

## ANALISIS BUTIR SOAL TES LITERASI NUMERASI DOMAIN GEOMETRI SISWA SMP MENGGUNAKAN RASCH MODEL BERBANTUAN MINISTEP

Erene Nazharita<sup>1\*</sup>, Rifqi Hidayat<sup>2</sup>, Jajang Rahmatudin<sup>3</sup>  
Pendidikan Matematika<sup>1,2,3</sup>, Fakultas Keguruan Dan Ilmu Pendidikan<sup>1,2,3</sup>,  
Universitas Muhammadiyah Cirebon<sup>1,2,3</sup>  
erenenhara@gmail.com<sup>1</sup>, rifqi.hidayat@umc.ac.id<sup>2</sup>, jajang@umc.ac.id<sup>3</sup>

### Abstrak

Ketersediaan instrumen tes literasi numerasi yang berkualitas, khususnya pada domain geometri untuk siswa SMP, masih terbatas. Kondisi ini menjadi alasan utama dilakukannya penelitian ini. Fokus utama penelitian adalah menggambarkan mutu instrumen tes literasi numerasi pada materi geometri yang dirancang untuk siswa SMP. Penelitian menggunakan pendekatan kuantitatif deskriptif, dengan analisis mengacu pada pemodelan Rasch. Penelitian dilakukan dengan tiga tahap. Tahap pertama adalah penyusunan butir soal berdasarkan indikator yang telah ditentukan. Selanjutnya dilakukan uji coba terbatas pada siswa kelas IX. Tahap terakhir adalah menganalisis kualitas butir soal tes menggunakan model Rasch dengan bantuan perangkat lunak Ministep. Instrumen yang disusun terdiri dari enam soal uraian. Data penelitian diperoleh dari 36 siswa kelas IX yang dipilih menggunakan teknik purposive sampling. Hasil analisis menunjukkan bahwa 5 dari 6 butir soal dinyatakan valid. Reliabilitas instrumen berdasarkan nilai *Cronbach's Alpha* sebesar 0,25 termasuk kategori lemah. Tingkat kesukaran butir soal bervariasi, terdiri dari 1 butir sangat mudah, 2 butir mudah, 2 butir sukar, dan 1 butir sangat sukar. Secara umum, instrumen telah memenuhi sebagian besar kriteria Model Rasch, namun masih memerlukan revisi pada butir soal yang tidak valid dan penyesuaian tingkat kesukaran untuk meningkatkan kualitas instrumen. Sehingga, dapat digunakan secara efektif dalam evaluasi kemampuan literasi numerasi dengan domain geometri siswa SMP.

*Kata Kunci: Instrumen Tes, Literasi Numerasi, Geometri, Model Rasch, Ministep*

---

### A. Pendahuluan

Literasi numerasi merupakan salah satu kompetensi esensial yang diukur dalam Asesmen Kompetensi Minimum (AKM) dan menjadi fokus penting dalam pembelajaran matematika abad ke-21. Literasi numerasi adalah kemampuan menggunakan konsep bilangan serta keterampilan berhitung dalam berbagai situasi kehidupan sehari-hari (Latifah & Rahmawati, 2022). Kemampuan ini tidak hanya mencakup keterampilan menghitung, tetapi juga kemampuan menginterpretasikan

informasi kuantitatif dalam berbagai bentuk representasi, termasuk grafik, tabel, diagram, dan bangun geometri. Ketersediaan instrumen tes literasi numerasi yang berkualitas, khususnya pada domain geometri untuk jenjang SMP, masih terbatas (Mulyani & Huriaty, 2016). Padahal, geometri memiliki peran penting karena mampu menumbuhkan apresiasi terhadap lingkungan dan alam semesta, sekaligus mengasah keterampilan pemecahan masalah melalui penalaran spasial (Budiarto & Artiono, 2019). Salah satu topik dalam domain geometri adalah Teorema Pythagoras. Materi ini menjadi landasan penting dalam berbagai perhitungan matematika, baik dalam kaitannya dengan konsep lain maupun penerapannya pada penyelesaian masalah kehidupan sehari-hari (Wulandari & Riajanto, 2020).

Upaya pengembangan instrumen tes numerasi pada materi geometri telah dilakukan oleh beberapa peneliti di Indonesia. Tenriawaru & Sumarni (2018) mengembangkan alat penilaian autentik geometri untuk siswa kelas VIII SMP, yang terbukti sahih dan andal dalam mengukur capaian pembelajaran. Sementara itu, Ijtahidah dan Nisa' (2023) berhasil merancang soal literasi numerasi SMP terintegrasi nilai-nilai Islam yang mendapatkan validitas ahli dan tingkat reliabilitas tinggi, menunjukkan bahwa pengembangan instrumen dapat disesuaikan dengan konteks budaya dan karakter siswa. Pendekatan berbasis konteks lokal juga menjadi perhatian dalam pengembangan instrumen literasi numerasi. Aprilia et al., (2023) menyusun instrumen numerasi AKM yang memuat nilai budaya lokal, sehingga lebih relevan dengan pengalaman siswa sehari-hari. Relevansi konteks ini penting untuk meningkatkan keterlibatan siswa dan validitas butir soal. Selain itu, Dewi & Defitriani (2024) mengembangkan instrumen literasi numerasi berbasis etnomatematika yang menggabungkan unsur budaya dalam soal geometri, sehingga mendorong keterhubungan antara konsep matematika dan kehidupan nyata.

Penelitian terkini juga menguatkan pentingnya pengembangan instrumen numerasi pada domain geometri. Rahmawati & Rahayu (2024) menguji validitas instrumen penilaian berbasis literasi numerasi pada materi bangun ruang prisma dan limas kelas VIII SMP, yang menghasilkan butir soal dengan kategori valid tinggi. Fakhriyani et al., (2025) menemukan variasi kemampuan numerasi siswa SMP pada konten geometri dan pengukuran jika ditinjau dari gaya kognitif, yang menunjukkan adanya kebutuhan instrumen yang mampu mengukur kemampuan

secara lebih akurat. Bahkan pada materi aljabar, Nur Hikmah et al., (2024) mendapati banyak siswa SMP mengalami kesulitan menyelesaikan soal cerita literasi numerasi, sehingga semakin menegaskan perlunya pengembangan instrumen yang tepat dan terukur di berbagai topik matematika.

Pendekatan model *Item Response Theory* (IRT) semakin banyak digunakan dalam pengembangan instrumen tes karena mampu memisahkan parameter kemampuan peserta dan karakteristik butir secara independen. Keunggulan ini menjadikan hasil pengukuran lebih akurat dan tidak bergantung pada kelompok sampel tertentu (Wulandari et al., 2025). Model Rasch termasuk salah satu model IRT yang paling banyak digunakan. Prinsip dasarnya adalah memodelkan probabilitas peserta menjawab benar suatu butir berdasarkan perbedaan antara kemampuan peserta dan tingkat kesulitan butir soal. Estimasi parameter dalam skala *logit* yang dihasilkan bersifat objektif dan memungkinkan perbandingan lintas kelompok peserta tes (Nurfitriandini et al., 2024).

Perangkat lunak Winsteps dan Ministep umum digunakan untuk mempermudah analisis Model Rasch. Efektivitas penerapan Model Rasch berbantuan Ministep terbukti pada beberapa penelitian. Wulandari et al., (2025), menunjukkan bahwa analisis Rasch mampu mengidentifikasi distribusi tingkat kesulitan dari kategori sangat mudah hingga sangat sulit, serta memetakan kemampuan siswa secara lebih rinci dibanding teori tes klasik. Ramadhan & Hidayatullah (2023), menegaskan bahwa Ministep dapat membantu guru menemukan butir yang memerlukan revisi, sehingga instrumen menjadi lebih valid dan reliabel

Ketersediaan instrumen literasi numerasi pada materi geometri di tingkat SMP yang memenuhi persyaratan validitas, reliabilitas dan tingkat kesukaran masih menjadi tantangan. Sejumlah penelitian sebelumnya memang telah menghasilkan instrumen serupa, namun kajian yang menelaah secara rinci menggunakan Model Rasch dengan bantuan perangkat Ministep pada konteks literasi numerasi geometri SMP masih terbatas. Penelitian ini bertujuan untuk mendeskripsikan mutu instrumen tes literasi numerasi pada doian geometri yang dirancang untuk siswa SMP. Analisis mutu dilakukan berdasarkan kriteria validitas, reliabilitas, dan tingkat kesukaran dengan menggunakan Model Rasch berbantuan perangkat lunak Ministep.

## B. Metode Penelitian

Penelitian ini menggunakan pendekatan kuantitatif deskriptif yang bertujuan menggambarkan kualitas instrumen tes literasi numerasi pada domain geometri untuk siswa SMP. Analisis butir soal dilakukan menggunakan Model Rasch dengan bantuan perangkat lunak Ministep. Penelitian dilaksanakan dalam tiga tahap. Tahap pertama adalah penyusunan instrumen tes berdasarkan indikator literasi numerasi yang mengacu pada panduan Asesmen Kompetensi Minimum (AKM) dan Kurikulum Merdeka. Indikator soal disusun dengan mempertimbangkan kompetensi dasar dan capaian pembelajaran Kurikulum Merdeka untuk mata pelajaran matematika SMP, khususnya domain geometri pada topik Teorema Pythagoras. Instrumen yang dikembangkan berjumlah enam butir soal uraian untuk mengukur kemampuan literasi numerasi siswa SMP pada materi geometri. Tahap kedua adalah uji coba instrumen pada kelompok terbatas, yaitu siswa kelas IX A SMPN 1 Plumbon pada tanggal 29 Juli 2025 yang dipilih menggunakan teknik *purposive sampling*. Teknik ini dipilih karena subjek penelitian harus memiliki karakteristik tertentu, yakni telah mempelajari materi geometri termasuk Teorema Pythagoras. Dengan demikian, dapat memberikan respons yang relevan terhadap instrumen yang diuji. Tahap ketiga adalah analisis kualitas butir soal yang mencakup uji validitas, reliabilitas dan tingkat kesukaran menggunakan pemodelan Rasch. Subjek penelitian terdiri dari 36 siswa kelas IX SMP yang telah mengikuti tes menggunakan instrumen yang dikembangkan. Uji coba dilaksanakan secara tatap muka di ruang kelas dengan durasi 60 menit, di bawah pengawasan langsung peneliti. Data hasil tes dianalisis menggunakan perangkat lunak Ministep untuk memperoleh informasi mengenai kelayakan setiap butir soal. Analisis dilakukan berdasarkan tiga aspek berikut:

### Validitas Butir Soal

Validitas butir soal dianalisis menggunakan tiga indikator pada Model Rasch, yaitu *outfit mean square* (MNSQ), *outfit z-standard* (ZSTD), dan *point measure correlation* (PT Measure Corr). Kriteria kelayakan butir mengacu pada Sumintono dan Widhiarso (2015), yaitu:

1. Nilai MNSQ berada pada rentang  $0,5 < \text{MNSQ} < 1,5$
2. Nilai ZSTD berada pada rentang  $-2,0 < \text{ZSTD} < +2,0$

3. Nilai *PT Measure Corr* berada pada rentang  $0,4 < PT Measure Corr < 0,85$   
 Reliabilitas Instrumen

Reliabilitas instrumen diuji menggunakan *Cronbach's Alpha* melalui perangkat lunak Ministep. Interpretasi nilai  $\alpha$  mengacu pada Ernanda et al., (2024), seperti pada tabel berikut:

**Tabel 1.** Kriteria Nilai *Cronbach's Alpha*

Nilai <i>Cronbach's Alpha</i>	Kriteria
$0,8 \leq \alpha$	Istimewa
$0,7 \alpha < 0,8$	Bagus Sekali
$0,6 \leq \alpha < 0,7$	Bagus
$0,5 \leq \alpha < 0,6$	Cukup
$\alpha < 0,5$	Lemah

Tingkat Kesukaran Butir Soal

Tingkat kesukaran butir soal ditentukan berdasarkan nilai *measure* (logit) pada *Wright Map* yang dihasilkan oleh analisis Rasch. Kriteria kesukaran mengacu pada Sumintono dan Widhiarso (2015), seperti pada tabel berikut:

**Tabel 2.** Kriteria Nilai Tingkat Kesukaran

Nilai <i>Meaasure (Logit)</i>	Kriteria
$Measure Logit < -SD logit$	Sangat Mudah
$-SD logit = Measure logit = 0$	Mudah
$0 = Measure logit = SD logit$	Sukar
$Measure logit > SD logit$	Sangat Sukar

Dalam Model Rasch, nilai logit yang tinggi menunjukkan tingkat kesulitan butir yang besar. Tingkat kesukaran setiap butir diwakili oleh nilai *measure*, yaitu hasil konversi data ordinal berupa skor jawaban menjadi satuan *logit* yang memiliki jarak setara (*equal interval*).

### C. Hasil Dan Pembahasan

Penelitian ini bertujuan untuk menggambarkan kualitas instrumen tes literasi numerasi pada domain geometri bagi siswa SMP, khususnya pada topik Teorema Pythagoras. Instrumen yang digunakan berupa enam butir soal uraian yang disusun berdasarkan indikator literasi numerasi yang mengacu pada panduan Asesmen Kompetensi Minimum (AKM) dan Kurikulum Merdeka. Subjek penelitian terdiri dari 36 siswa kelas IX A SMPN 1 Plumbon, yang mengikuti uji coba instrumen

pada tanggal 29 Juli 2025 secara tatap muka di ruang kelas dengan durasi 60 menit di bawah pengawasan peneliti. Analisis kualitas butir soal dilakukan menggunakan Model Rasch dengan bantuan perangkat lunak Ministep untuk menilai validitas, reliabilitas, dan tingkat kesukaran setiap butir soal.

## Hasil

### Validitas Butir Soal

Analisis butir soal dilakukan menggunakan perangkat lunak Ministep dengan mengacu pada tiga kriteria:

1.  $0,5 < \text{MNSQ} < 1,5$
2.  $-2,0 < \text{ZSTD} < +2,0$
3.  $0,4 < \text{PT Measure Corr} < 0,85$

**Analysis Validity**

TABLE 13.1 C:\Users\azus\Downloads\DATA NILAI.pr EQ0133WS.TXT Jul 31 2025 11:48  
 INPUT: 36 Person 6 Item REPORTED: 36 Person 6 Item 5 CAIS MINISTEP-4.8.2.0  
 Person: REAL SEP.: .45 REL.: .17 ... Item: REAL SEP.: 3.33 REL.: .92

Item STATISTICS: MEASURE ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT (MNSQ)	OUTFIT (ZSTD)	PTMEASURE-AL (EXACT MATCH) CORR.	EXP.	OBS%	EXP%	Item		
6	42	36	1.2E	.24	1.46	1.19	1.08	.33	.64	.39	66.7	70.0	R6
4	70	36	.24	.16	.73	-1.24	1.62	-1.45	.70	.53	47.2	40.5	R4
2	75	36	.12	.15	2.19	4.34	2.58	4.28	-.03	.53	22.2	39.5	R2
3	89	36	-.18	.14	.95	-.22	.85	-.57	.53	.50	27.8	26.1	R3
5	108	36	-.58	.15	.60	-2.60	.61	-1.26	.52	.43	44.4	27.3	R5
1	120	36	-.86	.16	.55	-2.34	.47	-1.33	.56	.36	41.7	40.2	R1
MEAN	84.0	36.0	.00	.17	1.08	-.11	1.04	.01			41.7	40.6	
P.S.D	25.6	.0	.68	.03	.58	2.4	.72	2.0			14.3	14.4	

$0,5 < \text{MNSQ} < 1,5$   
 $-2,0 < \text{ZSTD} < +2,0$   
 $0,4 < \text{Pt Measure Corr} < 0,85$

**Description**  
 Valid/retained  
 Valid/retained  
 Invalid/Revised  
 Valid/retained  
 Valid/retained  
 Valid/retained

**Gambar 1.** Output analisis validitas butir soal

Berdasarkan output tersebut, ringkasan hasil analisis validitas disajikan pada Tabel berikut:

**Tabel 3.** Rekap Hasil Analisis Validitas Butir Soal

Kode Soal	MNSQ	ZSTD	PT Measure Corr	Kriteria
R1	0,47	-1,33	0,56	Valid/retained
R2	2,58	4,28	-0,03	Invalid/revised
R3	0,85	-0,57	0,53	Valid/retained
R4	0,62	-1,45	0,70	Valid/retained
R5	0,61	-1,26	0,52	Valid/retained
R6	1,08	0,33	0,64	Valid/retained

Reliabilitas Instrumen

Reliabilitas instrumen diuji menggunakan *Cronbach's Alpha* pada perangkat lunak Ministep.

**Item Reliability Analysis**

TABLE 3.1 C:\Users\asus\Downloads\DATA NILAI.prn ZOU133WS.TXT Jul 31 2025 11:48  
 INPUT: 36 Person 6 Item REPORTED: 36 Person 6 Item 5 CATS MINISTEP 4.8.2.0

SUMMARY OF 36 MEASURED Person

	TOTAL SCORE	COUNT	MEASURE	MODEL S.E.	MNSQ	ZSTD	MNSQ	ZSTD
MEAN	14.0	6.0	.78	.40	1.09	.07	1.04	.07
SEM	.6	.0	.09	.01	.11	.18	.12	.15
P.SD	3.3	.0	.51	.04	.66	1.07	.68	.88
S.SD	3.3	.0	.52	.04	.67	1.09	.69	.89
MAX.	21.0	6.0	1.98	.52	2.97	1.99	2.55	1.54
MIN.	9.0	6.0	-.01	.37	.10	-2.52	.11	-1.79

REAL RMSE	.47	TRUE SD	.21	SEPARATION	.45	Person RELIABILITY	.17
MODEL RMSE	.40	TRUE SD	.32	SEPARATION	.80	Person RELIABILITY	.39
S.E. OF Person MEAN = .09							

Person RAW SCORE-TO-MEASURE CORRELATION = 1.00  
**CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .25 SEM = 2.83**  
 STANDARDIZED (50 ITEM) RELIABILITY = .84

Lemah

**Gambar 2.** Output uji reliabilitas instrumen

Berdasarkan *output* tersebut, Nilai *Cronbach's Alpha (KR-20)* yang diperoleh adalah 0,25. Berdasarkan kriteria Ernanda et al., (2024), nilai ini termasuk kategori lemah ( $\alpha < 0,5$ ), yang menunjukkan reliabilitas antarbutir soal masih rendah.

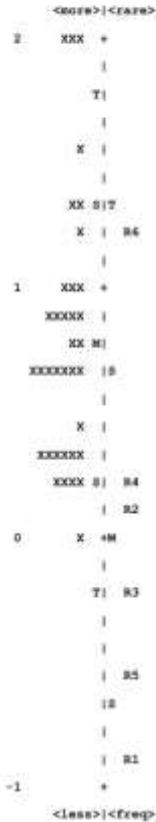
Tingkat Kesukaran Butir Soal

Tingkat kesukaran dianalisis berdasarkan nilai *measure (logit)* pada *Wright Map* yang dihasilkan oleh Ministep.

**Indeks Kesukaran**

TABLE 12.2 C:\Users\asus\Downloads\DATA NILAI.pr EOU133WS.TXT Jul 31 2025 11:48  
 INPUT: 36 Person 6 Item REPORTED: 36 Person 6 Item 5 CATS MINISTEP 4.8.2.0

MEASURE Person - MAP - Item



**Gambar 3. Output Item Wright Map**

TABLE 10.1 C:\Users\asus\Downloads\DATA NILAI.pr EOU536WS.TXT Aug 10 2025 12:57  
 INPUT: 36 Person 6 Item REPORTED: 36 Person 6 Item 5 CATS MINISTEP 4.8.2.0

Person: REAL SEP.: .45 REL.: .17 ... Item: REAL SEP.: 3.33 REL.: .92

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT  MNSQ	OUTFIT  MNSQ	PTMEASUR-AL  CORR.	EXACT MATCH  EXP.	EXACT MATCH  OBS%	EXACT MATCH  EXP%	Item		
2	75	36	.12	.15	2.19	4.34	2.58	4.28	A-.03	.53	22.2	39.5	R2
6	42	36	1.28	.24	1.46	1.19	1.08	.33	B-.64	.39	66.7	70.0	R6
3	89	36	-.18	.14	.95	-.22	.85	-.57	C-.53	.50	27.8	26.1	R3
4	70	36	.24	.16	.73	-1.24	.62	-1.45	C-.70	.53	47.2	40.5	R4
5	108	36	-.58	.15	.60	-2.60	.61	-1.26	B-.52	.43	44.4	27.3	R5
1	120	36	-.86	.16	.55	-2.34	.47	-1.33	A-.56	.36	41.7	40.2	R1
MEAN	84.0	36.0	.00	.17	1.08	-.1	1.04	.0			41.7	40.6	
P.S.D	25.6	.0	.68	.03	.58	2.4	.72	2.0			14.3	14.4	

**Gambar 4. Output analisis data tingkat kesukaran**

Berdasarkan nilai *measure logit* pada *output* tersebut, ringkasan hasil tingkat kesukaran disajikan pada tabel berikut:

**Tabel 4.** Rekap Hasil Analisis Tingkat Kesukaran Butir Soal

Kode Soal	<i>Measure Logit</i>	<i>SD Logit</i>	Kriteria
R1	-0,86	0,68	Sangat Mudah
R2	0,12	0,68	Sukar
R3	-0,18	0,68	Mudah
R4	0,24	0,68	Sukar
R5	-0,58	0,68	Mudah
R6	1,26	0,68	Sangat Sukar

#### **Pembahasan**

Analisis validitas butir soal menggunakan Model Rasch pada perangkat lunak Ministep menunjukkan bahwa dari enam butir soal yang diuji, lima di antaranya memenuhi kriteria kelayakan validitas, sedangkan satu butir soal dinyatakan tidak valid. Butir R2 memiliki nilai *outfit* MNSQ sebesar 2,58 (melebihi batas atas 1,5), nilai ZSTD sebesar 4,28 (di atas +2,0), dan nilai PT *Measure Corr* negatif (-0,03). Kondisi ini mengindikasikan bahwa respons peserta terhadap butir R2 menyimpang dari prediksi model Rasch, sehingga butir ini perlu direvisi atau dihilangkan. Butir soal lain (R1, R3, R4, R5, R6) memiliki nilai MNSQ dalam rentang yang direkomendasikan (0,5–1,5), nilai ZSTD antara -2,0 hingga +2,0, dan PT *Measure Corr* pada kisaran 0,4–0,85 sehingga dinyatakan valid dan dapat dipertahankan.

Hasil uji reliabilitas menggunakan *Cronbach's Alpha* (KR-20) pada perangkat lunak Ministep menunjukkan nilai sebesar 0,25. Berdasarkan kriteria Ernanda et al., (2024), nilai ini berada pada kategori lemah ( $\alpha < 0,5$ ), yang berarti konsistensi internal antarbutir soal masih rendah. Rendahnya reliabilitas dapat disebabkan oleh beberapa faktor, antara lain jumlah butir soal yang relatif sedikit (hanya enam butir), keberadaan butir yang tidak valid (R2) yang berpotensi menurunkan konsistensi pengukuran, serta heterogenitas kemampuan responden yang kurang terdistribusi merata.

Hasil analisis tingkat kesukaran berdasarkan nilai *measure logit* menunjukkan bahwa dari enam butir soal yang dianalisis, terdapat satu butir soal (R1) dengan kategori sangat mudah, dua butir soal (R3 dan R5) dengan kategori mudah, dua butir soal (R2 dan R4) dengan kategori sukar, serta satu butir soal (R6)

dengan kategori sangat sukar. Penyebaran tingkat kesukaran yang bervariasi ini menunjukkan bahwa instrumen tes mampu menjangkau kemampuan peserta didik yang berbeda-beda. Meskipun demikian, proporsi butir dengan kategori sangat mudah dan sangat sukar perlu dikendalikan agar tidak terlalu mendominasi, karena keberadaan butir-butir tersebut yang terlalu banyak dapat mengurangi efektivitas tes dalam membedakan kemampuan peserta didik pada level menengah.

Secara umum, instrumen tes literasi numerasi geometri yang dikembangkan telah memenuhi sebagian besar kriteria kelayakan menurut Model Rasch. Namun, perlu adanya revisi pada butir yang tidak valid (R2) serta peningkatan reliabilitas instrumen. Penambahan jumlah butir dan pemerataan tingkat kesukaran disarankan agar instrumen lebih representatif dan mampu mengukur kemampuan siswa secara akurat. Instrumen ini berpotensi menjadi alat evaluasi yang efektif dalam mengukur kemampuan literasi numerasi geometri siswa SMP, khususnya jika telah melalui tahap revisi dan pengujian lanjutan.

#### **D. Kesimpulan**

Penelitian ini berhasil mengembangkan instrumen tes literasi numerasi pada domain geometri untuk siswa SMP dengan menggunakan Model Rasch berbantuan perangkat lunak Ministep. Hasil analisis menunjukkan bahwa dari enam butir soal yang disusun, lima butir memenuhi kriteria validitas sedangkan satu butir (R2) dinyatakan tidak valid dan memerlukan revisi. Reliabilitas instrumen berdasarkan nilai *Cronbach's Alpha* sebesar 0,25 termasuk kategori lemah, yang mengindikasikan konsistensi internal antarbutir soal masih rendah. Tingkat kesukaran butir soal bervariasi, terdiri atas satu butir sangat mudah, dua butir mudah, dua butir sukar, dan satu butir sangat sukar. Variasi ini menunjukkan bahwa instrumen mampu menjangkau rentang kemampuan siswa yang beragam, meskipun proporsinya perlu disesuaikan untuk efektivitas pengukuran. Secara umum, instrumen telah memenuhi sebagian besar kriteria Model Rasch, namun masih memerlukan revisi pada butir tidak valid dan penyesuaian tingkat kesukaran untuk meningkatkan kualitas instrumen. Sehingga, dapat digunakan secara efektif dalam evaluasi kemampuan literasi numerasi dengan domain geometri siswa SMP.

## Daftar Pustaka

- Aprilia, N., Setiani, Y., & Hadi FS, C. A. (2023). Pengembangan Instrumen Tes Numerasi Pada Asesmen Kompetensi Minimum Yang Bernilai Budaya Lokal. *Jurnal Educatio FKIP UNMA*, 9(2), 850–857. <https://doi.org/10.31949/educatio.v9i2.4824>
- Budiarto, M. T., & Artiono, R. (2019). Geometri Dan Permasalahan Dalam Pembelajarannya (Suatu Penelitian Meta Analisis). *JUMADIKA : Jurnal Magister Pendidikan Matematika*, 1(1), 9–18. <https://doi.org/10.30598/jumadikavol1iss1year2019page9-18>
- Dewi, S., & Defitriani, E. (2024). Pengembangan Instrumen Tes Literasi Numerasi Berbasis Etnomatematika untuk Siswa SMP Negeri 2 Jambi. *Jurnal Ilmiah Universitas Batanghari Jambi*, 24(2), 1284. <https://doi.org/10.33087/jiubj.v24i2.5231>
- Ernanda, M. R. D., Rizqiah, S., Binta, S. A., Oktarisa, Y., & Sukma Suryana, T. G. (2024). Pengembangan instrumen tes diagnostik tingkat konsep pada materi dinamika. *Jurnal Nebula*, 1(2), 28–30. <https://jurnal.untirta.ac.id/index.php/nebula/>
- Fakhriyani, L., Subarinah, S., Novitasari, D., & Sridana, N. (2025). Analisis Kemampuan Numerasi Siswa SMP Pada Konten Geometri dan Pengukuran Ditinjau Dari Gaya Kognitif. *Journal of Classroom Action Research*, 8(1). <http://jppipa.unram.ac.id/index.php/jcar/index>
- Ijtahidah, F. F., & Nisa', R. (2023). Pengembangan soal literasi numerasi SMP yang terintegrasi Islam. *Primatika : Jurnal Pendidikan Matematika*, 12(1), 1–8. <https://doi.org/10.30872/primatika.v12i1.1596>
- Latifah, L., & Rahmawati, F. P. (2022). Penerapan Program CALISTUNG untuk Meningkatkan Literasi Numerasi Siswa Kelas Rendah di Sekolah Dasar. *Jurnal Basicedu*, 6(3), 5021–5029. <https://doi.org/10.31004/basicedu.v6i3.3003>
- Muliyani, T., & Huriaty, D. (2016). *Pengukuran Pada Jenjang Smp*. 2(2), 91–98.
- Nurfitriandini, S., Zulkarnaen Ilham, M., Sahara, F., Aryani, L., Oktarisa, Y., & Sultan Ageng Tirtayasa, U. (2024). *8 Nebula 1(1) (2024) Pengembangan Instrumen Tes Diagnostik Tingkat Konsepsi Pada Materi Kinematika*. 1(1), 18–23. <https://jurnal.untirta.ac.id/index.php/nebula/>
- Nur Hikmah, A. A., Budiman, I., & Kartika, H. (2024). Analisis Kemampuan Literasi Numerasi Siswa SMP dalam Menyelesaikan Soal Cerita pada Persamaan Linier Satu Variabel. *Didactical Mathematics*, 6(2), 314–322. <https://doi.org/10.31949/dm.v6i2.10672>
- Rahmawati, S., & Rahayu, P. (2024). *Validitas Instrumen Penilaian Berbasis Literasi Numerasi Pada Materi Bangun Ruang Prisma Dan Limas Kelas 8 Smp*. 9(1), 1255–1261.

- Ramadhan, A. F., & Hidayatullah, R. S. (2023). Analisis Kualitas Butir Soal Ujian Satuan Pendidikan (Usp) Materi C2 Teknik Pemesinan Kelas Xii Di Smk PGRI 1 Lamongan Melalui Model Rasch. *Jptm*, 12(03), 1–10.
- Sumintono, B., & Widhiarso, B. (2015). Aplikasi Pemodelan Rasch Pada Assessment Pendidikan. Cimahi: Trim Komunikata
- Tenriawaru, A., & Sumarni. (2018). Pengembangan alat penilaian autentik dalam pembelajaran geometri di kelas viii smp negeri sungguminasa kabupaten gowa. *Jurnal Ilmiah Pendidikan Matematika Equals*, 1(2), 1–9.
- Wulandari, L., & Riajanto, M. L. E. J. (2020). Analisis Kesulitan Siswa SMP dalam Menyelesaikan Soal Materi Teorema Pythagoras. *Jurnal Riset Pendidikan Dan Inovasi Pembelajaran Matematika (JRPIPM)*, 3(2), 61. <https://doi.org/10.26740/jrpijm.v3n2.p61-67>
- Wulandari, O., Muhtarom, M., & Sumarno, S. (2025). Analisis butir soal pengetahuan matematika kelas V sekolah dasar menggunakan model Rasch. *Jurnal Pengembangan dan Penelitian Pendidikan*, 7(1), 1-10. <https://journalversa.com/s/index.php/jppp>