

## **Analisis Butir Soal Ulangan Harian Pembelajaran IPA Pada Kelas VIII Di SMPN 21 Semarang**

**Wike Aprilia Ningrum<sup>1</sup>, Rivanna Citraning Rahmawati\*<sup>2</sup>, Ipah Budi Minarti<sup>3</sup>, Roch.  
Mayang Mekar<sup>4</sup>**

Email: [wikeaprilial5@gmail.com](mailto:wikeaprilial5@gmail.com)<sup>1</sup>, [rivanna.biologi@gmail.com](mailto:rivanna.biologi@gmail.com)\*<sup>2</sup>

<sup>1</sup>Pendidikan Profesi Guru Universitas PGRI Semarang

<sup>2,3</sup>Pendidikan Biologi Universitas PGRI Semarang

<sup>4</sup>SMP Negeri 21 Semarang

### **Abstract**

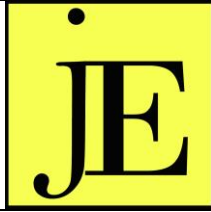
*To determine whether the prepared goods are worthy of being preserved or not, it is necessary to analyze the quality of the items. The purpose of this study is to evaluate the validity, reliability, level of difficulty, and differentiating power of the items on daily exam questions for scientific classes on the digestive system for class VIII G at SMPN 21 Semarang. The investigation was conducted on November 24, 2022. Research is being conducted using a quantitative descriptive approach. The questions, answer choices, and student responses are the focus of this study. The study's findings showed that while 28% of the multiple-choice questions and 20% of the essay questions were found in the items that were deemed invalid, 72% of the multiple-choice questions and 80% of the essay questions were found to be valid. In the multiple choice questions, the reliability of the item items shown a high consistency of 0.816, while the description questions demonstrated a sufficient consistency of 0.441. The results of the multiple choice questions and descriptions' discriminating capacity were very good, moderate, good, and not good. The results in the categories of medium, moderate, and too easy are displayed at the multiple-choice difficulty level and in the explanations. According to the study's findings, the items that were excellent enough or valid can be reused, while those that weren't could be improved upon or created from scratch using the indicators. The KKM score has been achieved for overall learning outcomes, though.*

**Keywords:** *Items analysis, Validity, Reliability, Difficulty Level, Discrimination.*

### **Abstrak**

*Menganalisis kualitas butir soal merupakan kegiatan yang harus dilakukan untuk mengetahui butir soal yang disusun layak dipertahankan atau tidak. Penelitian ini bertujuan untuk mengetahui tingkat kualitas butir soal ulangan harian pembelajaran IPA materi sistem pencernaan kelas VIII G di SMPN 21 Semarang yang ditinjau dari aspek validitas, reliabilitas, tingkat kesukaran, dan daya pembeda. Penelitian dilaksanakan pada 24 November 2022. Metode penelitian yang digunakan adalah penelitian deskriptif kuantitatif. Objek penelitian ini adalah butir soal, kunci jawaban, dan jawaban peserta didik. Hasil dari penelitian menunjukkan bahwa validitas butir soal menunjukkan 72% soal pilihan ganda dan 80% soal uraian dinyatakan valid, sedangkan butir soal yang dinyatakan tidak valid menunjukkan 28% soal pilihan ganda dan 20% soal uraian. Reliabilitas butir soal menunjukkan kekonsistennya tinggi sebesar 0,816 pada soal pilihan ganda, sedangkan soal uraian menunjukkan kekonsistennya cukup sebesar 0,441. Daya pembeda soal pilihan ganda dan uraian menunjukkan hasil sangat baik, cukup baik, baik dan tidak baik. Pada tingkat kesukaran pilihan ganda dan uraian menunjukkan hasil kategori sedang, mudah, dan terlalu mudah. Hasil penelitian disimpulkan bahwa pada butir soal yang valid bisa digunakan kembali sedangkan butir soal yang tidak valid atau kurang baik bisa diperbaiki atau dibuatkan baru sesuai dengan indikator. Namun secara keseluruhan hasil belajar sudah mencapai nilai KKM.*

**Kata kunci:** *Analisis Butir Soal, Validitas; Reliabilitas; Tingkat Kesukaran; Daya Pembeda*



## **PENDAHULUAN**

Pendidikan merupakan suatu hal yang penting bagi manusia. Pentingnya pendidikan sebagai upaya untuk mencerdaskan generasi bangsa yang dapat diperoleh dengan mengikuti pendidikan di sekolah, mengikuti kegiatan pelatihan maupun penelitian. Seiring dengan perkembangan zaman, pentingnya peningkatan kualitas pendidikan untuk dapat disesuaikan dengan perkembangan zaman dan kebutuhan dari peserta didik. Peningkatan kualitas pendidikan ini tidak terlepas dari berbagai komponen yang ada di dalamnya, seperti kompetensi guru, perangkat kurikulum, sarana dan prasarana serta metode dan strategi pembelajaran yang lebih inovatif (Sari et al., 2021). Upaya lain yang dapat dilakukan untuk meningkatkan kualitas pendidikan pada proses pembelajaran juga dapat ditempuh melalui peningkatan kualitas asesmen atau sistem penilaian (Zainal, 2020). Sistem penilaian yang digunakan dalam proses pembelajaran yaitu berupa teknik tes dan non tes. Penilaian non tes menggunakan teknik seperti, pemberian angket, wawancara, ataupun teknik observasi. Sedangkan penilaian berupa tes, teknik yang digunakan seperti soal-soal yang harus dijawab oleh peserta didik, soal bisa berupa pilihan ganda, isian singkat ataupun uraian. Salah satu alat penilaian yang sering digunakan oleh guru untuk mengukur ketercapaian tujuan pembelajaran adalah dengan pelaksanaan tes ulangan harian. Ulangan harian merupakan suatu kegiatan yang dilakukan oleh guru untuk mengukur ketercapaian tujuan belajar peserta didik setelah menyelesaikan satu kompetensi dasar (KD) (Depdiknas, 2013)

Tes digunakan untuk mengetahui kemampuan seseorang dalam memahami materi-materi yang telah pelajari sebelumnya. Tes merupakan salah satu alat evaluasi yang sistematis, komprehensif, dan objektif yang dapat dijadikan acuan oleh guru untuk mengambil keputusan pembelajaran yang telah dilakukan (Hariyatmi & Marsiyah, 2018). Sesuai dengan pendapat (Muluki, 2020) bahwa tes merupakan suatu sarana yang digunakan untuk mengukur keberhasilan peserta didik dalam mencapai tujuan pendidikan dan pengajaran. Menurut Sudijono dalam (Muluki, 2020) bahwa tes memiliki dua fungsi, diantaranya sebagai alat untuk mengukur kemampuan peserta didik dari proses pembelajaran yang telah diperolehnya dan sebagai alat ukur keberhasilan program pengajaran yang dilakukan oleh guru. Peranan penting dari alat ukur penilaian seperti tes hendaknya soal-soal yang disusun disesuaikan dengan kemampuan dan kebutuhan peserta didik, sesuai dengan konten atau materi, serta sesuai dengan tujuan pembelajaran yang akan dicapai, dan soal tersebut dapat dijadikan acuan yang akurat untuk mendapatkan hasil belajar peserta didik. Menurut (Sopiah et al., 2019) bahwa salah satu syarat soal tes yang baik terdiri dari soal-soal yang ditulis dengan baik dan dapat dijadikan sebagai alat untuk mengukur kemampuan dan keterampilan peserta didik untuk mencapai tujuan pembelajaran. Soal yang bermutu adalah soal yang dapat memberikan sebuah informasi mengenai sejauh mana pemahaman peserta didik terhadap materi yang telah diajarkan oleh guru (Fiska et al., 2021).

Cara yang dapat dilakukan untuk mengetahui soal tes yang telah dibuat tergolong layak dan efektif maka dibutuhkan kegiatan analisis butir soal. Analisis butir soal adalah suatu kegiatan yang dilakukan untuk mengidentifikasi kekurangan dalam suatu soal pada tes hasil belajar peserta didik (Virginia et al., 2021). Menurut Daryanto dalam (Salvina et al., 2019) bahwa tujuan menganalisis butir soal untuk mengidentifikasi soal-soal yang baik, kurang baik, dan soal jelek serta dapat digunakan untuk melakukan sebuah perbaikan pada soal yang telah dibuat. Ada beberapa syarat yang diperhatikan dalam menganalisis butir soal diantaranya, dari segi, 1) validitas, 2) reliabilitas, 3) objektivitas, 4) praktisibilitas dan (5) ekonomis.

Tujuan dari penelitian ini untuk mengetahui kualitas butir soal ulangan harian yang telah dikerjakan oleh peserta didik yang dilihat dari aspek validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan yang dapat dijadikan sebuah perbaikan butir soal untuk mengukur dan meningkatkan tingkat pemahaman peserta didik. Penelitian ini diharapkan bisa bermanfaat untuk berbagai pihak serta penelitian ini dapat dijadikan sebagai informasi untuk mengetahui kualitas dari butir soal yang dibuat sehingga dapat diketahui dianalisis kualitas butir soal apakah layak dipertahankan atau tidak digunakan, dan juga dapat menemukan penyebab permasalahannya. Sedangkan manfaat bagi peneliti, bahwa penelitian ini berfungsi untuk menambahkan wawasan mengenai penilaian khususnya pada bidang analisis butir soal.

## **METODE PENELITIAN**

Pada penelitian ini menggunakan metode deskriptif. Pendekatan yang digunakan pada penelitian ini adalah pendekatan kuantitatif karena hasil analisis pada butir soal dilakukan secara sistematis, terencana, dan terstruktur. Pada penelitian ini menghasilkan data berupa angka-angka dari perhitungan statistik (Sudijono, 2018). Penelitian dilakukan di SMP Negeri 21 Semarang pada tanggal 24 November 2022. Subjek dalam penelitian ini yaitu 33 peserta didik dari kelas VIII G. Objek dalam penelitian ini yaitu 25 butir soal pilihan ganda dan 5 butir soal uraian ulangan harian IPA pada materi “Sistem Pencernaan pada Manusia”. Data yang digunakan pada penelitian ini yaitu soal ulangan harian IPA materi “Sistem Pencernaan pada Manusia”, kunci jawaban, dan juga jawaban dari peserta didik.

Metode yang dapat digunakan untuk menganalisis butir soal ulangan harian dilakukan secara kuantitatif. Analisis secara kuantitatif biasanya ditinjau dari pendekatan klasik yang meliputi aspek validitas, reliabilitas, tingkat kesukaran, daya pembeda. Data yang diperoleh kemudian dianalisis dengan menggunakan *Microsoft Excel*.

### **Validitas**

Validitas yang digunakan untuk mengukur kemampuan hasil belajar peserta didik adalah validitas isi. Suatu soal dinyatakan valid apabila  $r_{hitung} \geq r_{tabel}$  (Rahayu & Djazari, 2016)

### **Reliabilitas**

Tinggi rendahnya reliabilitas ditunjukkan oleh angka yang disebut koefisien reliabilitas, koefisien reliabilitas berkisar antara 0 - 1, semakin tinggi angka reliabilitas artinya akan semakin konsisten hasil pengukurannya (Ida & Musyarofah, 2021).

**Tabel 1. Klasifikasi Realibilitas**

Rentang	Kualitas Butir Soal
$r_{11} < 0.20$	Kekonsistenannya sangat rendah
$0.20 \leq r_{11} < 0.40$	Kekonsistenannya rendah
$0.40 \leq r_{11} < 0.70$	Kekonsistenannya cukup
$0.70 \leq r_{11} < 0.90$	Kekonsistenannya tinggi
$0.90 \leq r_{11} < 1.00$	Kekonsistenannya sangat tinggi

(Azis, 2016)

#### Tingkat Kesukaran

Tingkat Kesukaran soal adalah peluang peserta didik dapat menjawab pertanyaan dengan benar pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks.

**Tabel 2. Indeks Kesukaran**

Indeks Kesukaran	Kategori
$IK=0.00$	Terlalu sukar
$0.30 < IK \leq 0.30$	Sukar
$0.30 < IK \leq 0.70$	Sedang
$0.70 < IK \leq 1.00$	Mudah
$IK=1.00$	Terlalu mudah

(Subana & Sudrajat, 2011)

#### Daya Pembeda

Daya pembeda digunakan untuk mengetahui peserta didik yang telah menguasai kompetensi materi dan peserta didik yang belum menguasai kompetensi materi setelah melakukan proses pembelajaran (Kurniawan, 2015).

**Tabel 3. Klasifikasi Daya Pembeda**

Rentang	Kualitas Butir Soal
$0,40 \leq D \leq 1,0$	Sangat Baik
$0,3 D < 0,4$	Baik
$0,2 D < 0,3$	Cukup Baik
$D < 0,2$	Tidak Baik

(Fiska et al., 2021)

## HASIL DAN PEMBAHASAN

Data hasil penelitian mengenai analisis kualitas butir soal ulangan harian pembelajaran IPA pada materi sistem pencernaan kelas VIII G di SMPN 21 Semarang adalah sebagai berikut:

### 1. Validitas

Hasil analisis data pada aspek validitas butir soal pilihan ganda dan uraian dapat dilihat pada Tabel 4. dan Tabel 5. di bawah ini:

**Tabel 4.** Validitas Butir Soal Pilihan Ganda

No	Validitas	Nomor Butir Soal	Jumlah	Persentase
1	Valid	2,4,5,6,7,8,9,11,13,16,17,18,20,21,22,23,24	17	68%
2	Tidak Valid	1,3,10,12,14,15,19,25	8	32%

**Tabel 5.** Validitas Butir Soal Uraian

No	Validitas	Nomor Butir Soal	Jumlah	Persentase
1	Valid	1,2,3,5	4	80%
2	Tidak Valid	4	1	20%

Berdasarkan Tabel 4. soal pilihan ganda dan Tabel 5. soal uraian diperoleh rtabel pada taraf signifikansi 5% yaitu 0,344. Sehingga berdasarkan hasil penelitian pada uji validitas berpatokan pada rtabel 0,344, apabila nilai rhitung  $\geq 0,344$  maka soal tersebut dinyatakan valid, sedangkan apabila rhitung  $\leq 0,344$  maka soal dinyatakan tidak valid. Hasil analisis pada Tabel 4. bahwa sejumlah 17 butir soal pilihan ganda (68%) dan pada Tabel 5. sejumlah 4 butir soal uraian (32%) dinyatakan valid dengan nilai rhitung  $\geq 0,344$ . Pada butir soal yang dinyatakan tidak valid sejumlah 8 butir soal pilihan ganda (32%) dan 1 butir soal uraian (20%) dengan nilai rhitung  $\leq 0,344$ .

Berdasarkan analisis data tersebut dapat disimpulkan bahwa butir soal yang telah dibuat memperoleh validitas dengan kualitas yang cukup baik karena jumlah soal yang dinyatakan valid lebih banyak dibandingkan dengan soal yang dinyatakan tidak valid. Butir soal dinyatakan valid karena mencakup materi yang mewakili sasaran ukurnya serta memberikan hasil yang sesuai dengan fungsi dan tujuannya. Menurut (Fiska et al., 2021) pada soal yang dinyatakan valid bisa disimpan pada bank soal untuk digunakan kembali, sedangkan pada butir soal yang tidak valid, maka harus direvisi dan disesuaikan dengan teknik penyusunan butir soal maupun dengan indikator pencapaian. Pada butir soal yang dinyatakan tidak valid akan memberikan dampak pada hasil yang diperoleh oleh peserta didik (Nurhasanah & Ahmad, 2017). Hasil yang diperoleh peserta didik bisa tidak mencapai KKM, sehingga perlu dikaji ulang sebelum soal tes digunakan. Apabila butir soal yang dinyatakan tidak valid digunakan, maka tidak dapat mengukur apa yang akan diukur, sehingga tidak memunculkan hasil yang sesuai dengan

kemampuan peserta didik yang sesungguhnya (Fiska et al., 2021). Hal ini selaras dengan teori yang dinyatakan oleh (Gronlund dalam Arifin, 2017) menyebutkan bahwa ada tiga faktor yang mempengaruhi validitas hasil tes seperti, instrumen yang digunakan, pedoman penskoran, serta faktor dari jawaban peserta didik.

## 2. Reliabilitas

Hasil analisis data pada aspek reliabilitas butir soal pilihan ganda dan uraian dapat dilihat pada Tabel 6. di bawah ini:

**Tabel 6.** Reliabilitas Butir Soal Pilihan Ganda dan Uraian

No	Reliabilitas	Koefisien	Jenis Soal
1	Reliabel	0,811	Pilihan Ganda
2	Reliabel	0,441	Uraian

Berdasarkan Tabel 6. dapat dilihat bahwa reliabilitas butir soal pilihan ganda memperoleh hasil 0,811 dinyatakan reliabel dengan kategori kekonsistennya tinggi. karena pada butir soal yang baik harus memiliki nilai reliabilitas  $\geq 0.70$ . Pada butir soal uraian diperoleh nilai 0,441 yang dinyatakan reliabel dengan kategori kekonsistennya cukup. Menurut (Rahmasari & Ismiyati, 2016) apabila suatu butir soal memiliki nilai reliabilitas  $< 0,70$ , maka reliabilitas tersebut masih dikategorikan belum baik dan perlu dilakukan revisi. Tinggi rendahnya nilai reliabilitas yang dihasilkan dapat dipengaruhi oleh beberapa faktor diantaranya, penyelenggaraan tes, jumlah peserta didik yang melaksanakan tes dan tingkat kesukaran pada butir soal. Beberapa bagian yang termasuk dalam penyelenggaraan tes seperti, petunjuk pengerjaan yang kurang jelas, kurangnya kesiapan peserta didik, dan tempat tes. Apabila bagian-bagian tersebut tidak mendukung maka akan mempengaruhi nilai reliabilitas. Semakin banyak jumlah peserta didik maka semakin beragam jawaban yang dihasilkan dan semakin rendah nilai reliabilitas yang dihasilkan. Semakin tinggi tingkat kesukaran, maka nilai reliabilitas akan semakin rendah karena butir soal yang sukar akan membuat peserta didik menjawab soal dengan cara menebak (Anita et al., 2018).

## 3. Tingkat Kesukaran

Hasil analisis data pada aspek tingkat kesukaran butir soal pilihan ganda dan uraian dapat dilihat pada Tabel 7. dan Tabel 8. di bawah ini:

**Tabel 7.** Hasil Tingkat Kesukaran Butir Soal Pilihan Ganda

No	Indeks Kesukaran	Nomor Butir Soal	Jumlah	Persentase
1	$0.30 < IK \leq 0.70$ Sedang	6,9,11,18	4	16%



2	0.70<IK≤1.00 Mudah	2,3,4,5,7,8,12,13,15,16,17,19,20, 21,22,23,24,25	18	72%
3	IK=1.00 Terlalu mudah	1,10,14	3	12%

**Tabel 8.** Hasil Tingkat Kesukaran Butir Soal Uraian

No	Kategori Soal	Nomor Butir Soal	Jumlah	Persentase
1	0.30<IK≤0.70 Sedang	2,3,4	3	60%
2	0.70<IK≤1.00 Mudah	1,5	2	40%

Berdasarkan Tabel 7. bahwa tingkat kesukaran soal pilihan ganda terdapat 18 butir soal (72%) termasuk dalam kategori mudah. Butir soal dengan kategori mudah dapat disebabkan karena banyak peserta didik yang dapat menjawab soal dengan benar. Butir soal dengan kategori mudah tidak akan merangsang kemampuan peserta didik dalam menjawab soal. Suatu butir soal dianggap baik apabila memiliki tingkat kesukaran antara 0,30 – 0,70 (Arikunto, 2013). Sehingga dapat disimpulkan bahwa butir soal yang dibuat kurang baik karena tingkat kesukaran dominan pada kategori mudah yaitu 0.70 – 1.00, sedangkan butir soal kategori sedang hanya 4 butir soal (16%) pada nomor 6,9,11, dan 18. Butir soal tersebut merupakan soal yang cukup sulit dijawab oleh sebagian besar peserta didik, bahasan soal mengenai karakteristik zat makanan dan mekanisme proses pencernaan. Hal tersebut sama dengan hasil pada Tabel 8. yang menunjukkan kategori sedang pada butir soal uraian nomor 2,3,4 dengan bahasan soal yang sama mengenai mekanisme proses pencernaan pada manusia. Hal tersebut menunjukkan bahwa sebagian besar peserta didik kurang memahami materi mekanisme proses pencernaan.

Butir soal yang baik yaitu butir soal dengan kategori sedang, sehingga butir soal yang dibuat sebaiknya tidak terlalu sukar dan tidak terlalu mudah. Butir soal dengan kategori sedang sebaiknya disimpan dalam bank soal. Hal ini sejalan dengan penelitian (Fatimah & Alfath, 2019) bahwa soal yang baik adalah soal yang kategorinya tidak terlalu mudah serta tidak terlalu sukar. Pada butir soal pilihan ganda juga menghasilkan data dengan tingkat kesukaran terlalu mudah sejumlah 3 butir soal (12%) yakni pada nomor 1, 10 dan 14, bahwa seluruh peserta didik mampu menjawab pertanyaan dengan benar. Butir soal dengan kategori terlalu mudah sebaiknya dilakukan revisi. Apabila butir soal terlalu mudah maka dapat menurunkan minat belajar peserta didik, sedangkan butir soal yang terlalu sukar dapat membuat peserta didik menjadi tidak bersemangat (Virginia et al., 2021). Sehingga, tingkat kesukaran pada setiap butir soal harus disesuaikan dengan tahap perkembangan peserta didik agar tes yang dilakukan dapat mengukur tingkat perkembangan maupun kemajuan peserta didik setelah melakukan suatu proses pembelajaran (Fiska et al., 2021).

#### 4. Daya Pembeda

Hasil analisis data pada aspek daya pembeda butir soal pilihan ganda dan uraian dapat dilihat pada Tabel 9. dan Tabel 10. di bawah ini:

**Tabel 9.** Daya Pembeda Butir Soal Pilihan Ganda

No	Daya Pembeda	Nomor Butir Soal	Jumlah	Persentase
1	$0,40 \leq D \leq 1,0$ Sangat Baik	6,8,9,11,13,17,18,21	8	32%
2	$0,3 D < 0,4$ Baik	2,5,7,16,20,22,23	7	28%
3	$0,2 D < 0,3$ Cukup Baik	3,4,24,25	4	16%
4	$D < 0,2$ Tidak Baik	1,10,12,14,15,19	6	24%

**Tabel 10.** Daya Pembeda Butir Soal Uraian

No	Daya Pembeda	Nomor Butir Soal	Jumlah	Persentase
1	$0,40 \leq D \leq 1,0$ Sangat Baik	2	1	20%
2	$0,2 D < 0,3$ Cukup Baik	1,3,5	3	60%
3	$D < 0,2$ Tidak Baik	4	1	20%

Berdasarkan Tabel 9. dapat dilihat bahwa aspek daya pembeda pada butir soal pilihan ganda menunjukkan kategori yang beragam, terdiri dari kategori sangat baik yang berjumlah 8 butir soal (32%), kategori baik berjumlah 7 butir soal (28%), kategori cukup baik berjumlah 4 butir soal (16%) dan kategori tidak baik berjumlah 6 butir soal (24%). Sedangkan berdasarkan Tabel 10. bahwa butir soal uraian hanya menunjukkan 3 kategori yang terdiri dari kategori sangat baik 1 butir soal (20%), cukup baik 3 butir soal (60%) dan tidak baik 1 butir soal (20%). Daya pembeda pada suatu soal dapat digunakan untuk membedakan antara peserta didik yang pencapaiannya tinggi dan peserta didik yang pencapaiannya rendah pada tes pencapaian hasil belajar (Rahmasari & Ismiyati, 2016). Berdasarkan hasil analisis bahwa ada keterkaitan antara daya pembeda dengan tingkat kesukaran pada butir soal. Pada butir soal pilihan ganda dengan kategori daya pembeda tidak baik pada nomor soal 1,10,12,14,15,19 merupakan butir soal yang paling banyak dijawab dengan benar oleh peserta didik. Sehingga butir soal tersebut tidak mampu membedakan peserta didik yang sudah menguasai materi ataupun peserta didik yang belum menguasai materi. Hal tersebut sejalan dengan penelitian (Fatimah & Alfath, 2019) bahwa daya pembeda dan tingkat kesukaran memiliki keterkaitan yang dimana apabila butir soal dengan tingkat kesukaran yang sangat sukar maka daya pembedanya akan rendah karena sedikitnya peserta didik yang mampu menjawab soal dengan benar, sedangkan pada butir soal



dengan tingkat kesukarannya mudah tidak akan memiliki daya pembeda karena banyak peserta didik yang mampu menjawab soal dengan benar. Butir soal dengan daya pembeda tidak baik harus dilakukan revisi, sedangkan butir soal dengan daya pembeda sangat baik, baik, dan cukup baik bisa disimpan di bank soal.

## SIMPULAN

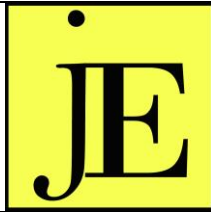
Berdasarkan hasil penelitian dan pembahasan mengenai kualitas analisis butir soal pilihan ganda dan soal uraian ulangan harian pembelajaran IPA pada materi sistem pencernaan semester ganjil pada kelas VIII G di SMP Negeri 21 Semarang tahun pelajaran 2022/2023 maka dapat disimpulkan bahwa kualitas butir soal tergolong valid karena persentase butir soal dinyatakan valid lebih banyak dibandingkan dengan butir soal yang dinyatakan tidak valid. Pada aspek reliabilitas bahwa butir soal tergolong reliabel dengan kekonsistennya tinggi pada pilihan ganda dan kekonsistennya cukup pada butir soal uraian. Pada aspek daya pembeda dapat disimpulkan bahwa butir soal memiliki daya pembeda yang beragam, namun tergolong cukup baik karena dominan nilai yang mendekati 1 lebih banyak dibandingkan dengan kategori tidak baik. Pada tingkat kesukaran butir soal lebih banyak dihasilkan dengan kategori mudah, hal ini kualitas butir soal tergolong kurang baik karena nilai yang dijadikan patokan berkisar 0.30-0.70 dengan kategori sedang. Hasil penelitian disimpulkan bahwa pada butir soal yang valid bisa digunakan kembali sedangkan tidak valid atau kurang baik bisa diperbaiki atau dibuatkan baru sesuai dengan indikator. Namun secara keseluruhan hasil belajar Sebagian besar peserta didik sudah mencapai nilai KKM.

## SARAN

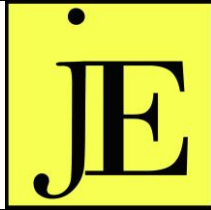
Berdasarkan hasil penelitian bahwa peneliti memberikan saran sebagai berikut: (1) menganalisis butir soal sebelum soal digunakan untuk tes dengan menggunakan instrument untuk menguji kevalidannya, (2) butir soal yang disusun disesuaikan dengan tahap perkembangan peserta didik dan indikator pencapaiannya, (3) penyusunan soal disesuaikan dengan kaidah penulisan butir soal yang baik dan benar.

## DAFTAR PUSTAKA

- Anita, A., Tyowati, S., & Zuldafrial, Z. (2018). Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Edukasi: Jurnal Pendidikan*, 16(1), 35. <https://doi.org/10.31571/edukasi.v16i1.780>
- Arifin, Z. (2017). *Evaluasi Pembelajaran* (10th ed.). PT Remaja Rosdakarya.
- Arikunto, S. (2013). *Dasar-Dasar Evaluasi Pendidikan* (R. Damayanti (ed.); 2nd ed.). Bumi Aksara.
- Azis, A. (2016). Analisis Tes Buatan Guru Bidang Studi Matematika Kelas V SD 1 Katobengke. *Edumatica: Jurnal Pendidikan Matematika*, 6(1), 15–24. <https://doi.org/10.22437/edumatica.v6i01.2998>



- Depdiknas. (2013). *Permendikbud No 66 Tahun 2013 tentang Standar Penilaian Pendidikan. 2011*, 1–6. <https://doi.org/10.1016/j.metabol.2009.10.012>
- Fatimah, L. U., & Alfath, K. (2019). Analisis Kesukaran Soal, Daya Pembeda dan Fungsi Distraktor. *AL-MANAR: Jurnal Komunikasi Dan Pendidikan Islam*, 8(2), 37–64. <https://doi.org/https://doi.org/10.36668/jal.v8i2.115>
- Fiska, J. M., Hidayati, Y., Qomaria, N., & Hadi, W. P. (2021). Analisis Butir Soal Ulangan Harian Ipa Menggunakan Software Anates Pada Pendekatan Teori Tes Klasik. *Natural Science Education Research*, 4(1), 65–76. <https://doi.org/10.21107/nser.v4i1.8133>
- Hariyatmi, & Marsiyah, F. (2018). Profil Soal Buatan Guru Biologi Berdasarkan Taksonomi Bloom Di SMA Muhammadiyah Surakarta. *Seminar Nasional Pendidikan Biologi Dan Saintek III*, 3(7), 572–579.
- Ida, F. F., & Musyarofah, A. (2021). Validitas dan Reliabilitas dalam Analisis Butir Soal. *Al-Mu'Arrib: Journal of Arabic Education*, 1(1), 34–44. <https://doi.org/10.32923/al-muarrib.v1i1.2100>
- Muluki, A. (2020). Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran IPA Kelas IV Mi Radhiatul Adawiyah. *Jurnal Ilmiah Sekolah Dasar*, 4(1), 86. <https://doi.org/10.23887/jisd.v4i1.23335>
- Nurhasanah, & Ahmad, H. (2017). Analisis Soal Ujian Tengah Semester (Uts) Mata Kuliah Geometri. *Jurnal Pendidikan Papatudzu*, 13(1), 30–42.
- Rahayu, R., & Djazari, M. (2016). Analisis Kualitas Soal Pra Ujian Nasional Mata Pelajaran Ekonomi Akuntansi. *Jurnal Pendidikan Akuntansi Indonesia*, 14(1). <https://doi.org/10.21831/jpai.v14i1.11370>
- Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran. *Economic Education Analysis Journal*, 5(1), 317–330.
- Salvina, N., Sidauruk, S., & Asi, N. B. (2019). Kualitas Soal Penilaian Akhir Semester (PAS) Buatan Guru Mata Pelajaran Kimia Kelas X SMK Jurusan Teknologi Dan Rekayasa Di Kabupaten Kotawaringin Timur Pada Semester Ganjil Tahun Ajaran 2018/2019. *Jurnal Ilmiah Kanderang Tingang*, 10(1), 46–56. <https://doi.org/10.37304/jikt.v10i1.20>
- Sari, I. N., Anggraini, M. P., Maryoni, R., & Walid, A. (2021). Analisis Kualitas Butir Soal Uts Pembelajaran Ipa Kelas Vii Di Smp Bp Pancasila. *ISEJ: Indonesian Science ...*, 2(3), 164–170. <https://www.siducat.org/index.php/isej/article/view/358%0Ahttps://www.siducat.org/index.php/isej/article/download/358/271>
- Sopiah, A., Sidauruk, S., & Asi, N. B. (2019). Kualitas Soal Penilaian Akhir Semester (PAS) Buatan Guru Mata Pelajaran Kimia Kelas X MIA SMA Di Kabupaten Kotawaringin Barat



Pada Semester Ganjil Tahun Ajaran 2018/2019. *Jurnal Ilmiah Kanderang Tingang*, 10(2), 110–126. <https://doi.org/https://doi.org/10.37304/jikt.v10i2.27>

Subana, M., & Sudrajat. (2011). *Dasar-Dasar Penelitian Ilmiah* (4th ed.). CV Pustaka Setia.

Sudijono, A. (2018). *Pengantar Statistik Pendidikan* (27th ed.). Rajawali Press.

Virginia, S., Angraini, W., Pratesya, W., & Walid, A. (2021). Analisis Butir Soal Ulangan Harian IPA Terpadu Kelas VII SMP 05 Kota Bengkulu Tahun Pelajaran 2020/2021. *Jurnal PTK Dan Pendidikan*, 6(2), 61–65. <https://doi.org/10.18592/ptk.v6i2.4173>

Zainal, N. F. (2020). Pengukuran, Assessment dan Evaluasi dalam Pembelajaran Matematika. *Laplace : Jurnal Pendidikan Matematika*, 3(1), 8–26. <https://doi.org/10.31537/laplace.v3i1.310>