

Seleksi dan Validasi Butir Soal Literasi Matematis Materi Perbandingan Menggunakan Analisis Karakteristik Psikometrik

Indah Wahyuni^{1*}, Junarti², Dian Ratna Puspananda³

^{1,2,3}Pendidikan Matematika, IKIP PGRI Bojonegoro, Jalan Panglima Polim No. 46

*Korespondensi Penulis. E-mail: iw4489925@gmail.com, Telp: +6285718884230

Abstrak

Penelitian ini bertujuan untuk menyeleksi dan memvalidasi draf butir soal tes literasi matematis materi perbandingan pada siswa kelas VII berdasarkan analisis karakteristik psikometrik. Penelitian deskriptif kuantitatif ini melibatkan 36 siswa kelas VII A SMP Negeri 1 Soko yang dipilih melalui teknik *cluster random sampling*. Instrumen yang diujicobakan berupa 10 draf butir soal uraian pada tes awal (pretest) dan tes akhir (posttest). Analisis data psikometrik meliputi uji validitas empiris (*Pearson Product Moment*), tingkat kesukaran, daya pembeda, dan uji reliabilitas (*Cronbach's Alpha*). Hasil penelitian menunjukkan bahwa pada uji pretest, terdapat 6 soal valid, namun hanya 5 soal (Soal 1, 2, 3, 5, dan 8) yang terpilih karena memiliki daya pembeda minimal "Cukup" dengan $DP \geq 0,305$ dan sebaran kesukaran yang proporsional (1 mudah, 2 sedang, 2 sukar) dengan koefisien reliabilitas $r_{11} = 0,80$ (kategori tinggi). Pada uji posttest, diperoleh 5 soal valid dan layak (Soal 1, 2, 5, 8, dan 9) dengan sebaran kesukaran 2 mudah, 2 sedang, dan 1 sukar serta nilai reliabilitas $r_{11} = 0,70$ (kategori tinggi). Dengan demikian, paket soal terpilih terbukti memenuhi standar kriteria psikometrik dan siap digunakan sebagai instrumen evaluasi keterampilan literasi matematis siswa.

Kata kunci: Literasi matematis, Materi perbandingan, karakteristik psikometrik

Abstract

This study aims to select and validate draft items for a mathematical literacy test on the topic of ratios for seventh-grade students based on an analysis of psychometric characteristics. This quantitative descriptive study involved 36 seventh-grade students from Class 7A at SMP Negeri 1 Soko, who were selected using a cluster random sampling technique. The instruments tested consisted of 10 draft open-ended test items for the pretest and posttest. Psychometric data analysis included empirical validity testing (Pearson Product-Moment correlation), difficulty level, discriminant validity, and reliability testing (Cronbach's Alpha). The results of the study indicate that in the pretest, there were 6 valid items; however, only 5 items (Items 1, 2, 3, 5, and 8) were selected because they had a minimum discriminant power of "Fair" with $DP \geq 0.305$ and a proportional distribution of difficulty (1 easy, 2 moderate, 2 difficult) with a reliability coefficient of $r_{11} = 0.80$ (high category). In the posttest, five valid and appropriate items were identified (Items 1, 2, 5, 8, and 9), with a difficulty distribution of 2 easy, 2 moderate, and 1 difficult, and a reliability coefficient of $r_{11} = 0.70$ (high category). Thus, the selected test item set has been proven to meet psychometric criteria and is ready for use as an instrument for evaluating students' mathematical literacy skills.

Keyword: Mathematical literacy, Ratio and proportion,

PENDAHULUAN

Literasi matematis merupakan salah satu kompetensi utama yang harus dimiliki siswa pada abad ke-21 (Kiawati et al., 2023) hal ini juga diperkuat oleh (Inayah et al., 2026) bahwa kemampuan ini sangatlah penting. Kemampuan ini tidak sekadar menuntut siswa untuk menghafal rumus dan melakukan kalkulasi rutin, melainkan menekankan pada kapasitas individu untuk merumuskan, menggunakan, dan menafsirkan matematika dalam berbagai konteks kehidupan nyata (Junarti et al., 2021). Dalam konteks kurikulum matematika sekolah menengah, materi perbandingan merupakan salah satu topik yang kaya akan konteks dunia nyata, seperti skala peta, laju kendaraan, konversi nilai mata uang, hingga estimasi kebutuhan bahan (Marina et al., 2025). Oleh karena itu, materi perbandingan sangat potensial untuk mengukur dan mengasah keterampilan literasi matematis siswa (Syahdela et al., 2023).

Namun, fakta di lapangan menunjukkan bahwa banyak siswa kelas VII masih mengalami kesulitan ketika dihadapkan pada soal-soal matematika yang berbasis pemecahan masalah konteks riil. Di sisi lain, proses asesmen di sekolah sering kali belum didukung oleh instrumen tes yang teruji secara statistik. Banyak guru membuat soal tes hanya berdasarkan pertimbangan subjektif tanpa melalui pengujian parameter psikometrik yang ketat. Penggunaan instrumen tes yang belum teruji kevalidan dan reliabilitasnya berisiko menghasilkan data capaian siswa yang tidak akurat (bias).

Untuk mengatasi permasalahan tersebut, diperlukan ketersediaan instrumen tes yang tidak hanya selaras dengan indikator literasi matematis, tetapi juga memenuhi kriteria kualitas evaluasi pendidikan. Sebelum suatu paket tes digunakan secara luas untuk mengukur kemampuan siswa, perlu dilakukan proses seleksi dan validasi butir soal dari draf awal yang dikembangkan. Berdasarkan urgensi tersebut, penelitian ini bertujuan untuk melakukan analisis karakteristik psikometrik yang mencakup validitas empiris dengan korelasi *Pearson Product Moment*, reliabilitas, tingkat kesukaran, dan daya pembeda terhadap 10 butir draf soal literasi matematis materi perbandingan. Dari hasil analisis tersebut, dilakukan seleksi ketat untuk menghasilkan butir soal terbaik yang valid, reliabel, dan terkalibrasi secara presisi untuk digunakan pada siswa kelas VII.

METODE

A. Jenis Penelitian dan Subjek

Penelitian ini menerapkan pendekatan kuantitatif dengan data penelitian yang berwujud angka (Syahrizal & Jailani, 2023). Melalui pendekatan deskriptif kuantitatif, kondisi variabel penelitian dipaparkan secara faktual dan didukung oleh analisis data numerik dari lapangan

(Nurhabiba et al., 2023). Penelitian berfokus pada pengujian parameter psikometrik *instrument* tes tanpa memberikan perlakuan eksperimental pada subjek. Populasi dalam penelitian ini adalah seluruh siswa kelas VII di SMP Negeri 1 Soko yang terbagi ke dalam beberapa paralel kelas dari kelas VII A hingga VII I. Namun karena terdapat kekurangan ruangan sehingga kelas VII I di tambahkan di beberapa kelas sehingga hanya terdapat kelas VII A-VII H Teknik pengambilan sampel yang digunakan adalah *Cluster Random Sampling*, karena populasi secara alami sudah terkelompokkan dalam bentuk kelas-kelas homogen. Penentuan sampel dilakukan dengan cara mengacak seluruh kelompok kelas yang ada, sehingga terpilih Kelas VII A secara acak sebagai kelas sampel uji coba yang mewakili karakteristik populasi siswa kelas VII di sekolah tersebut. Seluruh siswa di dalam kelas VII A sebanyak 36 siswa dijadikan subjek uji coba instrumen tes literasi matematis.

B. Prosedur Penelitian

Prosedur penelitian dilaksanakan melalui tiga tahap utama:

1. Tahap Persiapan

Merancang kisi-kisi instrumen berdasarkan indikator literasi matematis (merumuskan, menggunakan, dan menafsirkan) pada materi perbandingan, serta menyusun draf awal sebanyak 10 butir soal uraian beserta rubrik penskorannya.

2. Tahap Uji Coba

Menguji coba draft 10 butir soal tersebut kepada 36 siswa kelas VII A dalam alokasi waktu yang sudah ditentukan.

3. Tahap Analisis dan Seleksi

Mengolah data skor jawaban siswa menggunakan analisis karakteristik psikometrik untuk memilih soal terbaik.

C. Teknik Analisis Data Psikometrik

Data hasil jawaban siswa dianalisis menggunakan empat kriteria psikometrik utama:

1. Uji Validitas Butir (*pearson product moment*)

Validitas empiris dihitung dengan mengkorelasikan skor tiap butir soal (X) dengan skor total (Y) menggunakan rumus korelasi *pearson product moment* (Sari et al., 2023):

$$r_{xy} = \frac{N \sum XY - (\sum X)(Y)}{\sqrt{(N \sum X^2 - (\sum X)^2)(N \sum Y^2 - (\sum Y)^2)}}$$

Butir soal dinyatakan valid jika $r_{hitung} > r_{tabel}$ pada taraf signifikan 0,05 serta nilai r tabel yang diperoleh dari $df = 36 - 2 = 34$ ($r_{tabel} = 0,329$).

2. Analisis Tingkat Kesukaran

Indeks tingkat kesukaran dihitung dengan membandingkan rata-rata skor butir ($\bar{X}$) dengan skor maksimalnya (S_{maks}) berikut rumus tingkat kesukaran (Rismawati et al., 2024):

$$P = \frac{\bar{X}}{S_{maks}}$$

Kategori kesukaran diklasifikasikan menjadi: $P > 0,70$ (mudah), $0,31 \leq P \leq 0,70$ (sedang), dan $P \leq 0,30$ (sukar).

3. Analisis Daya Pembeda

Daya pembeda dihitung dengan membagi sampel menjadi 18 siswa kelompok atas (A) dan 18 siswa kelompok bawah (B) berdasarkan total skor, berikut rumus daya pembeda menurut Yusrizal dalam (Rismawati et al., 2024):

$$DP = \frac{\bar{X}_A - \bar{X}_B}{S_{maks}}$$

Kategori daya pembeda:

$DP \geq 0,40$ (Baik), $0,20 \leq DP \leq 0,39$ (cukup), dan $DP \leq 0,19$ (tidak baik).

4. Uji Reabilitas

Menurut (Rismawati et al., 2024) uji konsistensi internal dilakukan terhadap paket soal terpilih menggunakan rumus *Cronbach's Alpha*:

$$r_{11} = \left(\frac{k}{k-1} \right) \left(1 - \frac{\sum \sigma_b^2}{\sigma_t^2} \right)$$

HASIL DAN PEMBAHASAN

A. Hasil Rekapitulasi Analisis Psikometrik 10 Butir Soal Pretest dan Posttest

Tabel 1. Rekapitulasi Hasil Analisis Psikometrik 10 Draf Butir Soal Pretest Tes Literasi

No	$r - hitung$	Status Validitas	Tingkat Kesukaran	Kategori Kesukaran	Daya Pembeda	Kategori Daya Pembeda	Keputusan Akhir
1	0,722	Valid	0,80	Mudah	0,319	Cukup	Terpilih
2	0,767	Valid	0,69	Sedang	0,458	Baik	Terpilih
3	0,747	Valid	0,00	Sukar	0,306	Cukup	Terpilih
4	0,332	Valid	0,93	Mudah	0,056	Tidak Baik	Dibuang
5	0,675	Valid	0,28	Sukar	0,347	Cukup	Terpilih
6	-0,032	Tidak Valid	0,47	Mudah	-0,097	Tidak Baik	Dibuang
7	-0,033	Tidak Valid	0,67	Sedang	0,014	Tidak Baik	Dibuang
8	0,667	Valid	0,58	Sedang	0,444	Baik	Terpilih
9	0,241	Tidak Valid	0,67	Sedang	0,000	Tidak Baik	Dibuang
10	0,310	Tidak Valid	0,81	Mudah	0,097	Tidak Baik	Dibuang

Tabel 2. Rekapitulasi Hasil Analisis Psikometrik 10 Draf Butir Soal Posttest Tes Literasi

No	$r - hitung$	Status Validitas	Tingkat Kesukaran	Kategori Kesukaran	Daya Pembeda	Kategori Daya Pembeda	Keputusan Akhir
----	--------------	------------------	-------------------	--------------------	--------------	-----------------------	-----------------

1	0,710	Valid	0,83	Mudah	0,32	Cukup	Terpilih
2	0,870	Valid	0,49	Sedang	0,49	Baik	Terpilih
3	0,266	Tidak Valid	0,69	Sedang	0,11	Tidak Baik	Dibuang
4	0,324	Tidak Valid	0,73	Mudah	-0,01	Tidak Baik	Dibuang
5	0,740	Valid	0,29	Sukar	0,44	Baik	Terpilih
6	0,198	Tidak Valid	0,70	Sedang	0,10	Tidak Baik	Dibuang
7	0,067	Tidak Valid	0,67	Sedang	0,03	Tidak Baik	Dibuang
8	0,656	Valid	0,51	Sedang	0,57	Baik	Terpilih
9	0,573	Valid	0,72	Mudah	0,26	Cukup	Terpilih
10	0,279	Tidak Valid	0,81	Mudah	0,13	Tidak Baik	Dibuang

B. Pembahasan Parametrik Psikometrik

Berdasarkan uji coba instrumen tes literasi matematis materi perbandingan kepada 36 siswa, diperoleh data pengujian parameter psikometrik sebagai berikut:

1. Analisis Validitas Butir

a. Uji Validitas Butir Soal Pretest

Dengan taraf signifikan 5% dan $df = 36 - 2 = 34$, diperoleh nilai $r - tabel$ 0,329. Dari 10 draft butir soal yang diuji:

1. Terdapat 6 butir soal valid: yakni soal nomor 1 (0,722), soal 2 (0,767), soal 3 (0,747), soal 4 (0,332), soal 5 (0,675), dan soal 8 (0,667).
2. Terdapat 4 butir soal tidak valid:

Soal nomor 6 (-0,032), soal nomor 7 (-0,033), soal nomor 9 (0,241), dan soal 10 (0,310).

b. Uji Validitas Butir Soal Post-test

Pada uji posttest dengan jumlah subjek 36 siswa dengan taraf signifikan 0,05 diperoleh $r - tabel$ 0,329. Dari 10 butir soal yang diujikan, diperoleh:

- 1) 5 butir soal valid ($r_{hitung} > 0,329$). Soal 1 ($r = 0,710$), soal 2 ($r = 0,870$), soal 5 ($r = 0,740$), soal 8 ($r = 0,656$), dan soal 9 ($r = 0,573$).
- 2) 5 butir soal tidak valid, yakni soal 3 ($r = 0,266$), soal 4 ($r = 0,324$), soal 6 ($r = 0,198$), soal 7 ($r = 0,067$), dan soal 10 ($r = 0,279$).

2. Analisis Tingkat Kesukaran

a. Tingkat Kesukaran Butir Soal Pretest

Indeks tingkat kesukaran dari 10 soal bervariasi 0,00 hingga 0,93. Terdapat 4 soal kategori Mudah (soal 1, 4, 6, 10), 4 soal kategori sedang (Soal 2, 7, 8, 9), dan 2 soal kategori sukar/sulit (soal 3 dan 5).

b. Tingkat Kesukaran Butir Soal Post-Test

Tingkat kesukaran pada data posttest tersebar ke dalam tiga kategori:

Kategori Mudah, soal dikategorikan mudah ketika tingkat kesukaran $P > 0,70$. Dari hasil uji tingkat kesukaran terdapat 4 soal posttest yang memiliki kategori mudah yakni

soal 1 (0,83), soal 4 (0,73), soal 9 (0,72), dan soal 10 (0,81). Kategori Sedang, dikategorikan sedang ketika tingkat kesukaran soal tersebut $0,31 \leq P \leq 0,70$). Dari hasil uji terdapat soal 2 (0,49), soal 3 (0,69), soal 6 (0,70), soal 7 (0,67), dan soal 8 (0,51). Kategori Sukar $P \leq 0,30$ hanya terdapat 1 butir soal yang masuk dalam kategori ini, yaitu soal 5.

3. Analisis Daya Pembeda

a. Daya Pembeda Butir Soal Pretest

Pengujian daya pembeda membagi sampel menjadi 18 siswa kelompok atas dan 18 siswa kelompok bawah (50%-50%). Hasil analisis menunjukkan 2 soal berkategori baik (soal 2 dan 8), 3 soal berkategori cukup (soal 1, 3, dan 5), serta 5 soal berkategori tidak baik (soal 4, 6, 7, 9, dan 10) karena memiliki nilai $DP \leq 0,097$

b. Daya Pembeda Butir Soal *Post-test*

Daya pembeda dihitung dengan membandingkan 18 siswa kelompok atas dan 18 siswa kelompok bawah. Kategori baik dengan $DP \geq 0,40$ yakni soal 2, 5, dan 8 dengan masing-masing DP 0,49, 0,44, 0,57. Kategori cukup ketika soal memiliki daya pembeda $0,20 \leq DP \leq 0,39$ yakni soal 1 dan 9 dengan DP 0,32 dan 0,26. Dan kategori tidak baik ketika $DP \leq 0,19$ yakni soal 3, 4, 6, 7 dan 10 dengan $DP \leq 0,13$.

4. Analisis Uji Reliabilitas

Dari hasil uji reabilitas soal pretest dan posttest mendapatkan 5 butir soal pretest final yang terpilih (soal 1, 2, 3, 5, dan 8), dilakukan pengujian reliabilitas *Cronbach's Alpha*. Diperoleh nilai varians total sebesar 21, 57 dengan jumlah varians butir 7,59, sehingga menghasilkan koefisien reliabilitas sebesar $r_{11}=0,80$. Selain itu terdapat 5 butir soal posttest yang dinyatakan valid dan berkinerja baik yakni soal 1, 2, 5, 8, dan 9, dilakukan uji reliabilitas menggunakan *Cronbach's Alpha*. Diperoleh jumlah varians butir sebesar 9, 10 dan varians total sebesar 21,31, sehingga menghasilkan koefisien reliabilitas sebesar $r_{11}=0,70$. Hal ini menunjukkan bahwa instrumen pretest dan posttest terpilih memiliki reliabilitas berkategori tinggi dan konsisten sebagai alat ukur.

C. Proses Seleksi dan Karakteristik 5 Soal Terpilih

Proses penyaringan 5 soal dilakukan dengan mengeliminasi soal yang tidak memenuhi syarat kualitatif maupun kuantitatif, yakni soal 6, 7, 9, dan 10 digugurkan karena nilai $r_{hitung} < r_{tabel}$ dan daya pembedanya sangat rendah $DP \leq 0,097$. Dan soal 4 memiliki status valid secara statistik dengan skor $r = 0,332$, namun digugurkan karena tingkat kesukarannya terlalu tinggi/mudah yakni sebesar 0,93 dan daya pembedanya tidak baik yakni sebesar 0,056. Sehingga

terpilih 5 butir soal final soal 1, 2, 3, 5, dan 8 sehingga yang memenuhi seluruh syarat validitas empiris, daya pembeda minimal cukup dengan nilai $DP \geq 0,305$. Serta membentuk komposisi tingkat kesukaran yang proporsional (1soal mudah, 2 soal sedang, dan 2 soal sukar).

Penentuan 5 butir soal posttest terbaik dilakukan secara langsung dan lugas, yaitu mengeliminasi soal gugur: soal 3, 4, 6, 7, dan 10 dieliminasi karena tidak memenuhi nilai ambang batas validitas ($r_{hitung} < 0,329$) serta memiliki indeks daya pembeda yang tidak baik $DP \leq 0,13$. Penetapan paket final diperoleh 5 butir soal terpilih, yaitu soal 1, 2, 5, 8, dan 9. Karakteristik paket terpilih didasarkan pada proporsi kesukaran berdasarkan tingkat kesukarannya terdapat 2 soal kategori mudah yakni soal 1 & 9, 2 soal sedang yakni soal 2 & 8, dan 1 soal sukar yakni soal nomor 5. dan indikator kualitas kelima soal yang terpilih memiliki validitas empiris yang signifikan, daya pembeda minimal “cukup” hingga “baik”, serta memiliki tingkat konsistensi internal/reliabilitas yang tinggi yakni sebesar 0,70.

SIMPULAN

Berdasarkan analisis karakteristik psikometrik terhadap draf instrument tes literasi matematis materi perbandingan yang diujicobakan pada 36 siswa kelas VII A SMP Negeri 1 Soko, penelitian ini menghasilkan beberapa temuan utama:

1. Generalisasi hasil validasi pretest dan posttest

Pengujian statistik membuktikan bahwa tidak semua draf soal yang disusun secara teoritis memiliki perform psikometrik yang baik di lapangan. Dari 10 draf butir soal yang diujikan, diperoleh 5 butir soal final pretest (soal 1, 2, 3, 5, dan 8) serta 5 butir soal final posttest (soal 1, 2, 5, 8, dan 9) yang terbukti valid dan empiris $r_{hitung} > 0,329$, berdaya pembeda memadai ($DP \geq 0,26$), serta memiliki sebaran tingkat kesukaran yang proporsional. Kedua paket instrument final ini memiliki koefisien reliabilitas *cronbach's alpha* yang tergolong tinggi ($r_{11} = 80$ untuk pretest dan $r_{11} = 70$ untuk posttest), sehingga terkalibrasi secara presisi dan konsistensi untuk mengukur keterampilan literasi matematis siswa kelas VII.

2. Rekomendasi langkah selanjutnya

a. Bagi pendidik dan praktisi

Paket 5 butir soal terpilih ini direkomendasi untuk digunakan sebagai instrument tes baku dalam mengukur pencapaian literasi matematis siswa pada materi perbandingan, baik untuk asesmen diagnostic maupun evaluasi hasil belajar.

b. Bagi peneliti selanjutnya

Peneliti disarankan untuk melakukan uji coba instrument pada skala populasi yang lebih luas dan bervariasi guna memperkuat generalisasi instrument. Selain itu, penggunaan

analisis teori respons butir dapat dipertimbangkan pada penelitian selanjutnya untuk mengevaluasi parameter butir soal secara lebih.

DAFTAR PUSTAKA

- Inayah, P., Novianti, D. E., & Pupananda, D. R. (2026). *Pengaruh Literasi Digital Terhadap Kemampuan Pemecahan Masalah Matematis Siswa Pondok dan Non-pondok di Era. 4*(2), 215–222.
- Junarti, J., Sujiran, S., & Utami, A. D. (2021). *Strategi Analogi-Abstraksi untuk Membangun Literasi Matematis pada Eksistensi Elemen Identitas dan Elemen Invers pada Materi Grup. 2*(1), 8–21.
- Kiawati, E. S., Junaedi, B., & Tabrani, M. B. (2023). Penerapan Model Pembelajaran Berbasis Masalah untuk Meningkatkan Kemampuan Literasi Matematis Siswa. *Jurnal Cendikia, 7*(3), 2465–2474. <https://doi.org/https://doi.org/10.31004/cendekia.v7i3.2213>
- Marina, R., Zulkardi, Z., Susanti, E., & Meryansumayeka, M. (2025). Analisis Kemampuan Representasi Matematis Siswa SMP pada Materi Perbandingan Menggunakan Konteks Jajanan. *Jurnal Pendidikan Matematika Dan Sains, 13*(1), 31–46.
- Nurhabiba, F. D., Misdalina, M., & Tanzimah, T. (2023). Kemampuan Higher Order Thingking Skill (HOTS) dalam Pembelajaran Berdiferensiasi SD 19 Palembang. *Didaktik: Jurnal Ilmiah PGSD FKIP Univversitas Mandiri, 9*(3), 492–504.
- Rismawati, R., Mukhlisa, N., & Lukman, L. (2024). *Analisis Butir Soal Penilaian Akhir Ssemester (PAS) Ganjil Tahun Ajaran 2023/2024 Pada Mata Pelajaran Matematika UPT SD Negeri 228 Pinrang. 9*(3), 189–202.
- Sari, M. F., Hadiati, N. R., & Sihotang, P. W. (2023). Pearson Correlation Analysis of Total Population and Number of Motorized Vehicles in Jambi Province. *Jurnal Statistika Universitas Jambi, 2*(1), 39–44. <https://online-journal.unja.ac.id/multiproximity39https://doi.org/10.22437/multiproximity.v2i1.25568>
- Syahdela, A. A., Junarti, J., & Zuhriah, F. (2023). *Pelevelan Kemampuan Literasi Matematis Siswa Sekolah Menengah Pertama pada Materi Perbandingan. 1*(1), 94–101.
- Syahrizal, H., & Jailani, M. S. (2023). Jenis-Jenis Penelitian dalam Penelitian Kuantitatif dan Kualitatif. *JQOSIIM: Jurnal Pendidikan, Sosial & Humaniora, 1*(1), 13–23. <http://ejournal.yayasanpendidikandzurriyatulquran.id/index.php/qosim>.