

Uji Kelayakan Soal Pilihan Ganda *HOTS* Melalui Penerapan Anates Kelas XII MPLB SMKN 1 Surabaya

Artika Iesa Karunia Illahi¹, Narulita Diva Saraswati², Vivi Pratiwi³, Luqman Hakim⁴
^{1,2,3,4} Pendidikan Akuntansi, Universitas Negeri Surabaya
e-mail: artika.23045@mhs.unesa.ac.id

Abstrak

Penelitian ini bertujuan mengevaluasi kualitas 10 butir soal mata pelajaran Dasar-Dasar Manajemen Perkantoran pada materi dokumen berbasis digital pada 15 siswa/i di kelas XII SMK Negeri 1 Surabaya. Penelitian ini menggunakan metode penelitian kuantitatif dengan pendekatan deskriptif. Hasil analisis menggunakan perangkat lunak Anates 4.0 menunjukkan reliabilitas soal sebesar 0,49, mengindikasikan korelasi yang cukup. Sebanyak 60% butir soal memiliki daya beda yang baik ($\geq 0,5$), sedangkan sisanya perlu direvisi. Tingkat kesukaran soal berada pada kategori sedang (0,33-0,80). Validitas soal juga tergolong cukup (0,33-0,62). Namun, 62,5% pengecoh dalam soal dinilai kurang baik. Secara keseluruhan, hasil penelitian ini menunjukkan bahwasanya perlu adanya perbaikan kualitas instrumen penilaian pada mata pelajaran tersebut, terutama pada aspek daya beda dan kualitas pengecoh, untuk meningkatkan akurasi dan reliabilitas penilaian.

Kata kunci: *Analisis Butir Soal, HOTS, Anates*

Abstract

This research aims to evaluate the quality of 10 Basic Office Management test questions related to digital-based document material for 15 class XII students at SMK Negeri 1 Surabaya. Using quantitative methods with a descriptive approach, the analysis shows sufficient reliability (0.49) but indicates the need for improvement. 60% of the items effectively differentiated between high and low achieving students, while the remaining items required revision. The level of difficulty is moderate (0.33-0.80), and the validity is sufficient (0.33-0.62). However, 62.5% of distractors were deemed ineffective. Overall, these findings suggest a need to improve the quality of assessment instruments, particularly focusing on improving item differentiation and distractor effectiveness to ensure more accurate and reliable assessments.

Keywords : *Analysis of Question, HOTS, Anates*

PENDAHULUAN

Pendidikan merupakan pilar utama pembangunan nasional. Dalam konteks ini, evaluasi berperan krusial untuk mengukur capaian pembelajaran peserta didik. Ujian menjadi instrumen evaluasi yang lazim digunakan. Untuk menghasilkan ujian yang valid dan reliabel, analisis butir soal menjadi langkah yang sangat penting. Penilaian secara komprehensif akan memberikan gambaran yang akurat tentang pemahaman serta penguasaan materi akademik peserta didik. Di

antara berbagai metode penilaian yang ada, soal pilihan ganda mendominasi sebagai instrumen penilaian yang paling sering digunakan oleh guru SMK dalam mengukur pemahaman siswa terhadap materi pelajaran. Untuk mengukur kompetensi siswa. Pemilihan instrumen pilihan ganda didasarkan pada efisiensi, reliabilitas, dan kemudahan dalam proses standarisasi. Tes pilihan ganda merupakan penilaian yang dapat dinilai secara objektif berdasarkan jawaban benar-salah. Soal ini memiliki banyak pilihan jawaban dan biasanya menggunakan skala A sampai E. Keunggulannya adalah mudah dinilai dan fleksibel untuk berbagai mata pelajaran. Pada saat melakukan penilaian, ada beberapa metode yang bisa digunakan, seperti menjumlahkan jawaban benar (teknik konvensional). Kualitas suatu tes pilihan ganda dapat dinilai dari tingkat kesulitan soal, kemampuan membedakan siswa yang pandai dan kurang pandai, serta efektivitas pilihan jawaban yang salah (Putri Mulyani et al., 2024.). Dengan merancang soal pilihan ganda yang baik, kita dapat mendorong siswa agar mampu mencapai level berpikir yang lebih kompleks, seperti menganalisis, mengevaluasi, dan menciptakan, sebagaimana yang dijelaskan dalam taksonomi Bloom dan Piramida Miller (Hidayatullah et al., 2023).

HOTS adalah kemampuan berpikir yang melibatkan proses mental yang kompleks. Siswa tidak hanya mengolah informasi secara pasif, tetapi juga aktif bernalar, berpikir kritis, menarik kesimpulan, dan mengembangkan strategi pemecahan masalah yang inovatif (Herman et al., 2022). Pada penilaian menggunakan soal HOTS, terdiri dari beberapa tahapan, diantaranya analisis yang mencakup kemampuan membedakan, mengorganisasi, dan mengasosiasikan informasi. Tahap evaluasi melibatkan keterampilan melihat kembali dan mengkritisi secara mendalam suatu informasi atau ide untuk menghasilkan suatu kesimpulan.

Arsip merupakan aset strategis bagi organisasi yang dapat mendukung pengambilan keputusan, evaluasi kinerja, dan inovasi. Keberhasilan pemanfaatan arsip sangat bergantung pada kualitas pengelolaannya. Pada era digitalisasi yang pesat, organisasi dari berbagai sektor dituntut untuk melakukan transformasi digital guna meningkatkan produktivitas. Pengarsipan digital merupakan salah satu pilar penting dalam transformasi ini. Di era 5.0 yang serba cepat pengelolaan dokumen berbasis teknologi menjadi suatu keharusan untuk meningkatkan efisiensi dan produktivitas (Francisca Lies Ambarwati et al., 2022.).

Teknologi pada era 5.0 membawa kita menuju masa depan di mana manusia dan mesin bekerja sama secara harmonis. Era digital yang ditandai dengan kemajuan pesat teknologi komputer dan internet telah menciptakan paradigma baru dalam kehidupan manusia, di mana informasi dan konektivitas menjadi elemen sentral dalam setiap aktivitas. Teknologi digital kini telah menjadi kebutuhan dasar dalam berbagai aspek kehidupan (Mahgafhira et al., 2023). Konektivitas yang semakin canggih akan memungkinkan terciptanya solusi inovatif untuk berbagai tantangan global, salah satunya di bidang pendidikan dan pekerjaan. Teknologi merupakan faktor penentu dalam meningkatkan standar pendidikan. Oleh karena itu, pendidikan harus terus berinovasi dengan mengadopsi teknologi terbaru. Konsep Society 5.0, yang menggabungkan teknologi digital dengan kehidupan sehari-hari, membuka peluang baru untuk menciptakan masyarakat yang lebih cerdas, inklusif, dan berkelanjutan yang dapat berkontribusi pada peningkatan kualitas hidup manusia secara signifikan. Dengan mengintegrasikan teknologi dalam pembelajaran, guru dapat membekali siswa dengan keterampilan abad 21 yang akan membantu mereka beradaptasi dengan perubahan zaman (Nasution, et al., 2023). Transformasi pendidikan di era modern, utamanya di era Teknologi 5.0 perlu disesuaikan dengan mengintegrasikan

teknologi yang semakin maju ke dalam pembelajaran, sehingga mata pelajaran dokumen berbasis digital ini sangat relevan dengan kebutuhan di era 5.0. Untuk mempersiapkan generasi masa depan, diperlukan kerja sama yang solid antara sekolah, pemerintah, organisasi masyarakat, dan seluruh lapisan masyarakat.

Tujuan utama analisis butir soal adalah untuk mengukur kualitas soal dari segi validitas, reliabilitas, daya pembeda, dan kejelasan, sehingga dapat menghasilkan data yang akurat dan reliabel. Analisis butir soal memungkinkan kita untuk menyempurnakan tes dengan cara merevisi soal-soal yang kurang efektif dan menambahkan soal-soal baru yang lebih relevan dengan tujuan pembelajaran (Marambaawang et al., 2023). Soal yang berkualitas adalah kunci untuk mendapatkan hasil asesmen yang valid dan reliabel (Putri Andini & Mukhlis, 2023). Dengan demikian, analisis butir soal berperan penting dalam meningkatkan kualitas penilaian dan memberikan umpan balik yang lebih akurat bagi guru dan peserta didik. Guru adalah kunci keberhasilan dalam menyusun evaluasi pembelajaran yang efektif (Puspa Hanan et al., 2023). Di era digital, sinergi yang kuat di antara ketiga komponen ini sangat penting untuk menghasilkan sumber daya manusia yang berkualitas dan siap kerja.

METODE

Penelitian ini bertujuan untuk mengukur seberapa baik soal-soal ujian Dasar-Dasar Manajemen Perkantoran dan Layanan Bisnis (khususnya materi dokumen digital) sesuai dan efektif untuk menilai pemahaman siswa kelas XII di SMK Negeri 1 Surabaya. Dengan menggunakan aplikasi Quizizz, penelitian ini akan mengukur kesesuaian dan keefektifan soal-soal ujian Dasar-Dasar Manajemen Perkantoran dan Layanan Bisnis (materi dokumen digital) pada 15 siswa kelas XII di SMK Negeri 1 Surabaya. Penelitian ini mengadopsi pendekatan kuantitatif deskriptif dengan memanfaatkan perangkat lunak statistik Anates Versi 4.0 untuk menganalisis data. Program ini memungkinkan peneliti untuk memperoleh hasil analisis yang akurat dan efisien. Dengan demikian, dapat meningkatkan keefektifan dalam mengukur pencapaian belajar siswa. Analisis yang dilakukan mencakup penilaian terhadap validitas, reliabilitas, tingkat kesulitan, daya pembeda, serta kualitas pengalih perhatian dari setiap butir soal, sehingga diperoleh gambaran yang komprehensif mengenai kualitas instrumen penelitian.

HASIL DAN PEMBAHASAN

```
DATA MENTAH
*****
Jumlah Subyek= 15
Jumlah Butir Soal= 10
Jumlah Pilihan Jawaban= 5
Nama berkas: D:\CULYEAH\EVAL BELAJAR\ANATES 10 SOAL PILGAN.ANA

Nomor Nomor No. Butir Baru ----> 1 2 3 4 5 6 7 8 9
Urut Subyek No. Butir Asli ----> 1 2 3 4 5 6 7 8 9
Nama Subyek | Kunci ->
1 1 NAGATA A E C A C D D B A
2 2 RAFA ALIF ARYA RAHMANA A E A A C D D D A
3 3 RAHAYU PURWATI E B C C D D D D A
4 4 NADHIFA F S A E B A B D C E D
5 5 NAHDHAH NAILAH A E B B B D C D E
6 6 NAJWA SABRINA A E B B B D C D E
7 7 NOVITA AULIA S A E B A B D C D E
8 8 NUR LAYLA F A D D A A D C D C
9 9 RAHMA DIASTY C E B B B E B D D
10 10 REFINA A A E B A B D D D D
11 11 RETNO WULAN LESTARI B E B A B B C E D
12 12 ISSAH C E B A B D C E D
13 13 NAJWA TALITHA A E B A A C C D E
14 14 RINDA NUR FAUZIAH A B B A B D C D E
15 15 NANDYA SHAVIRA AGUSTI... C E B A D D C B D
```

Gambar 1. Hasil Data Tes

Perkembangan teknologi telah memungkinkan analisis butir soal dilakukan secara lebih efisien dan akurat. Penelitian ini terdiri dari beberapa tahapan yakni, yang pertama tahap persiapan didalamnya melalui beberapa cara diantaranya, menentukan struktur soal melalui capaian pembelajaran, tujuan pembelajaran dan alur tujuan pembelajaran berdasarkan bahan ajar yang telah dibuat, menyiapkan media pendistribusian soal ke peserta didik yakni quizizz, menyiapkan alat uji validitas soal dengan software anates. Tahap yang kedua yakni tahap pelaksanaan yang dilakukan dengan menginput butir tes yang sudah disusun selanjutnya diinput pada aplikasi quizizz serta membagikan kode untuk bergabung atau link untuk bergabung kepada peserta didik ketika nilai sudah muncul maka dilakukan pengelolaan dan pengolahan data hasil penelitian berupa hasil tes peserta didik dengan merekap nilai serta jawaban peserta didik agar lebih mudah dalam menganalisis yang terakhir data hasil tes dimasukkan ke dalam program komputer Anates versi 4.0 untuk Windows untuk melakukan analisis validitas soal lebih lanjut.

Pada penerapan software anates perlu dilakukan analisis data dan kriteria validitas butir soal yang pertama dalam analisis kelayakan soal adalah item skor data dibobot. Skor data dibobot merupakan nilai yang diberikan pada suatu butir soal setelah mempertimbangkan faktor-faktor tertentu yang mempengaruhi tingkat kesulitan atau pentingnya butir soal tersebut. Dengan kata lain, biasanya tidak semua soal memiliki bobot yang sama. Pada penelitian ini skor bobot yang digunakan adalah apabila benar akan mendapatkan poin sebesar 10 sedangkan untuk jawaban salah akan mendapatkan poin sebesar 2. Hasil dari skor data dibobot yang diperoleh dari penelitian kami ditunjukkan pada gambar 2.

| No. Ut | No. Subyek | Kode/Nama | Benar | Salah | Kosong | Sko Asli | Sko Bobot |
|--------|------------|------------------------------|-------|-------|--------|----------|-----------|
| 1 | 7 | NOVITA AULIA S | 10 | 0 | 0 | 10 | 100 |
| 2 | 5 | NAHDAH NAILAH | 9 | 1 | 0 | 9 | 92 |
| 3 | 6 | NAJWA SABRINA | 9 | 1 | 0 | 9 | 92 |
| 4 | 14 | RINDA NUR FALZAH | 9 | 1 | 0 | 9 | 92 |
| 5 | 4 | NACHIFA F Z | 8 | 2 | 0 | 8 | 84 |
| 6 | 13 | NAJWA TALITHA | 8 | 2 | 0 | 8 | 84 |
| 7 | 10 | REFINA A | 7 | 3 | 0 | 7 | 76 |
| 8 | 2 | RAFFA ALIF ARIYA RAHMANA | 6 | 4 | 0 | 6 | 68 |
| 9 | 8 | NUR LAYLA F | 6 | 4 | 0 | 6 | 68 |
| 10 | 12 | IZZAH | 6 | 4 | 0 | 6 | 68 |
| 11 | 15 | NANDYA SHAVIRA AGUSTININGRUM | 6 | 4 | 0 | 6 | 68 |
| 12 | 9 | RAHMA DIASTY | 5 | 5 | 0 | 5 | 60 |
| 13 | 11 | RETNO WULAN LESTARI | 5 | 5 | 0 | 5 | 60 |
| 14 | 1 | NAGATA | 4 | 6 | 0 | 4 | 52 |
| 15 | 3 | RAHAYU PURWATI | 3 | 7 | 0 | 3 | 44 |

Gambar 2. Hasil Skor Dibobot

Reliabilitas, yang berasal dari kata "*reliability*", merujuk pada tingkat ketelitian dan konsistensi suatu pengukuran dalam menghasilkan data. Artinya, reliabilitas menunjukkan sejauh mana kita dapat mengandalkan hasil pengukuran untuk menggambarkan fenomena yang sebenarnya. Pada sebuah penelitian, reliabilitas dapat diukur dari berbagai aspek, seperti kesamaan hasil pengukuran yang dilakukan oleh penilai yang berbeda, stabilitas hasil pengukuran dari waktu ke waktu, serta kesamaan hasil pengukuran pada bentuk tes yang berbeda namun setara. Tingginya reliabilitas suatu soal dipengaruhi oleh beberapa faktor, seperti banyaknya soal yang diberikan, durasi waktu pengerjaan tes, variasi tingkat kesulitan soal, serta objektivitas dalam

memberikan nilai (Sijabat et al., 2024). Butir soal yang memenuhi nilai korelasi reliabilitas minimal yang tercantum pada Tabel 1 dapat digunakan dalam instrumen pengukuran.

Tabel 1: Kriteria Korelasi Koefisien

| Indeks Reliabilitas (r) | Tingkat Reliabilitas |
|-------------------------|-----------------------------------|
| $0,00 \leq r < 0,20$ | Korelasi Sangat Rendah |
| $0,20 \leq r < 0,40$ | Korelasi Rendah |
| $0,40 \leq r < 0,70$ | Korelasi Cukup |
| $0,70 \leq r < 0,90$ | Korelasi Tinggi |
| $0,90 \leq r \leq 1,00$ | Korelasi Sangat Tinggi (Sempurna) |

| No.Urut | No. Subyek | Kode>Nama Subyek | Skor Ganjil | Skor Genap | Skor Total |
|---------|------------|------------------|-------------|------------|------------|
| 1 | 1 | NAGATA | 1 | 3 | 4 |
| 2 | 2 | RAFFA ALIF AR... | 1 | 5 | 6 |
| 3 | 3 | RAHAYU PURWATI | 0 | 3 | 3 |
| 4 | 4 | NADHIFA F 2 | 4 | 4 | 8 |
| 5 | 5 | NAHDAN NALLAH | 5 | 4 | 9 |
| 6 | 6 | NAJWA SABRINA | 5 | 4 | 9 |
| 7 | 7 | NOVITA AULIA S | 5 | 5 | 10 |
| 8 | 8 | NUR LAVIA F | 2 | 4 | 6 |
| 9 | 9 | RAMA DIASTY | 2 | 3 | 5 |
| 10 | 10 | REFINA A | 3 | 4 | 7 |
| 11 | 11 | RETNO WULAN L... | 3 | 2 | 5 |
| 12 | 12 | ISDAH | 3 | 3 | 6 |
| 13 | 13 | NAJWA SALIHA | 4 | 4 | 8 |
| 14 | 14 | RINDA NUR FAU... | 5 | 4 | 9 |
| 15 | 15 | NANDYA SHAVIR... | 2 | 4 | 6 |

Gambar 3 : Hasil Reliabilitas Tes

Dari tabel 1 diketahui bahwasanya reliabilitas soal yang dapat dinyatakan reliabel yakni jika reliabel memiliki nilai $r > 0,50$ dapat dinyatakan reliabel. Pada penelitian ini hasil kelayakan reliabilitas soal ditunjukkan pada gambar 3. Berdasarkan data hasil uji kelayakan pada tabel 1 hasil dari uji kelayakan soal menunjukkan bahwa Reliabilitas Tes dari uji analisis soal sebesar 0,49. Berdasarkan penilaian terhadap soal, dapat diketahui bahwa butir soal yang telah disusun mempunyai **tingkat korelasi yang cukup** untuk digunakan sebagai instrumen pengukuran.

Setiap peserta didik memiliki kemampuan yang berbeda-beda. Oleh karena itu, pada software anates terdapat pembagian antara kelompok unggul dan kelompok asor (Dimana).

| No.Urut | No Subyek | Kode>Nama Subyek | Skor | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------------|-----------|------------------|------|---|---|---|---|---|---|---|
| 1 | 9 | RAHMA DIASTY | 5 | - | 1 | 1 | - | 1 | - | - |
| 2 | 11 | RETNO WULAN L... | 5 | - | 1 | 1 | 1 | 1 | 1 | - |
| 3 | 1 | NAGATA | 4 | 1 | 1 | - | 1 | - | 1 | - |
| 4 | 3 | RAHAYU PURWATI | 3 | - | - | - | - | - | 1 | - |
| Jml Jwb Benar | | | | 1 | 3 | 2 | 2 | 2 | 2 | 1 |

Gambar 4. Hasil Kelompok Asor

| No.Urut | No Subyek | Kode>Nama Subyek | Skor | 1 | 2 | 3 | 4 | 5 | 6 | 7 |
|---------------|-----------|------------------|------|---|---|---|---|---|---|---|
| 1 | 7 | NOVITA AULIA S | 10 | 1 | 1 | 1 | 1 | 1 | 1 | 1 |
| 2 | 5 | NAHDAN NALLAH | 9 | 1 | 1 | 1 | - | 1 | 1 | 1 |
| 3 | 6 | NAJWA SABRINA | 9 | 1 | 1 | 1 | - | 1 | 1 | 1 |
| 4 | 14 | RINDA NUR FAU... | 9 | 1 | - | 1 | 1 | 1 | 1 | 1 |
| Jml Jwb Benar | | | | 4 | 3 | 4 | 2 | 4 | 4 | 4 |

Gambar 5. Hasil Kelompok Unggul

Kelompok asor adalah mereka yang secara umum memiliki kesulitan dalam memahami materi pelajaran dan memperoleh nilai yang rendah pada evaluasi. Kelompok unggul adalah sekelompok individu yang memiliki kecerdasan dan pemahaman konsep yang mendalam (Dzil et al., 2024). Dimana hasil kelompok asor dapat dilihat pada gambar 4, sedangkan untuk kelompok unggul dapat dilihat pada gambar 5. Berdasarkan uji kelayakan soal yang telah dilakukan terhadap 15 peserta didik, terdapat 4 peserta yang masuk kedalam kelompok unggul atau kelompok nilai tertinggi yakni Novita Aulia S, Nahdah Nailah, Najwa Sabrina, dan Rinda Nur. Berdasarkan uji kelayakan soal yang telah dilakukan terhadap 15 peserta didik, terdapat 4 peserta yang masuk kedalam kelompok asor atau kelompok nilai terendah yakni Rahma Diasty, Retno Wulan L, Nagat, dan Rahayu Purwati.

Daya pembeda suatu soal menunjukkan seberapa baik soal tersebut dapat mengidentifikasi peserta didik yang telah memahami pembelajaran dengan peserta didik yang belum memahami pembelajaran. Analisis daya pembeda dilakukan dengan membandingkan proporsi jawaban benar pada kelompok peserta dengan skor tertinggi dan terendah. Kemudian, kita bandingkan proporsi jawaban benar pada kedua kelompok tersebut. Sebuah pertanyaan dikatakan mempunyai daya pembeda yang baik apabila memenuhi kriteria indeks daya pembeda sebagaimana tercantum pada Tabel 2.

Tabel 2. Kriteria Koefisien Butir Soal

| Koefisien Butir Soal | Daya Pembeda |
|----------------------|----------------------------------|
| 0,00 - 0,20 | Daya Pembeda Jelek |
| 0,20 - 0,40 | Daya Pembeda Cukup |
| 0,40 - 0,70 | Daya Pembeda Baik |
| 0,70 - 1,00 | Daya Pembeda Sangat Baik |
| Negatif | Daya pembeda tidak ada yang baik |

```

DAYA PEMBEDA
=====
Jumlah Subyek= 15
Klp Atas/bawah(n)= 4
Butir Soal= 10
Nama berkas: D:\CULYEAH\EVAL BELAJAR\ANATES 10 SOAL PILGAN.ANA

No Butir Baru  No Butir Asli  Kel. Atas  Kel. Bawah  Beda  Indeks DP (%)
1              1              4          1           3      75,00
2              2              3          3           0       0,00
3              3              4          2           2      50,00
4              4              2          2           0       0,00
5              5              4          2           2      50,00
6              6              4          2           2      50,00
7              7              4          1           3      75,00
8              8              4          2           2      50,00
9              9              4          0           4     100,00
10             10             4          2           2      50,00
    
```

Gambar 6. Hasil Indeks Daya Pembeda

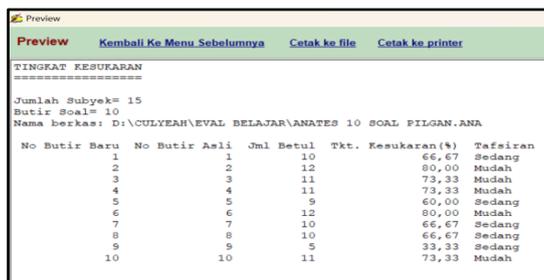
Berdasarkan hasil pada gambar 6 hasil daya pembeda pada butir soal dapat dirinci sebagai berikut, no 1 sebesar 0,75 dinyatakan sangat baik, no 2 sebesar 0,00 dinyatakan jelek, no 3 sebesar 0,50 dinyatakan baik, no 4 sebesar 0,00 dinyatakan jelek, no 5 sebesar 0,50 dinyatakan baik, no 6 sebesar 0,50 dinyatakan baik, no 7 sebesar 0,75 dinyatakan sangat baik, no 8 sebesar 0,50 dinyatakan baik, no 9 sebesar 1,00 dinyatakan sangat baik, no 10 sebesar 0,50 dinyatakan

baik. Sesuai dengan rincian hasil analisis indeks daya pembeda mengacu pada tabel 2 . Maka, butir soal yang memiliki indeks pembeda > 0,4 dianggap sebagai butir soal yang mencapai tujuan diferensiasi dengan baik. Dari hasil evaluasi butir soal, didapati bahwa delapan butir soal, yakni nomor 1, 3, 5, 6, 7, 8, 9, dan 10, berhasil membedakan peserta dengan kemampuan tinggi dan rendah secara efektif. Hal ini dibuktikan dengan nilai indeks daya beda masing-masing butir yang memenuhi kriteria $\geq 0,5$. Sedangkan nomor 2 dan 4 mempunyai daya beda yang jelek dikarenakan memiliki indeks daya pembeda < 0,5 sehingga soal tersebut perlu dilakukan perbaikan atau revisi.

Proporsi jawaban benar peserta didik menjadi tolok ukur yang valid untuk mengukur Soal yang terlalu mudah atau terlalu sukar akan menghambat proses pembelajaran. Butir soal dianggap memiliki tingkat kesukaran yang baik jika memenuhi kriteria indeks kesukaran soal yang ditunjukkan pada Tabel 3.

Tabel 3: Kriteria Tingkat Kesukaran

| Koefisien Butir Soal | Tingkat Kesukaran |
|--------------------------|-------------------|
| $0,00 \leq TK \leq 0,30$ | Sukar |
| $0,31 \leq TK \leq 0,70$ | Sedang |
| $0,71 \leq TK \leq 1,00$ | Mudah |



| No Butir Baru | No Butir Asli | Jml Betul | Tkt. Kesukaran (%) | Tafsiran |
|---------------|---------------|-----------|--------------------|----------|
| 1 | 1 | 10 | 66,67 | Sedang |
| 2 | 2 | 12 | 80,00 | Mudah |
| 3 | 3 | 11 | 73,33 | Mudah |
| 4 | 4 | 11 | 73,33 | Mudah |
| 5 | 5 | 9 | 60,00 | Sedang |
| 6 | 6 | 12 | 80,00 | Mudah |
| 7 | 7 | 10 | 66,67 | Sedang |
| 8 | 8 | 10 | 66,67 | Sedang |
| 9 | 9 | 5 | 33,33 | Sedang |
| 10 | 10 | 11 | 73,33 | Mudah |

Gambar 7. Hasil Tingkat Kesukaran

Analisis kuantitatif terhadap indeks kesulitan soal, sebagaimana tertuang dalam gambar 7, mengindikasikan adanya heterogenitas tingkat kesulitan pada tes pilihan ganda ini adalah 0,33 hingga 0,80 dimana indeks tersebut termasuk kedalam tingkat kesukaran “sedang”.

Validitas, yang berasal dari kata bahasa Inggris “*validity*” merupakan konsep yang menunjukkan sejauh mana suatu alat ukur (tes) berhasil menjalankan fungsi utamanya. Suatu tes dikatakan valid apabila mampu memberikan hasil pengukuran yang benar-benar mencerminkan karakteristik atau sifat sebenarnya dari objek yang sedang diteliti (Fakhri Ramadhan et al., 2024). Validitas yang tinggi menjamin bahwa interpretasi terhadap skor tes dapat dilakukan secara tepat. Terdapat dua jenis validitas utama yakni validitas logis yang didasarkan pada pertimbangan teoritis dan logis, serta validitas empiris yang didasarkan pada data empiris. Butir soal dianggap memiliki validitas yang baik apabila memenuhi kriteria korelasi validitas instrumen yang telah ditetapkan sesuai dengan tabel 4.

Tabel 4: Kriteria Korelasi Validitas Butir Tes

| Koefisien Butir Soal | Interpretasi |
|----------------------|-----------------------------|
| 0,800-1,00 | Validasi soal sangat tinggi |
| 0,600-0,800 | Validasi soal tinggi |
| 0,400-0,600 | Validasi soal cukup |
| 0,200-0,400 | Validasi soal rendah |
| 0,00-0,200 | Validasi soal sangat rendah |

| No Butir Baru | No Butir Asli | Jml Betul | Tkt. Kesukaran (%) | Tafsiran |
|---------------|---------------|-----------|--------------------|----------|
| 1 | 1 | 10 | 66,67 | Sedang |
| 2 | 2 | 12 | 80,00 | Mudah |
| 3 | 3 | 11 | 73,33 | Mudah |
| 4 | 4 | 11 | 73,33 | Mudah |
| 5 | 5 | 9 | 60,00 | Sedang |
| 6 | 6 | 12 | 80,00 | Mudah |
| 7 | 7 | 10 | 66,67 | Sedang |
| 8 | 8 | 10 | 66,67 | Sedang |
| 9 | 9 | 5 | 33,33 | Sedang |
| 10 | 10 | 11 | 73,33 | Mudah |

Gambar 8. Hasil Korelasi skor butir dengan skor total (Validitas)

Berdasarkan uji kelayakan soal pada software ANATES hasil validitas butir tes Seperti yang ditunjukkan pada Gambar 8 dapat diketahui bahwa, soal nomor 1 korelasi validitas sebesar 0,618 artinya validitas soal tinggi, soal nomor 2 korelasi validitas sebesar 0,185 artinya validasi soal sangat tinggi, soal nomor 3 korelasi validitas sebesar 0,603 artinya validitas soal tinggi, soal nomor 4 korelasi validitas sebesar 0,071 artinya validitas soal sangat rendah, soal nomor 5 korelasi validitas sebesar 0,508 artinya validitas soal cukup, soal nomor 6 korelasi validitas sebesar 0,185 artinya validitas soal sangat tinggi, soal nomor 7 korelasi validitas sebesar 0,618 artinya validitas soal tinggi, soal nomor 8 korelasi validitas sebesar 0,333 artinya validitas soal rendah.

Pengecoh dirancang untuk mengidentifikasi Peserta didik yang telah mencapai pemahaman mendalam terhadap konsep yang diuji dan membedakannya dengan yang hanya menghafal. Dengan menyajikan beberapa pilihan jawaban yang salah namun tampak terlihat sama atau benar, pengecoh dapat mengidentifikasi Peserta didik yang telah mencapai pemahaman mendalam terhadap konsep yang diuji. rendahnya kualitas pengecoh menunjukkan bahwa alternatif jawaban tersebut tidak mampu mengklasifikasikan peserta didik yang belum mencapai kompetensi pemahaman konsep (Nafs et al., 2023). Suatu butir soal dikatakan efektif apabila distribusi pilihan pengecoh pada peserta didik yang menjawab salah relatif seimbang (Yusuf, 2021). Butir soal ini dinilai memiliki pengecoh yang efektif apabila memenuhi seluruh kriteria yang tertera pada Tabel 5

Tabel 5: Kriteria Kualitas Pengecoh Soal

| Kriteria | Kualitas Pengecoh Soal |
|----------|------------------------|
| ** | Kunci Jawaban |
| ++ | Sangat Baik |
| + | Baik |
| - | Kurang Baik |
| -- | Buruk |
| - - - | Sangat Buruk |

| No Butir Baru | No Butir Asli | a | b | c | d | e | * |
|---------------|---------------|------------------|------------------|------------------|------------------|------------------|---|
| 1 | 1 | 10 ⁺⁺ | 1 ⁺⁺ | 3 ⁻⁻ | 0 ⁻⁻ | 1 ⁺⁺ | 0 |
| 2 | 2 | 0 ⁻⁻ | 2 ⁻⁻ | 0 ⁻⁻ | 1 ⁺ | 12 ⁺⁺ | 0 |
| 3 | 3 | 1 ⁺⁺ | 11 ⁺⁺ | 2 ⁻⁻ | 1 ⁺⁺ | 0 ⁻⁻ | 0 |
| 4 | 4 | 11 ⁺⁺ | 3 ⁻⁻ | 1 ⁺⁺ | 0 ⁻⁻ | 0 ⁻⁻ | 0 |
| 5 | 5 | 2 ⁺ | 9 ⁺⁺ | 2 ⁺ | 2 ⁺ | 0 ⁻⁻ | 0 |
| 6 | 6 | 0 ⁻⁻ | 1 ⁺ | 1 ⁺ | 12 ⁺⁺ | 1 ⁺ | 0 |
| 7 | 7 | 0 ⁻⁻ | 1 ⁺⁺ | 10 ⁺⁺ | 4 ⁻⁻ | 0 ⁻⁻ | 0 |
| 8 | 8 | 0 ⁻⁻ | 2 ⁻⁻ | 0 ⁻⁻ | 10 ⁺⁺ | 3 ⁻⁻ | 0 |
| 9 | 9 | 3 ⁺⁺ | 0 ⁻⁻ | 1 ⁻ | 6 ⁻⁻ | 5 ⁺⁺ | 0 |
| 10 | 10 | 1 ⁺⁺ | 11 ⁺⁺ | 0 ⁻⁻ | 3 ⁻⁻ | 0 ⁻⁻ | 0 |

Gambar 9. Hasil Korelasi skor butir dengan skor total (Validitas)

Berdasarkan hasil uji ANATES, efektivitas pengecoh soal tervisualisasi dalam Gambar 9 dapat disimpulkan bahwa kualitas pengecoh pada setiap soal sangat bervariasi. Hal ini terlihat dari penilaian terhadap setiap opsi jawaban, di mana terdapat kombinasi antara pengecoh yang sangat baik, baik, buruk, dan sangat buruk. Secara umum, jumlah pengecoh yang dinilai buruk atau sangat buruk lebih banyak dibandingkan dengan pengecoh yang baik atau sangat baik. Dari 50 opsi jawaban pengecoh (5 opsi/soal x 10 soal), terdapat 23 opsi yang dinilai kurang efektif dalam mengalihkan perhatian peserta didik. Kemudian 27 opsi jawaban dinilai cukup baik dalam mengecoh soal.

SIMPULAN

Hasil penelitian menunjukkan rata-rata nilai tes 15 responden adalah 6,73. Reliabilitas soal cukup baik (0,49), namun 4 responden termasuk kelompok unggul dan 4 lainnya di kelompok bawah. Analisis Anates menunjukkan 8 dari 10 soal berhasil membedakan peserta dengan kemampuan tinggi dan rendah, dengan indeks daya beda $\geq 0,5$. Soal nomor 2 dan 4 perlu direvisi karena daya bedanya rendah. Tingkat kesukaran soal berada pada kategori sedang (0,33-0,80). Validitas soal juga tergolong cukup (0,33-0,62). Analisis kualitas pengecoh menunjukkan hasil yang beragam, dengan sebagian besar pengecoh dinilai buruk. Meskipun terdapat beberapa kekurangan, secara umum hasil penelitian ini menunjukkan bahwa soal tes yang digunakan sudah cukup baik. Namun, dengan melakukan beberapa perbaikan, kualitas soal tes dapat ditingkatkan sehingga dapat memberikan hasil yang lebih akurat dan dapat diandalkan.

DAFTAR PUSTAKA

- Dzil, I., Hartono, I., Besse Tenriawaru, A., Ningsih (2024). ANALISIS BUTIR SOAL PENILAIAN SUMATIF IPA MENGGUNAKAN ANATES. 8, 162–171. <https://doi.org/10.26418/jurnalkpk.v8i2.78282>
- Fakhri Ramadhan, M., Siroj, R. A., Win Afgani, M., Raden Fatah Palembang, U., H Zainal Abidin Fikri, J. K., Kemuning, K., Palembang, K., & Selatan, S. (2024). Validitas and Reliabilitas. *Journal on Education*, 06(02), 10967–10975.
- Francisca Lies Ambarwati, M., Adianti, D., & Tinggi Ilmu Komunikasi dan Sekretari Tarakanita, S. (2022). EFISIENSI DALAM PENGELOLAAN DOKUMEN BERBASIS DIGITAL.
- Herman, T., Hasanah, A., Candra Nugraha, R., Harningsih, E., Aghniya Ghassani, D., Marasabessy, R., & Studi Pendidikan Matematika, P. (2022). Pembelajaran Berbasis Masalah-High Order Thinking Skill (HOTS) pada Materi Translasi.
- Hidayatullah, R. S., Ariyanto, S. R., Hafizh, A., Rasyid, A., Azizah, M., & Artikel, S. (2023). *Journal of Mechanical Engineering a ANALISIS SOAL HOTS PILIHAN GANDA: ALAT PENILAIAN KOMPETENSI PADA TEKNIK MESIN I N F O ARTIKEL*. In *JMEL* (Vol. 12, Issue 1).
- Inom Nasution, Aji Pramudya, Amaluddin Tanjung, Dina Oktapua, Khoirun Nisa, Nindya Azzahrah & Nurdahyanti. (2023). Supervisi Pendidikan Era Society 5.0. *Jurnal Riset Pendidikan dan Bahasa*, 2(2), 119–120. <https://doi.org/10.61132/manuhara.v1i4.115>
- Mahgafhira, I., Wahid, A., Parenreng, J. M., Makassar, U. N., Pettarani, J. A. P., Makassar, K., & Selatan, S. (2023). Pengembangan Sistem Otentikasi Dokumen Digital Jurusan Teknik Informatika Dan Komputer Fakultas Teknik UNM Berbasis Digital Signature 1.
- Marambaawang, D. N., Oktoviana Bano, V., Rambu, R., & Enda, H. (2023). ANALISIS KUALITAS BUTIR SOAL PENILAIAN AKHIR SEMESTER GASAL TAHUN 2021/2022 MENGGUNAKAN ITEMAN DI SMP NEGERI 1 KAMBERA. *Juni*, 4(1), 233. http://ejournal.undhari.ac.id/index.php/de_journal
- Nafs, H., Sridana, N., Hikmah, N., & Soeprianto, H. (2023). Analisis Kualitas Butir Soal Ulangan Akhir Semester Genap Mata Pelajaran Matematika Kelas VII SMPN 6 Mataram Tahun Ajaran 2022/2023. *Jurnal Ilmiah Profesi Pendidikan*, 8(4), 2324–2331. <https://doi.org/10.29303/jipp.v8i4.1785>
- Putri Mulyani, A., Septyanti, E., Putri Mustika, T (2024). MOTEKAR: Jurnal Multidisiplin Teknologi dan Arsitektur Analisis Butir Tes Pilihan Ganda Berbantuan Software Anates Mata Kuliah Menyimak Mahasiswa PBSI Universitas Riau.
- Putri Andini, D., & Mukhlis, M. (2023). Analisis Butir Soal pada Instrumen Asesmen Kompetensi Minimum Literasi Membaca di SMP IT Insan Utama Pekanbaru. *Diglosia: Jurnal Kajian Bahasa, Sastra, Dan Pengajarannya*, 6(2), 401–412. <https://doi.org/10.30872/diglosia.v6i2.658>
- Puspa Hanan, M., Raudhatul Jannah, R., & Alexander Alim, J. (2023). ANALISIS BUTIR SOAL MATEMATIKA MATERI KPK DAN FPB BERBASIS CERITA DI SDN 111 PEKANBARU. *Journal of Educational Learning and Innovation*, 3(1), 92. <https://doi.org/10.46229/elia.v3i1>
- Ningrum, W. A., Rahmawati, R. C., Minarti, I. B., & Mekar, R. M. (2023). Analisis Butir Soal Ulangan Harian Pembelajaran IPA Pada Kelas VIII Di SMPN 21 Semarang. *Educatio*, 18(1), 91–101. <https://doi.org/10.29408/edc.v18i1.18291>

- Ramadhan, W., Malahati, F., Romadhon, K., & Ramadhan, S. (2023). Analisis Butir Soal Tipe Multiple Choice Questions pada Penilaian Harian Sekolah Dasar. *Tarbiyah Wa Ta'lim: Jurnal Penelitian Pendidikan Dan Pembelajaran*, 10(2), 93–105. <https://doi.org/10.21093/twt.v10i2.6155>
- Sijabat, M. P., Hutabarat, K., Sitorus, L., & Syahrial, S. (2024). Analisis Soal Tes Hasil Belajar Siswa Soal Berstandar Nasional Bahasa Indonesia Kelas 5 Sekolah Dasar. *Jurnal Basicedu*, 8(2), 1265–1277. <https://doi.org/10.31004/basicedu.v8i2.7354>
- Yusuf, R. (2021). ANALISIS KUALITAS BUTIR SOAL UJIAN NASIONAL MATA PELAJARAN MATEMATIKA SEKOLAH MENENGAH PERTAMA. *Journal of Didactic Mathematics*, 1(3). <https://doi.org/10.34007/jdm.v1i3.417>