Yogyakarta: Fakultas Psikologi Universitas Gadjah Mada.
- Retnawati, H. (2016). *Analisis Kuantitatif Instrumen Penelitian*. Yogyakarta: Parama Publishing.
- Retnawati, H., Kartowagiran, B., Hadi, S., & Jaedun, A. (2021). A psychometric analysis of mathematics assessment using the Rasch model. *International Journal of Instruction*, 14(2), 663–678.
- DeMars, C. (2023). *Item Response Theory* (2nd ed.). New York: Oxford University Press.
- Boone, W. J., Staver, J. R., & Yale, M. S. (2014). *Rasch Analysis in the Human Sciences*. Dordrecht: Springer.
- Bond, T. G., Yan, Z., & Heene, M. (2021). *Applying the Rasch Model: Fundamental Measurement in the Human Sciences* (4th ed.). New York: Routledge.
- Niss, M., & Jablonka, E. (2020). *Mathematical Literacy*. In S. Lerman (Ed.), *Encyclopedia of Mathematics Education* (2nd ed.). Cham: Springer.
- Zulkardi. (2020). Pengembangan asesmen literasi matematika berbasis PISA di Indonesia. *Jurnal Pendidikan Matematika*, 14(2), 123–138.
- Wijaya, A. (2021). *Pendidikan Matematika Realistik: Suatu Alternatif Pendekatan Pembelajaran Matematika*. Yogyakarta: Graha Ilmu.