

## Evaluasi Butir Soal Pilihan Ganda Penilaian Tengah Semester dalam Pembelajaran Tematik untuk Kelas V di SDN Gladak Anyar 4 Pamekasan

Ria Kasanova<sup>1</sup>, Roni Sulistiyono<sup>2</sup>

<sup>1,2</sup>Universitas Sebelas Maret (UNS) Surakarta, Jl. Ir Sutami No.36, Kentingan, Kec. Jebres, Kota Surakarta, Jawa Tengah 57126

kasanovaria@uns.ac.id

### Abstract

This research was conducted with the aim of evaluating the validity, level of difficulty, discriminating power, effectiveness of the distractor, and reliability of multiple choice questions in the Mid Semester Assessment in fifth grade thematic learning at SDN Gladak Anyar 4 Pamekasan. The research method used is a quantitative descriptive approach. There are two themes in the Mid Semester Examination Questions, namely theme 6 with 20 questions and theme 7 with 19 questions. Evaluation of the validity of the questions, level of difficulty, discriminating power, effectiveness of the distractor, and reliability was carried out using Microsoft Excel 2010. The subjects of this study were fifth grade students, and data collection was carried out using documentation techniques. The results of this study indicate that the quality of the questions is high. (1) The validity of the questions in theme 6 were 19 questions (95%) and in theme 7 were 18 questions (94.74%) declared valid. (2) The difficulty level of the questions in theme 6 consisted of 13 questions (68.42%) which were categorized as easy and 2 questions (10.53%) which were categorized as difficult. In theme 7, there are 11 questions (61.11%) which are categorized as easy, so there are questions with difficulty levels that do not meet good quality. (3) The discriminating power of questions on theme 6 consisted of 10 items (52.63%) which were categorized as poor and 1 item (5.26%) which were categorized as good. In theme 7, there are 6 items (33.33%) which are categorized as not good and 3 items (16.67%) which are categorized as good. Therefore, the questions fall into the category of moderate discriminating power. (4) The effectiveness of the distractor in theme 6 consisted of 1 item (5.26%) which was categorized as very good, 8 items (42.11%) which were categorized as good, and 6 items (31.58%) which were categorized as poor. In theme 7, there were 4 items (22.22%) which were categorized as very good, 4 items (22.22%) which were categorized as good, and 3 items (16.67%) which were categorized as poor. Thus, the questions fall into the category of the effectiveness of a good distractor. (5) The reliability of the questions in theme 6 is 0.9592, while in theme 7 it is 0.8950, indicating that the questions have high reliability and high quality.

**Keywords:** Item Analysis, Multiple Choice, Thematic Learning

### Abstrak

Penelitian ini dilakukan dengan tujuan untuk mengevaluasi validitas, tingkat kesukaran, daya pembeda, keefektifan pengecoh, dan reliabilitas soal pilihan ganda dalam Penilaian Tengah Semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan. Metode penelitian yang digunakan adalah pendekatan deskriptif kuantitatif. Terdapat dua tema dalam Soal Ujian Tengah Semester, yaitu tema 6 dengan 20 soal dan tema 7 dengan 19 soal. Evaluasi validitas pertanyaan, tingkat kesulitan, daya pembeda, efektivitas pengecoh, dan reliabilitas dilakukan menggunakan Microsoft Excel 2010. Subjek penelitian ini adalah siswa kelas lima, dan pengumpulan data dilakukan melalui teknik dokumentasi. Hasil penelitian ini menunjukkan bahwa kualitas soal-soal tersebut tinggi. (1) Validitas soal pada tema 6 sebanyak 19 soal (95%) dan pada tema 7 sebanyak 18 soal (94,74%) dinyatakan valid. (2) Tingkat kesulitan soal pada tema 6 terdiri dari 13 soal (68,42%) yang dikategorikan mudah dan 2 soal (10,53%) yang dikategorikan sulit. Pada tema 7, terdapat 11 soal (61,11%) yang dikategorikan mudah, sehingga terdapat soal-soal dengan tingkat kesulitan yang belum memenuhi kualitas yang baik. (3) Daya pembeda soal pada tema 6 terdiri dari 10 item (52,63%) yang dikategorikan kurang baik dan 1 item (5,26%) yang dikategorikan baik. Pada tema 7, terdapat 6 item (33,33%) yang dikategorikan kurang baik dan 3 item (16,67%) yang dikategorikan baik. Oleh karena itu, soal-soal termasuk dalam kategori daya pembeda yang sedang. (4) Keefektifan pengecoh pada tema 6 terdiri dari 1 butir (5,26%) yang dikategorikan sangat baik, 8 butir (42,11%) yang dikategorikan baik, dan 6 butir (31,58%) yang dikategorikan kurang baik. Pada tema 7, terdapat 4 item (22,22%) yang dikategorikan sangat baik, 4 item (22,22%) yang dikategorikan baik, dan 3 item (16,67%) yang dikategorikan kurang baik. Dengan demikian, soal-soal termasuk dalam kategori keefektifan pengecoh yang baik. (5) Reliabilitas soal pada tema 6 sebesar 0,9592, sedangkan pada tema 7 sebesar 0,8950, menunjukkan bahwa soal-soal tersebut memiliki reliabilitas yang tinggi dan berkualitas tinggi.

**Kata Kunci:** Analisis Butir Soal, Pilihan Ganda, Pembelajaran Tematik

---

Corresponding author: Ria Kasanova, Roni Sulistiyono

Email Address: kasanovaria@uns.ac.id & roni.sulistiyono@student.uns.ac.id (Jl. Ir Sutami No.36, Kentingan, Kec. Jebres, Kota Surakarta, Jawa Tengah 57126)

Received 19 June 2023, Accepted 22 June 2023, Published 28 June 2023

## **PENDAHULUAN**

Guru memiliki tanggung jawab untuk tidak hanya mengajar dan membimbing, tetapi juga mengevaluasi hasil belajar peserta didik (Wulansari Ramadhany et al., 2022). Evaluasi pembelajaran merupakan kegiatan penting dalam proses pembelajaran yang bertujuan untuk melihat tingkat keberhasilan dalam pencapaian tujuan pembelajaran dan menjadi dasar dalam pengambilan keputusan akhir (Maulidiyah et al., 2021). Evaluasi dilakukan secara sistematis dengan menggunakan kriteria tertentu untuk menentukan nilai suatu tujuan, kegiatan, keputusan, unjuk kerja, proses, individu, atau objek (Stephenson et al., 2020). Dengan demikian, evaluasi hasil belajar merupakan kegiatan terstruktur yang meliputi penilaian dan pengukuran hasil belajar peserta didik, yang nantinya digunakan dalam pengambilan keputusan atau pembuatan kebijakan.

Evaluasi biasanya dilakukan setelah pelaksanaan kegiatan pembelajaran untuk mengukur sejauh mana tujuan pembelajaran tercapai (Amarila et al., 2021). Berdasarkan Pasal 1 Ayat 1 Peraturan Pemerintah RI Nomor 19 Tahun 2017 tentang Perubahan atas Peraturan Pemerintah Nomor 74 Tahun 2008 tentang Guru, disebutkan bahwa guru adalah pendidik profesional yang memiliki tanggung jawab utama dalam mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada berbagai tingkatan pendidikan. Dengan mengacu pada peraturan ini, diketahui bahwa salah satu tugas guru adalah melakukan evaluasi dan penilaian terhadap peserta didik (Mukarromah, 2018). Evaluasi dilakukan untuk memastikan hasil yang telah dicapai oleh peserta didik setelah mengikuti proses pembelajaran (Kundu et al., 2020). Melalui kegiatan evaluasi, guru dapat mengevaluasi efektivitas komponen pembelajaran yang telah diberikan.

Evaluasi umumnya dilakukan dengan menggunakan instrumen tes, baik yang disusun sendiri maupun yang berasal dari berbagai buku sumber (Dewi Widyaningrum & Sulistya Wardani, 2020). Penggunaan instrumen tes dianggap efektif dalam menginterpretasikan tingkat pemahaman peserta didik terkait materi yang diajarkan. Menurut Mardapi (Riyani et al., 2017), tes merupakan salah satu metode untuk mengestimasi kemampuan seseorang secara tidak langsung melalui respons terhadap stimulus atau pertanyaan. Tes dalam konteks pembelajaran digunakan untuk menilai kemampuan individu dan biasanya diberikan oleh guru dalam bentuk penilaian harian, penilaian tengah semester, dan penilaian akhir semester (Warju et al., 2020). Pemberian tes pada pertengahan semester bertujuan untuk mengukur tingkat penguasaan materi yang telah dipelajari selama setengah semester. Pada penilaian tengah semester, umumnya digunakan jenis tes objektif dan subjektif. Penting bagi soal-soal dalam penilaian tengah semester untuk secara akurat mengukur penguasaan materi selama setengah

semester. Oleh karena itu, guru perlu melakukan analisis butir soal untuk memastikan kualitas soal yang diberikan kepada peserta didik.

Daryanto (Mahmudah et al., 2016) menyatakan bahwa tujuan dari analisis butir soal adalah untuk mengidentifikasi kualitas soal, baik itu baik, kurang baik, atau jelek. Hasil analisis tersebut akan menentukan apakah sebuah butir soal layak untuk dimasukkan ke dalam bank soal, direvisi, atau dihapus. Oleh karena itu, sebagai seorang guru, penting untuk dapat membuat soal yang baik dengan melakukan analisis sebelum dan setelah soal diujikan kepada peserta didik. Sebuah soal dapat dianggap baik jika memenuhi karakteristik penilaian butir soal, seperti validitas, reliabilitas, tingkat kesulitan, daya pembeda, dan efektivitas pilihan-pilihan jawaban yang mengelabui (Sari et al., 2020). Validitas dinyatakan apabila soal dapat mengukur tujuan yang ingin dicapai dari tes dengan akurat.

Selain itu, reliabilitas juga merupakan karakteristik penting yang harus dimiliki oleh soal (Oktarina & Armariena, 2020). Reliabilitas mengacu pada kemampuan soal untuk tetap konsisten dalam tingkat kesulitan saat diujikan secara berulang pada waktu yang berbeda. Analisis soal yang dilakukan dengan menggunakan daya pembeda akan membantu dalam menentukan kualitas soal, apakah itu jelek, cukup, baik, atau sangat baik.

Dalam melakukan analisis soal, terdapat dua pendekatan yang dapat digunakan, yaitu pendekatan kualitatif dan pendekatan kuantitatif (Permatasari & Ahmad, 2022). Pendekatan kualitatif dalam analisis soal melibatkan pemeriksaan teliti terhadap soal-soal dengan mempertimbangkan kompetensi dasar, indikator pembelajaran, serta aspek materi, konstruksi, dan bahasa sebelum soal tersebut diberikan kepada peserta didik (Gomboo et al., 2019). Sementara itu, pendekatan kuantitatif dalam analisis soal dilakukan setelah soal diberikan kepada peserta didik, di mana jawaban peserta didik dianalisis menggunakan perangkat lunak komputer untuk mengevaluasi kualitas butir soal dengan lebih mudah dalam perhitungan.

Berdasarkan penelitian di Sekolah Dasar Negeri Gladak Anyar 4 Pamekasan pada tanggal 22 Desember 2022, ditemukan bahwa soal Penilaian Tengah Semester biasanya terdiri dari soal pilihan ganda dan soal isian (essay). Soal tersebut dibuat oleh tim pembuat soal Penilaian Tengah Semester yang terdiri dari guru-guru kelas V SDN Gladak Anyar 4 Pamekasan. Namun, guru belum melakukan analisis butir soal secara menyeluruh dari hasil Penilaian Tengah Semester peserta didik. Analisis yang dilakukan hanya berfokus pada tingkat kesulitan soal. Padahal, analisis butir soal seharusnya melibatkan pertimbangan aspek lain seperti validitas, reliabilitas, daya pembeda, dan keefektifan pengecoh. Guru menghadapi kendala dalam menggunakan perangkat lunak komputer untuk menganalisis butir soal secara menyeluruh. Oleh karena itu, peneliti melakukan penelitian dengan judul "Analisis Butir Soal Pilihan Ganda Penilaian Tengah Semester Pada Pembelajaran Tematik Kelas V Sekolah Dasar Negeri Gladak Anyar 4 Pamekasan" berdasarkan uraian di atas.

## **METODE**

Penelitian ini menggunakan pendekatan kuantitatif deskriptif (Yue & Xu, 2019). Penelitian kuantitatif melibatkan pengumpulan data dan analisis data yang dilakukan dengan prinsip-prinsip objektivitas, serta penyajian data dalam bentuk gambar, tabel, grafik, atau bentuk lainnya (Ujang Charda S, 2021). Fokus penelitian ini adalah menganalisis butir soal menggunakan Microsoft Excel 2010 sebagai alat pengukur kemampuan guru kelas V dalam membuat soal untuk peserta didik. Metode yang digunakan dalam penelitian ini adalah metode deskriptif. Peneliti bertanggung jawab sebagai perencana penelitian, pelaksana penelitian, pengumpul data, analisis data, dan pelaporan hasil penelitian. Data yang dikumpulkan berupa penilaian tengah semester peserta didik dari kelas VA, VB, dan VC di SDN Gladak Anyar 4 Pamekasan. Partisipan dalam penelitian ini adalah analisis butir soal pilihan ganda penilaian tengah semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan. Penelitian ini dilaksanakan pada semester kedua tahun ajaran 2021/2022.

Populasi penelitian ini terdiri dari 100 peserta didik, yang terdiri dari 34 peserta didik dari kelas VA, 33 peserta didik dari kelas VB, dan 33 peserta didik dari kelas VC di SDN Gladak Anyar 4 Pamekasan pada tahun ajaran 2021/2022. Pengambilan sampel dilakukan dengan menggunakan teknik simple random sampling, yaitu secara acak tanpa memprioritaskan level dalam populasi. Sampel yang digunakan adalah soal Penilaian Tengah Semester yang dijawab oleh 100 peserta didik dari kelas VA, VB, dan VC.

Teknik pengumpulan data yang digunakan dalam penelitian ini adalah dokumentasi. Dokumentasi dilakukan dengan mengumpulkan perangkat soal penilaian tengah semester genap kelas V tahun ajaran 2021/2022 di SDN Gladak Anyar 4 Pamekasan, termasuk lembar soal Penilaian Tengah Semester, kisi-kisi soal, dan kunci jawaban. Selain itu, juga dikumpulkan jawaban dari seluruh peserta didik yang terdiri dari kelas VA, VB, dan VC.

Analisis data dalam penelitian ini menggunakan teknik analisis kuantitatif (Fauzi & Pradipta, 2018). Dalam analisis, skor benar diberi nilai 1, sedangkan skor salah diberi nilai 0 (dikotomi). Peneliti menggunakan aplikasi Microsoft Excel 2010 untuk menganalisis data dan mencari tingkat validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dalam butir soal.

## **HASIL DAN PEMBAHASAN**

### ***Hasil***

Hasil analisis butir soal dari Penilaian Tengah Semester pada pembelajaran tematik kelas V SDN Gladak Anyar 4 Pamekasan melibatkan 39 butir soal yang terbagi menjadi 2 tema. Tema 6 terdiri dari 20 butir soal, sedangkan tema 7 terdiri dari 19 butir soal. Selain itu, juga melibatkan 100 lembar jawaban dari peserta didik dan kunci jawaban soal. Analisis butir soal dilakukan dengan mempertimbangkan beberapa aspek sebagai berikut:

#### **1. Validasi**

Pengujian validitas dilakukan dengan menghitung korelasi point biserial menggunakan Microsoft Excel 2010. Hasil perhitungan kemudian dibandingkan dengan nilai kritis pada tabel harga kritik  $r$  product moment dengan tingkat signifikansi 5%. Jumlah subjek dalam penelitian ini adalah 100 peserta didik, dan terdapat 2 variabel yang dikorelasikan, sehingga derajat kebebasan (df atau db) adalah 98 (100-2). Dengan menggunakan tingkat signifikansi 5%, nilai kritis pada tabel harga kritik adalah 0,195. Jika nilai korelasi hitung ( $y_{pbi}$ )  $\geq$  nilai kritis ( $r$  tabel), maka butir soal dikatakan valid. Berdasarkan hasil analisis, pada tema 6 terdapat 19 butir soal atau 95% yang dikatakan valid, sedangkan 1 butir soal atau 5% dikatakan tidak valid. Pada tema 7, terdapat 18 butir soal atau 94,74% yang dikatakan valid, dan 1 butir soal atau 5,26% dikatakan tidak valid. Distribusi soal berdasarkan validitasnya dapat dilihat pada Tabel 1 dan 2.

Tabel 1. Distribusi soal Tema 6 Kelas VA, VB, VC berdasarkan Indeks Validitas

No	Validitas	No Butir	Jumlah	Presentase
1.	$\geq 0,195$ Valid	2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20	19	95%
2.	$< 0,195$ Invalid	1	1	5%

Tabel 2. Distribusi soal Tema 7 Kelas VA, VB

No	Validitas	No Butir	Jumlah	Presentase
1.	$\geq 0,195$ Valid	1, 2, 3, 4, 7, 8, 11, 13, 14, 15, 16, 19, 20, 21, 22, 25, 27, 29	18	94,74%
2.	$< 0,195$ Invalid	10	1	5,26%

## 2. Realibilitas

Hasil perhitungan reliabilitas soal Penilaian Tengah Semester dalam pembelajaran tematik kelas V diinterpretasikan dengan menggunakan patokan bahwa jika nilai  $r_{11} \geq 0,70$ , maka reliabilitas dianggap tinggi. Perhitungan reliabilitas dilakukan menggunakan Microsoft Excel 2010. Hasil perhitungan menunjukkan bahwa soal Penilaian Tengah Semester pada tema 6 memiliki reliabilitas sebesar 0,962237416, sementara soal pada tema 7 memiliki reliabilitas sebesar 0,895019386. Analisis hasil menunjukkan bahwa kedua tema soal tersebut memiliki reliabilitas yang tinggi, mengingat nilai  $r_{11} \geq 0,70$ .

## 3. Tingkat Kesukaran

Dalam interpretasi perhitungan tingkat kesukaran, terdapat 3 kriteria yang digunakan, yaitu: soal dengan rentang P 0,00 - 0,30 dianggap sulit, rentang P 0,31 - 0,70 dianggap sedang, dan rentang P 0,71 - 1,00 dianggap mudah. Berdasarkan perhitungan tingkat kesukaran menggunakan Microsoft Excel 2010, pada tema 6 soal Penilaian Tengah Semester terdapat 2 butir soal atau 10,53% yang tergolong sulit, 4 butir soal atau 21,05% tergolong sedang, dan 13 butir soal atau 68,42% tergolong mudah. Sementara itu, pada tema 7 tidak ada butir soal yang tergolong sulit, terdapat 7 butir soal atau 38,89% tergolong sedang, dan 11 butir soal atau 61,11% tergolong mudah. Berikut ini adalah distribusi soal berdasarkan tingkat kesukarannya.

Tabel 3. Distribusi soal Tema 6 Kelas VA, VB, VC berdasarkan Tingkat Kesukaran

No	Tingkat Kesukaran	No Butir	Jumlah	Presentase
1.	Mudah	3, 5, 6, 9, 10, 12, 14, 15, 16, 17, 18, 19, 20	13	68,42%
2.	Sedang	2, 4, 8, 13	4	21,05%
3.	Sukar	7, 11	2	10,53%

Tabel 4. Distribusi soal Tema7 Kelas VA, VB, VC berdasarkan Tingkat Kesukaran

No	Tingkat Kesukaran	No Butir	Jumlah	Presentase
1.	Mudah	1, 3, 4, 7, 11, 13, 14, 16, 19, 21, 22	11	61,11%
2.	Sedang	2, 8, 15, 20, 25, 27, 29	7	38,89%
3.	Sukar		0	0%

#### 4. Daya Pembeda

Hasil perhitungan daya pembeda diberikan penafsiran berdasarkan empat kriteria sebagai berikut: jika nilai D berada dalam rentang 0,00 - 0,20, maka soal dikategorikan sebagai jelek. Rentang 0,21 - 0,40 menunjukkan bahwa soal cukup baik. Jika nilai D berada dalam rentang 0,41 - 0,70, maka soal dikategorikan sebagai baik. Rentang 0,71 - 1,00 menunjukkan bahwa soal dikategorikan sebagai baik sekali. Jika nilai D bernilai negatif, maka soal dikategorikan sebagai tidak baik.

Dari perhitungan menggunakan Microsoft Excel 2010, pada tema 6 penilaian tengah semester terdapat 10 butir soal atau 52,63% yang memiliki daya pembeda jelek, 8 butir soal atau 41,11% memiliki daya pembeda cukup, 1 butir soal atau 5,26% memiliki daya pembeda baik, dan tidak terdapat soal dengan daya pembeda baik sekali. Sementara itu, pada tema 7 terdapat 6 butir soal atau 33,33% dengan daya pembeda jelek, 9 butir soal atau 50% dengan daya pembeda cukup, 3 butir soal atau 16,67% dengan daya pembeda baik, dan tidak terdapat soal dengan daya pembeda baik sekali. Berikut ini adalah distribusi soal berdasarkan daya pembedanya.

Tabel 5. Distribusi soal Tema 6 Kelas VA, VB, VC berdasarkan Daya Pembeda

No	Daya Pembeda	Nomor Butir	Jumlah	Presentase
1.	Jelek	3, 6, 7, 9, 12, 14, 15, 17, 18, 19	10	52,63%
2.	Cukup	2, 4, 5, 8, 10, 11, 16, 20	8	42,11%
3.	Baik	13	1	5,26%
4.	Baik Sekali		0	0%

Tabel 6. Distribusi soal Tema 7 Kelas VA, VB, VC berdasarkan Daya Pembeda

No	Daya Pembeda	Nomor Butir	Jumlah	Presentase
1.	Jelek	1, 3, 7, 13, 16, 21	6	33,33%
2.	Cukup	4, 8, 11, 14, 19, 22, 25, 27, 29	8	50%
3.	Baik	2, 15, 20	3	16,67%
4.	Baik Sekali		0	0%

#### 5. Efektivitas Pengecoh

Efektivitas pengecoh dalam suatu soal dapat dikategorikan baik jika dipilih oleh peserta didik sebesar atau lebih dari 5%. Perhitungan efektivitas pengecoh dilakukan menggunakan Microsoft

Excel 2010. Penafsiran efektivitas pengecoh pada setiap butir soal didasarkan pada kriteria Skala Likert sebagai berikut:

- Jika keempat jawaban pengecoh dalam soal berfungsi dengan baik, maka efektivitas pengecoh dikategorikan sebagai sangat baik.
- Jika tiga jawaban pengecoh dalam soal berfungsi dengan baik, maka efektivitas pengecoh dikategorikan sebagai baik.
- Jika dua jawaban pengecoh dalam soal berfungsi dengan baik, maka efektivitas pengecoh dikategorikan sebagai cukup.
- Jika hanya satu jawaban pengecoh dalam soal yang berfungsi dengan baik, maka efektivitas pengecoh dikategorikan sebagai kurang baik.
- Jika soal tidak memiliki jawaban pengecoh, maka efektivitas pengecoh dikategorikan sebagai sangat tidak baik.

Dari hasil perhitungan menggunakan Microsoft Excel 2010, pada tema 6 penilaian tengah semester terdapat 1 butir soal atau 5,26% yang memiliki efektivitas pengecoh sangat baik, 8 butir soal atau 42,11% memiliki efektivitas pengecoh baik, 4 butir soal atau 21,05% memiliki efektivitas pengecoh cukup baik, 6 butir soal atau 31,58% memiliki efektivitas pengecoh kurang baik, dan tidak ditemukan soal yang memiliki efektivitas pengecoh tidak baik. Sedangkan pada tema 7, terdapat 4 butir soal atau 22,22% yang memiliki efektivitas pengecoh sangat baik, 4 butir soal atau 22,22% memiliki efektivitas pengecoh baik, 7 butir soal atau 38,89% memiliki efektivitas pengecoh cukup baik, 3 butir soal atau 16,67% memiliki efektivitas pengecoh kurang baik, dan tidak ditemukan soal yang memiliki efektivitas pengecoh tidak baik. Informasi lebih detail mengenai distribusi efektivitas pengecoh dapat dilihat pada tabel berikut.

Tabel 7. Distribusi soal Tema 6 Kelas VA, VB, VC berdasarkan Efektivitas Pengecoh

No	Efektivitas Pengecoh	No Butir	Jumlah	Presentase
1.	Sangat Baik (4)	8	1	5,26%
2.	Baik (3)	2, 4, 5, 7, 10, 11, 13, 18	8	42,11%
3.	Cukup Baik (2)	3, 6, 16, 20	4	21,05%
4.	Kurang Baik (1)	9, 12, 14, 15, 17 19	6	31,58%
5.	Tidak Baik (0)		0	0%

Tabel 8. Distribusi soal Tema 7 Kelas VA, VB, VC berdasarkan Efektivitas Pengecoh

No	Efektivitas Pengecoh	No Butir	Jumlah	Presentase
1.	Sangat Baik (4)	2, 8, 15, 29	4	22,22%
2.	Baik (3)	4, 11, 20, 27	4	22,22%
3.	Cukup (2)	7, 13, 14, 19, 21, 22, 25	7	38,89%
4.	Kurang Baik (1)	1, 3, 16	3	16,67%
5.	Tidak Baik (0)		0	0%

## 6. Kualitas Soal

Kualitas sebuah soal dapat dinilai berdasarkan kriteria dalam Skala Likert berikut ini: (1) Jika sebuah butir soal memenuhi 4 kriteria dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut memiliki kualitas sangat baik dan layak untuk dimasukkan ke dalam bank soal. (2) Jika sebuah butir soal memenuhi 3 kriteria dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut memiliki kualitas baik, namun belum sepenuhnya dapat dimasukkan ke dalam bank soal. (3) Jika sebuah butir soal memenuhi 2 kriteria dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut memiliki kualitas cukup dan perlu direvisi kembali. (4) Jika sebuah butir soal hanya memenuhi 1 kriteria dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut memiliki kualitas buruk dan perlu direvisi secara terperinci atau sebaiknya dibuang. (5) Jika tidak ada butir soal yang memenuhi kriteria dari segi validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh, maka soal tersebut memiliki kualitas sangat buruk dan perlu direvisi secara terperinci atau sebaiknya dibuang. (6) Secara keseluruhan, tes harus memiliki reliabilitas sesuai dengan ketentuan yang berlaku.

Berdasarkan analisis soal pada tema 6, terdapat 4 butir soal atau 20% yang memiliki kualitas sangat baik, 5 butir soal atau 25% yang berkualitas baik, 3 butir soal atau 15% yang berkualitas sedang, 7 butir soal atau 35% yang berkualitas tidak baik, dan 1 butir soal atau 5% yang berkualitas sangat tidak baik. Sementara itu, pada tema 7 terdapat 6 butir soal atau 31,58% yang berkualitas sangat baik, 6 butir soal atau 31,58% yang berkualitas baik, 3 butir soal atau 15,79% yang berkualitas sedang, 3 butir soal atau 15,79% yang berkualitas tidak baik, dan 1 butir soal atau 5,26% yang berkualitas sangat tidak baik. Informasi mengenai distribusi butir soal berdasarkan kualitasnya dapat ditemukan pada tabel 9 dan 10.

Tabel 9. Distribusi butir soal Tema 6 PTS berdasarkan Kualitas Butir Soal

No	Kualitas Butir Soal	No Butir	Jumlah	Presentase
1.	Sangat Baik (4)	2, 4, 8, 13	4	20%
2.	Baik (3)	5, 10, 11, 16, 20	5	25%
3.	Sedang (2)	3, 6, 18	3	15%
4.	Tidak Baik (1)	7, 9, 12, 14, 15, 17, 19	7	35%
5.	Sangat Tidak Baik (0)		1	5%

Tabel 10. Distribusi butir soal Tema 7 PTS berdasarkan Kualitas Butir Soal



No	Kualitas Butir Soal	No Butir	Jumlah	Presentase
1.	Sangat Baik (4)	8, 15, 20, 25, 27, 29	6	31,58%
2.	Baik (3)	2, 4, 11, 14, 19, 22	6	31,58%
3.	Sedang (2)	7, 13, 21	3	15,79%
4.	Tidak Baik (1)	1, 3, 16	3	15,79%
5.	Sangat Tidak Baik (0)	10	1	5,26%

### Diskusi

#### 1. Tingkat Validitas

Hasil penelitian menunjukkan bahwa validitas butir soal Penilaian Tengah Semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan menunjukkan bahwa pada tema 6 terdapat 95% (19 butir) soal yang masuk dalam kategori "valid" dan 5% (1 butir) soal yang masuk dalam kategori "invalid". Sedangkan pada tema 7, terdapat 94,74% (18 butir) soal yang masuk dalam kategori "valid" dan 5,26% (1 butir) soal yang masuk dalam kategori "invalid". Berdasarkan tabel 1, pada tema 6, butir nomor 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, dan 20 termasuk dalam kategori "valid", yang berarti hasil perhitungan  $Y_{pbi} > r_{tabel}$ . Namun, butir nomor 1 termasuk dalam kategori "invalid", yang berarti hasil perhitungan  $Y_{pbi} > r_{tabel}$ . Sedangkan berdasarkan tabel 2, pada tema 7, butir nomor 1, 2, 3, 4, 7, 8, 11, 13, 14, 15, 16, 19, 20, 21, 22, 25, 27, dan 29 termasuk dalam kategori "valid", yang berarti hasil perhitungan  $Y_{pbi} > r_{tabel}$ . Namun, butir nomor 10 termasuk dalam kategori "invalid", yang berarti hasil perhitungan  $Y_{pbi} > r_{tabel}$ .

Berdasarkan paparan di atas, dapat disimpulkan bahwa butir soal Penilaian Tengah Semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan pada tema 6 dan 7 memiliki kualitas soal yang baik dari segi validitas. Hal ini dikarenakan 19 dari 20 soal pada tema 6 dinyatakan valid dan 18 dari 19 soal pada tema 7 dinyatakan valid. Oleh karena itu, dapat dikatakan bahwa butir soal Penilaian Tengah Semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan dapat mengukur apa yang seharusnya diukur, sesuai dengan pendapat Sudijono (2010, h.128). Hal ini juga sejalan dengan pendapat Bambang Subali (2010, h.107) yang menyatakan bahwa alat ukur dikatakan valid jika mampu menyampaikan informasi empirik dengan tepat sesuai dengan yang diukur.

#### 2. Realibilitas

Dari hasil penelitian, terlihat bahwa reliabilitas butir soal Penilaian Tengah Semester pada pembelajaran tematik kelas V di SDN Gladak Anyar 4 Pamekasan, pada tema 6 memiliki indeks reliabilitas yang tinggi atau reliabel, yaitu sebesar 0,959267547. Setelah melakukan penghapusan butir soal yang invalid, dilakukan analisis ulang dan ditemukan hasil indeks reliabilitas sebesar 0,867390953. Sementara itu, pada tema 7, soal-soal memiliki indeks reliabilitas yang tinggi atau reliabel, yaitu sebesar 0,8950119386. Setelah dilakukan penghapusan butir soal yang invalid dan

dilakukan analisis ulang, ditemukan hasil indeks reliabilitas sebesar 0,895610913. Hal ini menunjukkan bahwa nilai indeks reliabilitas pada tema 6 mengalami penurunan sebesar 0,09, sementara nilai indeks reliabilitas pada tema 7 mengalami peningkatan sebesar 0,0005989744. Hasil ini menunjukkan bahwa soal-soal tersebut mampu mengukur dengan konsisten, yang merupakan indikasi reliabilitas. Secara prinsip, reliabilitas digunakan untuk menunjukkan kemampuan pengukuran yang hasilnya relatif tidak berubah ketika pengukuran dilakukan kembali terhadap subjek yang sama (Mardapi, 2008, h.58).

### 3. Tingkat Kesukaran

Berdasarkan hasil penelitian, ditemukan bahwa pada tema 6, sebanyak 68,42% (13 butir) soal dikategorikan sebagai "mudah", 21,05% (4 butir) dikategorikan sebagai "sedang", dan 10,53% (2 butir) dikategorikan sebagai "sukar". Sementara itu, pada tema 7, sebanyak 61,11% (11 butir) soal dikategorikan sebagai "mudah", dan 38,89% (7 butir) dikategorikan sebagai "sedang".

Berdasarkan tabel 3 pada tema 6, ditemukan bahwa butir nomor 3, 5, 6, 9, 10, 12, 14, 15, 16, 17, 18, 19, 20 dikategorikan sebagai soal "mudah", yang berarti hampir semua peserta didik dapat menjawab dengan benar. Butir nomor 2, 4, dan 8, 13 dikategorikan sebagai soal dengan tingkat kesukaran "sedang", yang berarti tidak semua peserta didik dapat menjawab dengan benar. Sedangkan, butir nomor 7 dan 11 dikategorikan sebagai soal "sukar", yang berarti hanya beberapa peserta didik yang dapat menjawab dengan benar.

Dari tabel 4 pada tema 7, ditemukan bahwa butir nomor 1, 3, 4, 7, 11, 13, 14, 16, 19, 21, dan 22 dikategorikan sebagai soal "mudah", yang berarti hampir semua peserta didik dapat menjawab dengan benar. Butir nomor 2, 8, 15, 20, 25, 27, dan 29 memiliki tingkat kesukaran "sedang", yang berarti tidak semua peserta didik dapat menjawab dengan benar.

Secara keseluruhan, tematik pada kelas V memiliki tingkat kesukaran "mudah" pada tema 6 dan 7. Namun, disarankan untuk menggunakan soal dengan tingkat kesukaran "sedang" saat membuat Penilaian Tengah Semester.

Berdasarkan data yang telah disajikan, perlu dilakukan penelitian dan revisi terperinci pada soal dengan tingkat kesukaran "mudah" dan "sukar", atau bahkan mempertimbangkan untuk menghapusnya jika soal tersebut tidak memenuhi kriteria yang baik. Sementara itu, untuk soal dengan tingkat kesukaran "sedang", dapat disimpan dalam bank soal dan digunakan sebagai soal tes di masa yang akan datang (Sudijono, 2012, h.376-378). Pendapat ini sejalan dengan Arikunto (2016, h.222), yang menyatakan bahwa soal yang baik adalah soal yang tidak terlalu mudah maupun terlalu sulit, karena soal yang terlalu mudah tidak dapat membangkitkan kemampuan peserta didik dalam memecahkan masalah, sedangkan soal yang terlalu sulit dapat meruntuhkan semangat peserta didik dan membuat mereka merasa tidak mampu.

### 4. Daya Pembeda

Hasil dari penelitian menunjukkan bahwa daya pembeda soal pada Penilaian Tengah Semester tema 6 memiliki kategori "jelek" sebesar 52,63% (10 butir), kategori "cukup" sebesar

42,11% (8 butir), kategori "baik" sebesar 5,26% (1 butir), dan tidak terdapat soal yang termasuk dalam kategori "baik sekali" (0 butir). Sementara itu, pada tema 7, daya pembeda soal memiliki kategori "jelek" sebesar 33,33% (6 butir), kategori "cukup" sebesar 50% (8 butir), kategori "baik" sebesar 16,67% (3 butir), dan tidak terdapat soal yang termasuk dalam kategori "baik sekali" (0 butir).

Dari data yang terdapat pada tabel 5, dapat dilihat bahwa butir nomor 3, 6, 7, 9, 12, 14, 15, 17, 18, 19 pada tema 6 termasuk dalam kategori "jelek", yang berarti bahwa soal-soal tersebut tidak efektif dalam membedakan antara peserta didik yang memiliki kemampuan tinggi dan rendah. Sedangkan butir nomor 2, 4, 5, 8, 10, 11, 16, 20 masuk dalam kategori "cukup", yang berarti soal-soal tersebut mampu memberikan sedikit perbedaan antara peserta didik yang memiliki kemampuan tinggi dan rendah. Satu-satunya butir nomor 13 termasuk dalam kategori "baik", yang berarti butir soal tersebut memberikan sedikit perbedaan antara peserta didik yang memiliki kemampuan tinggi dan rendah.

Sementara itu, berdasarkan tabel 6 pada tema 7, butir nomor 1, 3, 7, 13, 16, 21 termasuk dalam kategori "jelek", yang berarti soal-soal tersebut tidak efektif dalam membedakan antara peserta didik yang memiliki kemampuan tinggi dan rendah. Sedangkan butir nomor 4, 8, 11, 14, 19, 22, 25, 27, 29 masuk dalam kategori "cukup", yang berarti soal-soal tersebut mampu memberikan sedikit perbedaan antara peserta didik yang memiliki kemampuan tinggi dan rendah. Terakhir, butir nomor 2, 15, 20 termasuk dalam kategori "baik", yang berarti butir soal tersebut memberikan sedikit perbedaan antara peserta didik yang memiliki kemampuan tinggi dan rendah.

Menurut Arikunto (2012, h.226), pembeda soal merujuk pada kemampuan soal dalam memisahkan peserta didik berdasarkan tingkat kemampuan mereka (baik atau kurang pandai). Analisis daya pembeda dilakukan untuk memahami sejauh mana soal dapat membedakan antara peserta didik yang memiliki kemampuan akademik tinggi dan rendah. Analisis daya pembeda ini diperoleh melalui perhitungan indeks diskriminasi soal tes. Dari hasil analisis tersebut, terdapat dua kemungkinan untuk soal dengan daya pembeda yang jelek, yaitu melakukan penelusuran untuk diperbaiki atau menghapusnya. Namun, jika indeks daya pembeda negatif, maka soal tersebut harus dihapus. Sementara itu, soal dengan kategori cukup, baik, dan baik sekali dapat disimpan dalam bank soal dan digunakan kembali untuk tes yang akan datang.

##### 5. Efektivitas Pengecoh

Dari hasil penelitian, ditemukan efektivitas pengecoh pada soal Penilaian Tengah Semester tema 6 dengan kategori "sangat baik" sebesar 5,26% (1 butir), kategori "baik" sebesar 42,11% (8 butir), kategori "cukup baik" sebesar 21,05% (4 butir), kategori "kurang baik" sebesar 31,58% (6 butir), dan kategori "tidak baik" sebesar 0% (0 butir). Sementara itu, efektivitas pengecoh pada soal tema 7 memiliki kategori "sangat baik" sebesar 22,22% (4 butir), kategori "baik" sebesar 22,22% (4 butir), kategori "cukup baik" sebesar 38,89% (7 butir), kategori "kurang baik" sebesar 16,67% (3 butir), dan kategori "tidak baik" sebesar 0% (0 butir).

Berdasarkan tabel 7 tema 6, butir nomor 8 dikategorikan sebagai "sangat baik," yang berarti keempat pilihan jawaban berfungsi dengan sangat baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 2, 4, 5, 7, 10, 11, 13, 18 dikategorikan sebagai "baik," yang berarti ketiga pilihan jawaban berfungsi dengan baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 3, 6, 16, 20 dikategorikan sebagai "cukup baik," yang berarti dua pilihan jawaban berfungsi cukup baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 9, 12, 14, 15, 17, 19 dikategorikan sebagai "kurang baik," yang berarti satu pilihan jawaban berfungsi kurang baik dengan persentase setiap pilihan  $\geq 5\%$ .

Sedangkan berdasarkan tabel 8 tema 7, butir nomor 2, 8, 15, 29 dikategorikan sebagai "sangat baik," yang berarti keempat pilihan jawaban berfungsi dengan sangat baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 4, 11, 20, 27 dikategorikan sebagai "baik," yang berarti ketiga pilihan jawaban berfungsi dengan baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 7, 13, 14, 19, 21, 22, 25 dikategorikan sebagai "cukup baik," yang berarti dua pilihan jawaban berfungsi cukup baik dengan persentase setiap pilihan  $\geq 5\%$ . Butir nomor 1, 3, 16 dikategorikan sebagai "kurang baik," yang berarti satu pilihan jawaban berfungsi kurang baik dengan persentase setiap pilihan  $\geq 5\%$ .

Efektivitas pengecoh soal tes mengacu pada kemampuan soal dalam mengecoh peserta didik dalam menjawab soal tes yang diberikan. Pandangan ini sejalan dengan pendapat Sudijono (2012, h.411) yang menyatakan bahwa pengecoh dikatakan baik jika pengecoh tersebut dipilih oleh setidaknya 5% dari jumlah peserta tes.

#### 6. Kualitas Butir soal

Untuk memahami kualitas butir soal, kita dapat melihat hasil analisis validitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Kualitas butir soal dapat diklasifikasikan menjadi lima kategori, yaitu sangat baik, baik, sedang, tidak baik, dan sangat tidak baik, berdasarkan Skala Likert.

Dalam analisis soal tema 6, terdapat 4 butir soal atau 20% yang memiliki kualitas sangat baik, 5 butir soal atau 25% dengan kualitas baik, 3 butir soal atau 15% dengan kualitas sedang, 7 butir soal atau 35% dengan kualitas tidak baik, dan 1 butir soal atau 5% dengan kualitas sangat tidak baik. Sementara itu, dalam analisis soal tema 7, terdapat 6 butir soal atau 31,58% dengan kualitas sangat baik, 6 butir soal atau 31,58% dengan kualitas baik, 3 butir soal atau 15,79% dengan kualitas sedang, 3 butir soal atau 15,79% dengan kualitas tidak baik, dan 1 butir soal atau 5,26% dengan kualitas sangat tidak baik.

Butir soal yang memiliki kualitas sangat baik harus disimpan di bank soal dengan menjaga kerahasiaannya agar dapat digunakan kembali dalam ujian mendatang. Butir soal dengan kualitas baik dan sedang perlu direvisi sesuai dengan indikatornya sebelum dapat dimasukkan ke dalam bank soal. Sementara itu, butir soal dengan kualitas tidak baik dan sangat tidak baik sebaiknya dihapus karena memerlukan revisi yang lebih mendalam.

Berdasarkan data analisis soal tema 6, terdapat 8 butir soal yang perlu direvisi dan 8 butir soal yang perlu dihapus. Sedangkan dalam hasil analisis soal tema 7, terdapat 9 butir soal yang perlu direvisi dan 4 butir soal yang lebih baik dihapus. Penyebab kegagalan butir soal dapat dilihat kembali

berdasarkan aspek validitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh pada masing-masing butir soal. Informasi lebih lanjut mengenai penyebab kegagalan butir soal dapat ditemukan pada tabel 8 dan 9.

Tabel 11. Penyebab kegagalan butir soal tema 6

No	Penyebab Kegagalan	Butir Soal	Jumlah	Presentase
1.	Validitas (Invalid)	1	1	5%
2.	Tingkat Kesukaran (Mudah dan Sukar)	3, 5, 6, 7, 9, 10, 11, 12, 14, 15, 16, 17, 18, 19, 20	15	78,95%
3.	Daya Pembeda (Jelek dan Negatif)	3, 6, 7, 9, 12, 14, 15, 17, 18, 19	10	52,63%
4.	Efektivitas Pengecoh (Kurang Baik dan Tidak Baik)	9, 12, 14, 15, 17, 19	6	31,58%

Tabel 12. Penyebab kegagalan butir soal tema 7

No	Penyebab Kegagalan	Butir Soal	Jumlah	Presentase
1.	Validitas (Invalid)	10	1	5,26%
2.	Tingkat Kesukaran (Mudah dan Sukar)	1, 3, 4, 7, 11, 13, 14, 16, 19, 21, 22	11	61,11%
3.	Daya Pembeda (Jelek dan Negatif)	1, 3, 7, 13, 16, 21	6	33,33%
4.	Efektivitas Pengecoh (Kurang Baik dan Tidak Baik)	1, 3, 16	3	16,67%

## KESIMPULAN

Berdasarkan hasil analisis soal pilihan ganda Penilaian Tengah Semester yang menunjukkan kualitas baik, dapat disimpulkan bahwa: (1) Soal penilaian tengah semester memiliki kualitas baik dalam hal validitas. Hal ini dibuktikan melalui hasil validasi, di mana butir soal pada tema 6 dan 7 memiliki indeks validitas yang valid, yaitu 95% (19 butir) dan 94,74% (18 butir). (2) Soal penilaian tengah semester memiliki kualitas baik dalam hal reliabilitas. Hal ini terbukti dari reliabilitas yang tinggi pada soal tema 6 dan 7, dengan reliabilitas sebesar 0,959267547 dan 0,8950119386 secara berturut-turut. (3) Soal penilaian tengah semester belum memiliki kualitas baik dalam hal tingkat kesukaran. Tingkat kesukaran soal pada tema 6 terdiri dari kategori "mudah" sebesar 68,42% (13 butir), "sedang" sebesar 21,05% (4 butir), dan "sukar" sebesar 10,53% (2 butir). Pada tema 7, tingkat kesukaran terdiri dari kategori "mudah" sebesar 61,11% (11 butir) dan "sedang" sebesar 38,89% (7 butir), menunjukkan bahwa soal cenderung mudah. (4) Daya pembeda soal pada tema 6 terdiri dari kategori "jelek" sebesar 52,63% (10 butir), "cukup" sebesar 42,11% (8 butir), dan "baik" sebesar 5,26% (1 butir). Pada tema 7, daya pembeda terdiri dari kategori "jelek" sebesar 33,33% (6 butir), "cukup" sebesar 50% (8 butir), dan "baik" sebesar 16,67% (3 butir). Dengan demikian, soal termasuk dalam kategori cukup baik dalam hal daya pembeda. (5) Soal penilaian tengah semester memiliki kualitas baik dalam hal efektivitas pengecoh. Hal ini terlihat dari efektivitas pengecoh pada tema 6 yang terdiri dari kategori "sangat baik" sebesar 5,26% (1 butir), "baik" sebesar 42,11% (8 butir),

"cukup baik" sebesar 21,05% (4 butir), dan "kurang baik" sebesar 31,58% (6 butir). Pada tema 7, efektivitas pengecoh terdiri dari kategori "sangat baik" sebesar 22,22% (4 butir), "baik" sebesar 22,22% (4 butir), "cukup baik" sebesar 38,89% (7 butir), dan "kurang baik" sebesar 16,67% (3 butir). Dengan demikian, soal termasuk dalam kategori baik dalam hal efektivitas pengecoh.

## REFERENSI

- Amarila, R. S., Subali, B., & Saptono, S. (2021). Analisis Kemampuan Berpikir Kritis Siswa Pada Pembelajaran IPA Terpadu Tema Lingkungan. *Jurnal Improvement*, 8(1), 82–91.
- Dewi Widyaningrum, M., & Sulistya Wardani, N. (2020). Efektivitas Pembelajaran Tematik (PS) Terhadap Kemampuan Berpikir Kritis Peserta Didik. *International Journal of Elementary Education*, 4(1), 90. <https://doi.org/10.23887/ijee.v4i1.24385>
- Fauzi, A., & Pradipta, I. W. (2018). Research methods and data analysis techniques in education articles published by Indonesian biology educational journals. *JPBI (Jurnal Pendidikan Biologi Indonesia)*, 4(2), 123–134. <https://doi.org/10.22219/jpbi.v4i2.5889>
- Gomboo, A., Gombo, B., Munkhgerel, T., Nyamjav, S., & Badamdorj, O. (2019). Item Analysis of Multiple Choice Questions in Medical Licensing Examination. *Cent Asian J Med Sci*, 5(2), 141–148.
- Kundu, S., Ughade, J. M., Sherke, A. R., Kanwar, Y., Tiwari, S., Jatwar, R., Gurudiwan, R., & Kundu, S. G. (2020). Impact Measurement on Medical Faculty for Adhering to Appropriate Guidelines in Framing Effective Multiple-Choice Questions for Item Analysis. *Journal of Medical Education*, 19(1), 1–15. <https://doi.org/10.5812/jme.103482>
- Mahmudah, R., Pramudya, Y., & Sulisworo, D. (2016). Analisis Validitas Butir Soal Certainty of Respons Index (Cri) Untuk Identifikasi Miskonsepsi Materi Tata Surya Dan Fenomena Astronomi. In *Seminar Nasional Pendidikan Sains 2016* (pp. 1–15).
- Maulidiyah, F., Widarti, H. R., & Utomo, Y. (2021). Analisis Kesalahan Konsep Mahasiswa Kimia Pada Kajian Pokok Hidrolisis Garam Menggunakan Tes Pilihan Ganda Empat Tingkat. *Jurnal Pendidikan: Teori, Penelitian, Dan Pengembangan*, 6(7), 1044. <https://doi.org/10.17977/jptpp.v6i7.14921>
- Mukarromah, A. (2018). Analisis Kemampuan Berpikir Kritis Pada Model Discovery Learning Berdasarkan Pembelajaran Tematik. *Indonesian Journal of Primary Education*, 2(1), 38. <https://doi.org/10.17509/ijpe.v2i1.11844>
- Oktarina, N., & Armariena, D. N. (2020). Analisis Soal Tipe Higher Order Thinking Skills (Hots) Pada Tes Objektif Pilihan Ganda Usbn Bahasa Indonesia Sd/Mi. *Wahana Didaktika : Jurnal Ilmu Kependidikan*, 18(2), 146. <https://doi.org/10.31851/wahanadidaktika.v18i2.4374>
- Permatasari, D., & Ahmad, M. (2022). Pengaruh Metode Pembelajaran Index Card Match Terhadap Kemampuan Memahami Soal Cerita Matematika Siswa Kelas IV SDN Rambutan 02. *Jurnal Cendekia : Jurnal Pendidikan Matematika*, 6(3), 3192–3202.

<https://doi.org/10.31004/cendekia.v6i3.1543>

- Sari, R. M. M., Luritawaty, I. P., Dahlan, J. A., & Siregar, G. M. A. (2020). ANALISIS SOAL UJIAN SEKOLAH BERSTANDAR NASIONAL MATEMATIKA SEKOLAH MENENGAH PERTAMA Tahun Ajaran 2018/2019. *Jurnal THEOREMS (The Original Research of Mathematics)*, 4(2), 195. <https://doi.org/10.31949/th.v4i2.1862>
- Stephenson, C. R., Bonnes, S. L., Sawatsky, A. P., Richards, L. W., Schleck, C. D., Mandrekar, J. N., Beckman, T. J., & Wittich, C. M. (2020). The relationship between learner engagement and teaching effectiveness: a novel assessment of student engagement in continuing medical education. *BMC Medical Education*, 20(1), 1–8. <https://doi.org/10.1186/s12909-020-02331-x>
- Ujang Charda S. (2021). Typology of Legal Research Methods in Normative and Sociological Thinking. *Fox Justi : Jurnal Ilmu Hukum*, 12(1), 111–118. <https://doi.org/10.58471/justi.v12i1.769>
- Warju, W., Ariyanto, S. R., Soeryanto, S., & Trisna, R. A. (2020). Analisis Kualitas Butir Soal Tipe Hots Pada Kompetensi Sistem Rem Di Sekolah Menengah Kejuruan. *Jurnal Pendidikan Teknologi Dan Kejuruan*, 17(1), 95. <https://doi.org/10.23887/jptk-undiksha.v17i1.22914>
- Wulansari Ramadhany, F., Hakim, A., & Sulaeman, N. (2022). Analisis Penguasaan Konsep Siswa pada Materi Elastisitas dan Hukum Hooke dengan Pembelajaran Discovery Learning Berbantuan Laboratorium Virtual PhET (The Physics education Technology). *Jurnal Literasi Pendidikan Fisika (JLPF)*, 3(2), 165–175. <https://doi.org/10.30872/jlpf.v3i2.1477>
- Yue, C., & Xu, X. (2019). Review of Quantitative Methods Used in Chinese Educational Research, 1978–2018. *ECNU Review of Education*, 2(4), 515–543. <https://doi.org/10.1177/2096531119886692>