

Analisis Kualitas Butir Soal Asesmen Sumatif Matematika Pilihan Ganda Semester Genap Kelas VIII MTs Baitul Ma'mur Tahun Ajaran 2023/2024

Analysis of the Quality of Multiple-Choice Mathematics Summative Assessment Items for the Even Semester of Grade VIII at MTs Baitul Ma'mur in the 2023/2024 Academic Year

Ikhsan Abdullah^{*1}, Mega Kusuma Listyotami², Rolina Amriyanti Ferita³, Lisa⁴, Dina Maulida⁵, Williza Yanti⁶, Farid Hidayat⁷

^{1,3,5,6,7}Universitas Nahdlatul Ulama Kalimantan Selatan, Jl. Ahmad Yani Km. 12.500, Banjar, Kalimantan Selatan 70652

²Universitas Muhammadiyah Kupang, Jl. KH Ahmad Dahlan No. 17 Walikota Baru, Kota Kupang, Provinsi Nusa Tenggara Timur

⁴MTs Baitul Ma'mur, Jalan Sumber Murni, RT.5/RW.2, Landasan Ulin Utara, Kec. Liang Anggang, Kota Banjar Baru, Kalimantan Selatan

¹ikhsanabdullah@gmail.com, ²megakusumalistyotami@gmail.com, ³rolina.amriyanti@unukase.ac.id, ⁴lisa6113@guru.sd.belajar.id,

⁵edu.dinams@gmail.com, ⁶willizayanti@gmail.com, ⁷ridhi.frd0989@gmail.com

Format Kutipan: Abdullah, I. Listyotami, M. K., Ferita, R. A., Lisa, L., Maulida, D., Yanti, W., & Hidayat, F. (2025). Analisis Kualitas Butir Soal Asesmen Sumatif Matematika Pilihan Ganda Semester Genap Kelas VIII MTs Baitul Ma'mur Tahun Ajaran 2023/2024. *Nusantara Journal of Education and Social Science*, 2(1), hal. 27-32. <https://doi.org/10.69959/nujess.v2n1.49>

RIWAYAT ARTIKEL

Dikirim: 4 Juli 2024

Revisi Akhir: 28 Januari 2025

Diterbitkan: Januari 2025

Tersedia Daring Sejak: 31 Januari 2025

KATA KUNCI

Analisis Butir Soal

Penilaian Sumatif

Matematika

KEYWORDS

Item Analysis

Summative Assessment

Math Questions

ABSTRAK

Tujuan penelitian ini yakni untuk melihat kualitas soal asesmen sumatif pilihan ganda mata pelajaran matematika di MTs Baitul Ma'mur semester genap kelas VIII tahun ajaran 2023/2024. Soal akan diuji dengan menggunakan uji validitas, tingkat kesulitan, reliabilitas, pengecoh, dan uji daya pembeda. Analisis deskriptif terhadap dua puluh siswa adalah metode kuantitatif yang digunakan dalam penelitian ini. Data disajikan dalam bentuk lembar soal, lembar jawaban, dan kunci jawaban. Evaluasi data dilakukan dengan teknik deskripsi persentase yang dihasilkan dari pengolahan data Excel. Berdasarkan temuan penelitian, uji validitas mencakup delapan soal dengan validitas sangat rendah dan tiga soal dengan validitas rendah. Namun, pada hasil uji reliabilitas pada soal tersebut memiliki tingkat konsistensi atau keajegan tinggi dan uji tingkat kesukaran pada soal tersebut bertaraf sedang, artinya tingkat kesukaran termasuk kategori baik. Selanjutnya, hasil uji daya pembeda pada soal tersebut dimasukkan ke dalam klasifikasi baik dan bermanfaat. Selanjutnya, hasil uji efektivitas pengalih perhatian pada soal tersebut dimasukkan ke dalam kriteria berfungsi. Jadi, soal asesmen sumatif pilihan ganda untuk mata pelajaran matematika di MTs Baitul Ma'mur semester genap kelas VIII tahun ajaran 2023/2024 dinilai baik, meskipun pada bagian validitasnya perlu ditingkatkan.

ABSTRACT

The goal of this study was to evaluate the quality of multiple-choice summative assessment questions in mathematics at MTs Baitul Ma'mur, even semester, grade VIII, academic year 2023/2024, using a validity test, difficulty level test, reliability test, distractor test, and discriminating power test. This study strategy employs a quantitative approach in the form of a descriptive analysis of 20 pupils. Data are supplied as question sheets, answer sheets, and answer keys. The data were analyzed using percentage description approaches obtained from Excel data processing. According to the findings of this study, the validity test contained three questions with low validity and eight questions with very poor validity. However, the results of the reliability test on the question showed a high level of consistency, and the difficulty level of the test on the question was moderate, indicating that the level of difficulty was in the good range. Furthermore, the findings of the discriminating power test on the question were incorporated into the good and useful category. Furthermore, the results of the distraction effectiveness test on the question were factored into the functional criteria. So, the multiple-choice summative assessment questions for mathematics topics at MTs Baitul Ma'mur, even semester, grade VIII, 2023/2024 academic year are deemed satisfactory, although the validity portion might be enhanced.

Artikel ini dapat diakses secara terbuka (open access) di bawah lisensi CC-BY-SA



PENDAHULUAN

Proses pembelajaran matematika tidak bisa lepas dari asesmen atau penilaian. Dengan adanya penilaian, guru dapat mengukur kemampuan siswa, menemukan kesulitan belajar yang dialami siswa, serta memperbaiki metode pengajaran (Ferita & Retnawati, 2016). Dalam satu kelas, tentunya terdapat beragam kemampuan siswa mengingat berasal dari latar belakang yang berbeda-beda sehingga memerlukan penilaian yang tepat (Riadi & Ferita, 2016). Asesmen sumatif merupakan salah satu bentuk penilaian yang digunakan untuk mengetahui sejauh mana siswa menguasai materi yang telah diajarkan. Oleh karena itu, kualitas butir soal dalam asesmen tertulis sangat penting karena akan berdampak langsung pada keakuratan hasil penilaian. Hasil tes digunakan untuk menilai pengetahuan siswa. Menurut Gronlund (Abdullah, 2015), hasil belajar digunakan untuk menilai tujuan pendidikan. Soal-soal yang digunakan harus dapat menentukan tercapai atau tidaknya tujuan kurikulum, sehingga dapat digunakan untuk menilai pencapaian tujuan pembelajaran.

Seluruh aktivitas pembelajaran yang dilaksanakan oleh guru dan siswa memberikan kontribusi terhadap keberhasilan proses pembelajaran, menurut Aunurrahman (Suprpta, 2020). Guru tidak hanya harus memiliki pengetahuan tentang materi dan strategi pembelajaran, tetapi juga harus memiliki pengetahuan tentang metode evaluasi yang tepat. Proses dan hasil belajar siswa dapat diukur secara signifikan ketika mengambil keputusan tentang cara meningkatkan kemajuan belajar siswa. Beberapa lembaga pendidikan masih mengandalkan uji coba dan analisis menyeluruh untuk menguji validitas isi, meskipun analisis butir soal sangat penting. Validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas pengecoh adalah faktor yang harus dipertimbangkan saat menganalisis butir soal (Amalia et al., 2021; Solichin, 2017), selain itu ada pula yang menekankan penilaian pada validitas, reliabilitas, dan praktikalitas (Arif, 2016). Kriteria penilaian validitas isi dihitung dengan menggunakan instrumen dan acuan kriteria tertentu (Ihsan, 2015). Soal pilihan ganda juga dapat dianalisis dengan Rasch Model untuk memperoleh tingkat abilitas siswa dan tingkat kesulitan soal (Kurniawan & Andriyani, 2018), atau menggunakan perangkat lunak yang sudah lazim ada di komputer yaitu Microsoft Office Excel (Purwanti, 2014). Penggunaan soal pilihan ganda dalam penilaian matematika banyak digunakan dan dapat untuk mengukur kemampuan berpikir tingkat tinggi siswa (Ferita & Fitria, 2019). Namun, tidak semua soal pilihan ganda memiliki kualitas yang sama. Agar hasil penilaian akurat, soal harus dirancang dengan cermat dan memastikan pengecoh berfungsi dengan baik. Soal yang baik harus memenuhi beberapa kriteria, seperti validitas, reliabilitas, tingkat kesulitan yang tepat, dan kemampuan membedakan siswa yang satu dengan yang lain.

MTs Baitul Ma'mur merupakan salah satu madrasah di wilayah Banjarbaru yang melakukan pengumpulan data tentang capaian tujuan pembelajaran siswa terhadap mata pelajaran matematika. Soal penilaian pilihan ganda sumatif untuk mata pelajaran matematika kelas VIII semester genap yang berjumlah 25 soal dikembangkan melalui diskusi antar guru matematika MTS Baitul Ma'mur. Kualitas setiap soal ujian dalam proses analisis saat asesmen sumatif berlangsung. Penelitian ini akan selesai dalam satu bulan, pada bulan Mei 2024.

Hasil ujian kurang ideal, dengan anak-anak yang menerima nilai rendah. Ini menunjukkan bahwa beberapa murid tidak dapat menjawab pertanyaan ujian. Akibatnya, tinjauan mendalam terhadap kriteria untuk pertanyaan yang disusun diperlukan. Analisis butir soal ini dimaksudkan untuk menjadi koreksi di masa mendatang untuk mengembangkan pertanyaan berkualitas lebih tinggi yang lebih efektif bila digunakan sebagai pertanyaan ujian siswa.

METODE PENELITIAN

Menurut Tilaar & Hasriyanti (2019) analisis deskriptif analitik adalah salah satu pendekatan kuantitatif yang dipakai pada sebuah penelitian. Analisis deskriptif menganalisis data dengan memberikan deskripsi kepada mereka. Studi ini mengumpulkan data dalam bentuk nilai-nilai numerik. Data yang dikumpulkan berasal dari soal asesmen sumatif pilihan ganda untuk mata pelajaran matematika di MTS Baitul Ma'mur selama semester genap tahun ajaran 2023/2024. Soal tersebut terdiri atas lembar soal, lembar jawaban, dan kunci jawaban. Kualitas soal termasuk validitas, keajegan, daya pembeda, tingkat kesulitan, dan efektivitas pengecoh.

Adapun teknik dalam penelitian ini sebagai berikut:

1. Uji Validitas

Semua pertanyaan dalam kuesioner diuji validitasnya. Teknik korelasi *product moment* yang dipakai, dengan tabel dengan taraf signifikansi 5% dan jumlah sampel sebanyak 25 butir (Kamilah, 2014). Menurut Arikunto (2013), metode korelasi momen produk digunakan dengan rumus berikut.

$$r_{xy} = \frac{N \sum xy - \sum x \sum y}{\sqrt{[N(\sum x^2) - (\sum x)^2][N(\sum y^2) - (\sum y)^2]}}$$

Keterangan:

r_{xy} = koefisiensi validitas

x = skor atau betul tiap butir soal

y = jumlah skor total

N = jumlah peserta tes

Soal dianggap valid jika r hitung (r_{xy}) lebih besar atau sama dengan r_{tabel} . Sebaliknya, jika r hitung (r_{xy}) lebih kecil dari r_{tabel} , soal dianggap tidak valid. Suherman & Sukjaya (Lovisia, 2018) menetapkan standar berikut untuk menilai koefisien validitas:

- 0,91 – 1,00 = sangat tinggi
- 0,71 – 0,90 = tinggi
- 0,41 – 0,70 = sedang
- 0,21 – 0,40 = rendah
- *negatif* – 0,20 = sangat rendah

2. Uji Reliabilitas

Uji reliabilitas menentukan tingkat konsistensi instrumen tes. Rumus Cronbach-Alpha, yang menyatakan bahwa:

$$\text{koefisien alpha } (r_{11}) = \left(\frac{k}{k-1}\right)\left(1 - \frac{\sum SD_i^2}{SD^2}\right)$$

Keterangan:

k = jumlah item

SD^2 = variansi tiap item

SD_i^2 = variansi skor total

Menurut Guilford (Soleh et al., 2017) kategori koefisien reliabilitas sebagai berikut.

- $0,80 < r_{11} \leq 1,00$ reliabilitas sangat tinggi
- $0,60 < r_{11} \leq 0,80$ reliabilitas tinggi
- $0,40 < r_{11} \leq 0,60$ reliabilitas sedang
- $0,20 < r_{11} \leq 0,40$ reliabilitas rendah
- $-1,00 \leq r_{11} \leq 0,20$ reliabilitas sangat rendah (tidak reliabel)

3. Uji Tingkat Kesukaran

Soal dinilai dengan indeks kesukaran (IK) yang berkisar dari 0,00 sampai 1,00. Farida & Musyarofah (2021) mengatakan bahwa tingkat kesulitan dilihat menggunakan rumus berikut.

$$P = \frac{B}{J_s}$$

Keterangan:

P = indeks kesukaran

B = jumlah siswa menjawab benar

J_s = total seluruh siswa

Menurut Sudjana (Hamimi et al., 2020) klasifikasi indeks tingkat kesukaran sebagai berikut:

- $0,00 \leq IK < 0,30$ = soal tergolong sukar
- $0,30 \leq IK < 0,70$ = soal tergolong sedang
- $0,70 \leq IK \leq 1,00$ = soal tergolong mudah

4. Uji Daya Pembeda

Uji daya pembeda bertujuan untuk menentukan apakah ada perbedaan dalam hasil jawaban siswa (Arifien & Wijayanto, 2021):

$$r_{pbis} = \frac{x_i - x_{total}}{SD} \sqrt{p_i/q_i}$$

Keterangan:

r_{pbis} = koefisiensi daya pembeda

x_i = rata-rata skor tiap butir

x_{total} = rata-rata skor total

SD = standar deviasi atau simpangan baku skor total

p_i = proporsi siswa yang menjawab benar butir ke- i

q_i = proporsi siswa yang menjawab salah butir ke- i

Adapun klasifikasinya sebagai berikut:

- $DP < 0$ = tidak baik dan diganti
- $0 \leq DP < 0,3$ = kurang baik dan diperbaiki
- $DP \geq 0,3$ = baik dan digunakan

5. Uji Pengecoh

Analisis pengecoh bertujuan untuk mengetahui seberapa baik pengalih perhatian berfungsi dalam soal yang telah dibuat. Angka persentase menentukan berhasil atau tidaknya pengalih perhatian, yaitu apakah paling sedikit 5% peserta tes dipilih oleh peserta tes. Wardoyo & Suprpto (2014) mengusulkan perhitungan tingkat efektivitas pengalih perhatian dengan menggunakan rumus berikut.

$$D = \frac{nD}{J_s}$$

Keterangan:

D = indeks pengecoh koefisiensi validitas

nD = jumlah siswa yang memilih pengecoh koefisiensi validitas

J_s = jumlah seluruh siswa yang mengikuti tes

Adapun klasifikasinya sebagai berikut.

- $D = 0,00$ = pengecoh tidak berfungsi
- $D > 0,00$ = pengecoh berfungsi

HASIL PENELITIAN DAN PEMBAHASAN

Hasil

Data soal, jawaban, dan kunci jawaban dari tes sumatif pilihan ganda mata kuliah matematika selama semester genap tahun ajaran 2023/2024 digunakan pada penelitian ini. Data tersebut kemudian dikaji menggunakan metode yang dijelaskan di bagian metode. Hasilnya akan disajikan dalam bentuk tabel.

1. Uji Validitas

Hasil uji validitas memakai teknik korelasi *product moment* menurut Arikunto (2013) disajikan dalam Tabel 1.

Tabel 1. Hasil Pengolahan Data Menggunakan Excel

No.	Nomor Soal	Keterangan	Jumlah Soal	Persentase
1	-	Validitas Sangat Tinggi	0	0%
2	1, 6, 8, 16	Validitas Tinggi	4	16%
3	3, 5, 7, 11, 12, 17, 18, 20, 21, 23	Validitas Sedang	10	40%
4	2, 4, 19	Validitas Rendah	3	12%
5	9, 10, 13, 14, 15, 22, 24, 25	Validitas Sangat Rendah	8	32%
Jumlah Soal			25	100%

2. Uji Reliabilitas

Pada uji reliabilitas ini memakai rumus Cronbach-Alpha. Hasil yang diperoleh yaitu 0,68143563.

Tabel 2. Hasil Pengolahan Data Menggunakan Excel

Koefisiensi Alpha	
0,68143563	

3. Uji Tingkat Kesukaran

Uji tingkat kesulitan dilakukan dengan menghitung persentase jawaban benar siswa. Hasil yang didapatkan seperti pada Tabel 3.

Tabel 3. Hasil Uji Tingkat Kesukaran

No.	Nomor Soal	Keterangan	Jumlah Soal	Persentase
1	12, 13, 21, 25	Soal Tergolong Sukar	4	16%
2	1, 4, 5, 6, 7, 8, 9, 11, 14, 15, 16, 17, 18, 19, 20, 23, 24	Soal Tergolong Sedang	17	68%
3	2, 3, 10, 22	Soal Tergolong Mudah	4	16%
Jumlah Soal			25	100%

4. Uji Daya Pembeda

Berdasarkan data yang telah dianalisis diperoleh daya pembeda yang diuraikan pada Tabel 4.

Tabel 4. Hasil Uji Daya Pembeda

No.	Nomor Soal	Keterangan	Jumlah Soal	Persentase
1	10	Tidak Baik dan Diganti	1	4%
2	14, 15, 25	Kurang Baik dan Diperbaiki	3	12%
3	1, 2, 3, 4, 5, 6, 7, 8, 9, 11, 12, 13, 16, 17, 18, 19, 20, 21, 22, 23, 24	Baik dan Digunakan	21	84%
Jumlah Soal			25	100%

5. Uji Keefektifan Pengecoh

Untuk menganalisis fungsi pengecoh yang dibuat, pada penelitian ini melihat pola jawaban yang dipilih yaitu A, B, C, dan D. Diperoleh hasil seperti Tabel 5.

Tabel 5. Hasil Pengolahan Data Menggunakan Excel

No.	Nomor Soal	Keterangan	Jumlah Soal	Persentase
1	3, 4, 12, 13, 14, 15, 16, 17, 18, 20, 21, 24, 25	Semua Pengecoh Berfungsi	13	52%
2	1, 2, 5, 6, 7, 8, 9, 11, 19, 22, 23	Satu Pengecoh Tidak Berfungsi	11	16%
3	-	Dua Pengecoh Tidak Berfungsi	0	0%
4	-	Tiga Pengecoh Tidak Berfungsi	0	0%
5	10	Semua Pengecoh Tidak Berfungsi	1	32%
Jumlah Soal			25	100%

Pembahasan

Berdasarkan hasil penelitian di atas terdapat satu soal yang tidak bisa dianalisis sebab adanya kesalahan redaksi. Sehingga soal tersebut dinilai bonus. Secara keseluruhan soal yang diujikan tergolong baik, namun perlu adanya perbaikan pada bagian validitas. Berikut ini adalah uraian hasil pengujian validitas, keajegan, daya pembeda, tingkat kesulitan, dan efektivitas pengecoh:

1. Uji Validitas

Tabel 1 menunjukkan bahwa hasil dari 25 soal yang diujikan di kelas VIII MTS Baitul Ma'mur tidak konsisten. Kriteria soal dengan validitas sangat tinggi adalah 0 dengan persentase 0%, 4 dengan persentase 16%, 10 dengan persentase 40%, 3 dengan persentase 12%, dan 8 dengan persentase 32%.

Berdasarkan hasil uji validitas tersebut lebih mendominasi pada kriteria cukup, yaitu sebesar 40% dengan jumlah 10 soal. Namun, perlu dilihat juga pada validitas rendah 3 soal dan validitas sangat rendah 8 soal. Jika digabungkan maka persentasenya sebesar 44% dengan jumlah soal sebanyak 11 soal. Maka dari itu perlu adanya perbaikan pada soal nomor 2, 4, 9, 10, 13, 14, 15, 19, 22, 24, dan 25.

2. Uji Reliabilitas

Tabel 2 menunjukkan hasil analisis uji reliabilitas yang menghasilkan koefisien α sebesar 0,68143563. Jika hasil tersebut diukur menggunakan Guilford (Soleh et al., 2017), Soal tersebut dapat diandalkan untuk digunakan karena telah menunjukkan bahwa memiliki tingkat keajegan tinggi.

3. Uji Tingkat Kesukaran

Tabel 3 menunjukkan empat soal dengan skor menantang (16%), 17 soal skala sedang (68%), dan lima soal skala mudah (16%). Soal-soal yang diujikan mendominasi kategori tingkat kesulitan sedang, seperti yang dikemukakan Sudjana (Hamimi, 2020). Jadi, tingkat kesulitan soal tergolong sedang (68%), dengan total 17 soal.

4. Uji Daya Pembeda

Tabel 4 menunjukkan bahwa setiap pertanyaan memiliki daya pembeda yang unik. Ada satu pertanyaan dengan daya pembeda yang buruk yang telah diganti dengan persentase 4%, tiga pertanyaan dengan daya pembeda yang buruk yang telah diperbaiki dengan persentase 12%, dan 21 pertanyaan dengan daya pembeda yang baik yang telah digunakan dengan rasio 84%. Uji daya pembeda mendominasi kategorisasi yang baik dan digunakan dengan persentase 35% dan 21 pertanyaan.

5. Uji Keefektifan Pengecoh

Berdasarkan Tabel 5 bisa kita tahu bahwa efektivitas pengecoh tiap soal tidak seragam. Terdapat 13 soal semua pengecoh berfungsi dengan persentase 52%, 11 soal satu pengecoh tidak berfungsi dengan persentase 44% dan 1 soal semua pengecoh tidak berfungsi dengan persentase 4% yaitu pada nomor 10. Dominasi efektivitas pengecoh terdapat pada kriteria semua pengecoh berfungsi, yaitu sebesar 52% dengan jumlah 13 soal.

SIMPULAN DAN SARAN

Simpulan

Setelah analisis dan reduksi, dapat disimpulkan sebagai berikut.

1. Uji validitas terdapat 3 soal validitas rendah dan 8 soal validitas sangat rendah. Hal ini perlu adanya perbaikan terhadap soal tersebut.
2. Uji reliabilitas pada soal tersebut menunjukkan tingkat konsistensi atau keajegan tinggi.
3. Tingkat kesulitan soal tergolong sedang.
4. Uji daya pembeda soal termasuk dalam klasifikasi baik dan diterapkan.
5. Uji keefektifan pengecoh pada soal tersebut termasuk dalam kriteria berfungsi.

Saran

Dalam penelitian ini, penulis mengajukan usulan, khususnya:

1. Sebelum menyusun soal ujian, kualitas soal diuji dengan menggunakan uji validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan keefektifan pengecoh.
2. Bagi guru matematika di MTS Baitul Ma'mur diharapkan dapat meningkatkan mutu soal matematika agar dapat digunakan secara efektif untuk menilai hasil belajar siswa.

DAFTAR PUSTAKA

- Abdullah, R. (2015). Urgensi Penilaian Hasil Belajar Berbasis Kelas Mata Pelajaran IPS di Madrasah Tsanawiyah. *Lantanida Journal*, 3(2), 168–181.
- Amalia, N. R., Halik, A., & Mukhlisa, N. (2021). Analisis Butir Soal Matematika Pada Siswa Sekolah Dasar. *PINISI JOURNAL OF EDUCATION*, 1(1), 219–230.
- Arif, Moh. (2016). Pengembangan Instrumen Penilaian Mapel Sains melalui Pendekatan Keterampilan Proses Sains SD/MI. *TA'ALLUM*, 4(1), 123–148.
- Arifien, Moch., & Wijayanto, P. A. (2021). Analisis Kualitas Butir Soal Ujian Tengah Semester Geografi SMA Islam Al Azhar 14 Kota Semarang. *Geodika: Jurnal Kajian Ilmu Dan Pendidikan Geografi*, 5(2), 206–215. <https://doi.org/10.29408/geodika.v5i2.3696>

Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan*. Bumi Aksara.

Farida, F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'arrib: Jurnal Pendidikan Bahasa Arab*, 1(1), 34–44. <https://jurnal.lp2msasbabel.ac.id/index.php/AL-MUARRIB>

Ferita, R. A., & Fitria, M. (2019). Pengembangan Instrumen Tes Pilihan Ganda Untuk Mengukur Tingkat Kemampuan Berpikir Matematika Siswa SMA. *AKSIOMA: Jurnal Program Studi Pendidikan Matematika*, 8(1), 1–10. <https://doi.org/10.24127/ajpm.v8i1.1628>

Ferita, R. A., & Retnawati, H. (2016). Pengembangan Perangkat Penilaian Autentik untuk Pembelajaran Matematika di Kelas VII Semester 1. *PYTHAGORAS: Jurnal Pendidikan Matematika*, 11(1), 69–76. <https://doi.org/10.21831/pg.v11i1.9672>

Hamimi, L., Zamharirah, R., & Rusydy, R. (2020). Analisis Butir Soal Ujian Matematika Kelas VII Semester Ganjil Tahun Pelajaran 2017/2018. *MATHEMA JOURNAL*, 2(1), 57–66.

Ihsan, H. (2015). Validitas Isi Alat Ukur Penelitian: Konsep dan Panduan Penilaiannya. *Pedagogia: Jurnal Ilmu Pendidikan*, 13, 173–179.

Kamilah, E. N. (2014). *Pengaruh Keterampilan Mengajar Guru Terhadap Hasil Belajar Siswa pada Mata Pelajaran Akuntansi* [Thesis S1]. Universitas Pendidikan Indonesia.

Kurniawan, U., & Andriyani, K. D. K. (2018). Analisis Soal Pilihan Ganda dengan Rasch Model. *Statistika*, 6(1), 34–39.

Lovisia, E. (2018). Pengaruh Model Pembelajaran Inkuiri Terbimbing terhadap Hasil Belajar. *Science and Physics Education Journal (SPEJ)*, 2(1), 1–10. <https://doi.org/10.31539/spej.v2i1.333>

Purwanti, M. (2014). Analisis Butir Soal Ujian Akhir Mata Pelajaran Akuntansi Keuangan Menggunakan Microsoft Office Excel 2010. *Jurnal Pendidikan Akuntansi Indonesia*, 12(1), 81–94.

Riadi, A., & Ferita, R. A. (2016). Inovasi Pembelajaran Matematika Berbasis Heterogenitas Siswa. *Math Didactic: Jurnal Pendidikan Matematika*, 2(1), 59–69.

Soleh, A., Khumaedi, M., & Pramono, S. E. (2017). Pengembangan Instrumen Penilaian Mata Pelajaran PKn Standar Kompetensi Memahami Kedaulatan Rakyat dalam Sistem Pemerintahan di Indonesia. *Journal of Educational Research and Evaluation*, 6(1), 71.

Solichin, M. (2017). Analisis Daya Beda Soal, Taraf Kesukaran, Validitas Butir Tes, Interpretasi Hasil Tes dan Validitas Ramalan dalam Evaluasi Pendidikan. *Dirasat: Jurnal Manajemen & Pendidikan Islam*, 2(2), 192–213. www.depdiknas.go.id/evaluasi-proses-

Suprpta, D. N. (2020). Penggunaan Model Pembelajaran Make a Match Sebagai Upaya Meningkatkan Hasil Belajar Bahasa Inggris Siswa. *Journal of Education Action Research*, 4(3), 240–246. <https://ejournal.undiksha.ac.id/index.php/JEAR/index>

Tilaar, A. L. F., & Hasriyanti, H. (2019). Analisis Butir Soal Semester Ganjil Mata Pelajaran Matematika pada Sekolah Menengah Pertama. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, 8(1), 57–68. <https://doi.org/10.15408/jp3i.v8i1.13068>

Wardoyo, W., & Suprptono, E. (2014). Rancang Bangun Program Analisis Butir Soal Pilihan Ganda sebagai Pendukung Proses Evaluasi Pembelajaran. *Jurnal Teknik Elektro*, 6(2).