

ANALISIS BUTIR SOAL DAN KUALITAS TES

Emrida Harahap¹, Marnis Khairati Hasibuan², Rahmat Saleh Rangkuti³, Syafnan⁴
^{1,2,3,4}Universitas Islam Negeri Syekh Ali Hasan Ahmad Addary Padangsidimpuan

¹emridahrp@uinsyahada.ac.id, ²marnishsb@uinsyahada.ac.id,

³salehrangkuti@uinsyahada.ac.id, ⁴syafnan@uinsyahada.ac.id

ABSTRACT

Item analysis is an essential step in educational evaluation to ensure the quality of test instruments used to measure students' learning outcomes. This study aims to examine the concept of item analysis and test quality as a foundation for developing valid, reliable, and objective evaluation instruments. The research employed a library research method with a qualitative descriptive approach. Data were collected from various literature sources, including books, scientific journals, and research findings relevant to item analysis and educational evaluation. The collected data were then analyzed using content analysis techniques. The findings indicate that item analysis plays a significant role in identifying the quality of each test item through validity, reliability, and difficulty level testing. A high-quality test should meet the criteria of validity, reliability, objectivity, balanced difficulty levels, good discrimination power, and practicality in its application. By conducting systematic item analysis, teachers can improve the quality of evaluation instruments, resulting in more accurate, fair, and trustworthy assessment outcomes. Therefore, item analysis should be carried out continuously to support the improvement of learning quality and the effectiveness of student assessment results.

Keywords: Item Analysis; Test Quality; Validity; Reliability; Educational Evaluation

ABSTRAK

Analisis butir soal merupakan salah satu langkah penting dalam evaluasi pembelajaran untuk memastikan kualitas instrumen tes yang digunakan dalam mengukur hasil belajar peserta didik. Penelitian ini bertujuan untuk mengkaji konsep analisis butir soal dan kualitas tes sebagai dasar dalam penyusunan instrumen evaluasi yang valid, reliabel, dan objektif. Metode yang digunakan adalah studi kepustakaan (library research) dengan pendekatan deskriptif kualitatif. Data diperoleh dari berbagai sumber literatur, seperti buku, jurnal ilmiah, dan hasil penelitian yang relevan dengan analisis butir soal dan evaluasi pembelajaran. Data kemudian dianalisis menggunakan teknik analisis isi (content analysis). Hasil kajian menunjukkan bahwa analisis butir soal berperan penting dalam mengidentifikasi kualitas setiap butir soal melalui pengujian validitas, reliabilitas, dan tingkat

kesukaran. Tes yang berkualitas harus memenuhi kriteria valid, reliabel, objektif, memiliki tingkat kesukaran yang seimbang, daya pembeda yang baik, serta praktis digunakan. Dengan melakukan analisis butir soal secara sistematis, guru dapat memperbaiki kualitas instrumen evaluasi sehingga hasil penilaian menjadi lebih akurat, adil, dan dapat dipercaya. Oleh karena itu, analisis butir soal perlu dilakukan secara berkelanjutan untuk mendukung peningkatan mutu pembelajaran dan kualitas hasil evaluasi peserta didik.

Kata Kunci: Analisis Butir Soal; Kualitas Tes; Validitas, Reliabilitas; Evaluasi Pembelajaran

A. Pendahuluan

Evaluasi merupakan salah satu komponen penting dalam proses pembelajaran, termasuk dalam pembelajaran Pendidikan Agama Islam. Melalui evaluasi, guru dapat mengetahui sejauh mana tujuan pembelajaran telah tercapai serta memahami tingkat penguasaan materi oleh peserta didik. Salah satu bentuk evaluasi yang sering digunakan adalah tes.

Tes merupakan salah satu instrumen penting dalam menilai tingkat keberhasilan peserta didik pada proses pembelajaran sejarah. Karena perannya yang sangat strategis, diperlukan butir-butir soal yang berkualitas agar tes yang diberikan mampu menghasilkan pengukuran yang akurat dan dapat dipercaya. Untuk memastikan bahwa soal yang disusun guru telah memenuhi kriteria kelayakan dan

mampu mengukur pemahaman peserta didik secara optimal, perlu dilakukan analisis terhadap setiap butir soal pilihan ganda. Melalui analisis tersebut, guru dapat mengetahui mutu masing-masing soal sehingga dapat melakukan perbaikan terhadap soal yang kurang efektif atau menghapus soal yang tidak layak digunakan. Selain itu, hasil analisis juga memberikan informasi mengenai tingkat penguasaan materi peserta didik terhadap materi yang telah diajarkan. Analisis butir soal pada dasarnya merupakan proses evaluasi untuk menilai kualitas setiap soal, baik dalam mengukur kemampuan kognitif maupun dalam membedakan peserta didik berdasarkan tingkat penguasaan materi. Dengan demikian, analisis butir soal menjadi langkah penting untuk menentukan soal yang layak dipertahankan, direvisi, atau bahkan dibuang demi meningkatkan kualitas

alat evaluasi pembelajaran (Basri & Karima, 2023: 21162).

Dalam konteks pendidikan modern, evaluasi pembelajaran tidak lagi hanya berorientasi pada hasil akhir, tetapi juga pada kualitas instrumen yang digunakan. Penyusunan instrumen evaluasi harus dilakukan secara cermat karena kualitas instrumen akan menentukan ketepatan hasil evaluasi. Jika instrumen yang digunakan kurang tepat, maka hasil evaluasi yang diperoleh juga tidak akurat. Oleh karena itu, guru dan pengembang tes perlu melakukan analisis terhadap setiap butir soal sebelum digunakan secara luas (Jannah, 2024: 332).

Namun, penelitian terbaru menunjukkan masih banyak ditemukan soal-soal yang belum memenuhi kriteria kualitas yang baik. Beberapa soal mungkin terlalu mudah atau terlalu sulit, tidak mampu mengukur kompetensi yang diharapkan, atau bahkan menghasilkan nilai yang tidak konsisten. Hal ini tentu dapat mempengaruhi keakuratan hasil evaluasi dan berdampak pada pengambilan keputusan dalam

pembelajaran (Abdullah et al., 2025: 28).

Oleh karena itu, diperlukan suatu proses yang disebut analisis butir soal. Analisis ini bertujuan untuk mengkaji setiap butir soal agar dapat diketahui kualitasnya, baik dari segi validitas, reliabilitas, maupun tingkat kesukaran. Dengan adanya analisis ini, guru dapat memperbaiki atau mengganti soal yang kurang baik sehingga tes yang digunakan benar-benar mampu mengukur kemampuan peserta didik secara tepat dan objektif.

Selain itu, dalam konteks Pendidikan Agama Islam, evaluasi tidak hanya bertujuan untuk mengukur aspek kognitif, tetapi juga harus mampu mendukung pembentukan sikap dan nilai-nilai keislaman. Oleh karena itu, penyusunan dan analisis soal harus dilakukan secara cermat agar hasil evaluasi dapat memberikan gambaran yang menyeluruh tentang perkembangan peserta didik. Dengan demikian, analisis butir soal dan kualitas tes menjadi hal yang sangat penting untuk dilakukan guna meningkatkan mutu pembelajaran serta menghasilkan evaluasi yang akurat, adil, dan dapat dipercaya.

B. Metode Penelitian

Penelitian ini menggunakan metode studi kepustakaan (library research) dengan pendekatan deskriptif kualitatif. Metode studi kepustakaan dipilih karena data yang digunakan berasal dari berbagai sumber literatur yang relevan dengan topik analisis butir soal dan kualitas tes, seperti buku, jurnal ilmiah, artikel penelitian, tesis, dan dokumen akademik lainnya. Teknik pengumpulan data dilakukan melalui kegiatan identifikasi, pengkajian, dan pencatatan berbagai referensi yang berkaitan dengan konsep analisis butir soal, validitas, reliabilitas, tingkat kesukaran soal, serta karakteristik tes yang berkualitas.

Data yang diperoleh kemudian dianalisis menggunakan teknik analisis isi (content analysis), yaitu dengan menelaah, membandingkan, dan menginterpretasikan informasi dari berbagai sumber untuk memperoleh pemahaman yang komprehensif mengenai analisis butir soal dan kualitas tes dalam evaluasi pembelajaran. Hasil analisis selanjutnya disajikan secara deskriptif untuk memberikan gambaran yang sistematis mengenai pentingnya

analisis butir soal dalam meningkatkan kualitas instrumen evaluasi dan efektivitas proses pembelajaran.

C. Hasil dan Pembahasan

Pengertian Analisis Butir Soal

Analisis butir soal adalah proses mengkaji dan menilai kualitas setiap butir soal dalam suatu tes secara sistematis. Kegiatan ini bertujuan untuk mengetahui apakah soal yang digunakan sudah sesuai dengan tujuan pembelajaran, mampu mengukur kemampuan siswa dengan tepat, serta layak digunakan dalam evaluasi. Jika ditemukan kelemahan, maka soal tersebut dapat diperbaiki atau diganti agar kualitas tes menjadi lebih baik (Tarumasely, 2024: 94).

Analisis butir soal tidak hanya dilakukan dengan melihat isi soal saja, tetapi juga berdasarkan hasil jawaban siswa. Dengan demikian, analisis ini bersifat objektif karena menggunakan data nyata dari hasil pengerjaan peserta didik. Melalui proses ini, guru dapat mengetahui bagaimana karakteristik setiap soal, apakah terlalu mudah, terlalu sulit, atau tidak mampu membedakan kemampuan

siswa (Bungai et al., 2025: 151).

Tujuan Analisis Butir Soal:

1. Mengetahui kualitas soal

Analisis dilakukan untuk memastikan bahwa setiap soal memiliki kualitas yang baik, baik dari segi isi, bahasa, maupun tingkat kesesuaiannya dengan materi yang diajarkan. Soal yang berkualitas akan mampu mengukur kompetensi siswa secara tepat.

2. Mengidentifikasi soal yang baik dan kurang baik

Dengan analisis, guru dapat membedakan soal yang sudah layak digunakan dan soal yang masih memiliki kekurangan. Soal yang baik dapat dipertahankan, sedangkan soal yang kurang baik dapat diperbaiki atau tidak digunakan lagi.

3. Meningkatkan mutu evaluasi pembelajaran

Analisis butir soal membantu guru dalam menyusun tes yang lebih berkualitas di masa mendatang. Dengan demikian, proses evaluasi menjadi lebih akurat, adil, dan dapat dipercaya sebagai dasar dalam menilai hasil belajar siswa (Amran Hapsan, 2024: 162).

Validitas Tes

Validitas tes merupakan salah satu aspek penting dalam menentukan kualitas suatu tes. Validitas menunjukkan sejauh mana suatu tes mampu mengukur apa yang seharusnya diukur. Dengan kata lain, tes yang valid adalah tes yang benar-benar sesuai dengan tujuan pembelajaran dan materi yang telah diajarkan kepada peserta didik.

Jika suatu tes memiliki validitas yang tinggi, maka hasil yang diperoleh dapat dipercaya karena mencerminkan kemampuan sebenarnya dari siswa. Sebaliknya, jika tes tidak valid, maka hasilnya tidak akurat dan dapat menyesatkan dalam pengambilan keputusan, seperti penentuan nilai atau evaluasi pembelajaran (Abdullah, 2025: 10).

1. Jenis-jenis Validitas Tes

a. Validitas Isi (Content Validity)

Validitas isi berkaitan dengan kesesuaian antara soal dengan materi atau kurikulum yang diajarkan. Soal harus mencakup indikator pembelajaran yang telah ditetapkan.

b. Validitas Konstruk (Construct Validity)

Validitas ini menunjukkan sejauh mana soal mampu mengukur aspek atau kemampuan tertentu,

seperti pemahaman, sikap, atau keterampilan sesuai dengan teori yang digunakan.

c. Validitas Empiris (Empirical Validity)

Validitas ini diperoleh melalui uji coba soal kepada siswa, kemudian dianalisis menggunakan teknik statistik, seperti korelasi (Sumardi, 2021: 77).

2. Cara Menentukan Validitas Tes

Salah satu cara yang sering digunakan adalah dengan rumus korelasi Product Moment:

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

Keterangan:

r_{xy} = koefisien validitas

X = skor butir soal

Y = skor total

N = jumlah peserta

Kriteria Penilaian Validitas

Jika $r_{hitung} > r_{tabel} \rightarrow$ soal valid

Jika $r_{hitung} < r_{tabel} \rightarrow$ soal tidak valid

Validitas tes sangat penting karena menentukan ketepatan hasil evaluasi. Semakin tinggi validitas suatu tes, maka semakin tepat tes tersebut dalam mengukur kemampuan peserta didik (Ketaren et al., 2024: 3283). Oleh karena itu, setiap soal perlu diuji validitasnya

sebelum digunakan dalam proses pembelajaran.

Reliabilitas Tes

Reliabilitas tes merupakan tingkat konsistensi atau keajegan suatu tes dalam mengukur kemampuan peserta didik. Suatu tes dikatakan reliabel apabila memberikan hasil yang relatif sama ketika digunakan berulang kali dalam kondisi yang sama. Dengan kata lain, reliabilitas menunjukkan sejauh mana suatu tes dapat dipercaya.

Tes yang reliabel sangat penting dalam evaluasi pembelajaran, karena hasilnya dapat dijadikan dasar yang kuat dalam pengambilan keputusan. Jika suatu tes tidak reliabel, maka hasil yang diperoleh bisa berubah-ubah dan tidak mencerminkan kemampuan sebenarnya dari peserta didik (Magdalena, 2022: 111).

1. Ciri-ciri Tes yang Reliabel

- a. Hasilnya konsisten dari waktu ke waktu
- b. Tidak dipengaruhi oleh faktor kebetulan
- c. Memberikan hasil yang stabil
- d. Dapat dipercaya sebagai alat ukur (Rahmat, 2019: 90).

2. Cara Mengukur Reliabilitas Tes

Salah satu cara yang sering digunakan adalah dengan rumus Alpha Cronbach:

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

Keterangan:

r_{11} = koefisien reliabilitas

k = jumlah butir soal

$\sum \sigma_i^2$ = jumlah varians tiap butir

σ_t^2 = varians total

Interpretasi Nilai Reliabilitas

0,80 – 1,00 → sangat tinggi

0,60 – 0,79 → tinggi

0,40 – 0,59 → sedang

0,20 – 0,39 → rendah

0,00 – 0,19 → sangat rendah

Reliabilitas tes menunjukkan tingkat kepercayaan terhadap hasil tes. Semakin tinggi nilai reliabilitas, maka semakin konsisten dan dapat dipercaya tes tersebut (Forester et al., 2024: 1820). Oleh karena itu, selain valid, suatu tes juga harus reliabel agar hasil evaluasi benar-benar akurat dan objektif.

Tingkat Kesukaran Soal

Tingkat kesukaran soal merupakan salah satu indikator penting dalam analisis butir soal yang digunakan untuk mengetahui apakah suatu soal tergolong mudah, sedang, atau sulit bagi peserta didik. Tingkat

kesukaran ini menunjukkan seberapa besar kemampuan siswa dalam menjawab suatu soal dengan benar.

Soal yang baik tidak hanya dilihat dari benar atau salahnya jawaban siswa, tetapi juga dari proporsi siswa yang dapat menjawab soal tersebut. Jika terlalu banyak siswa yang menjawab benar, maka soal tersebut tergolong mudah. Sebaliknya, jika hanya sedikit siswa yang mampu menjawab dengan benar, maka soal tersebut tergolong sulit (Maulani et al., 2024: 324).

1. Rumus Tingkat Kesukaran

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

Keterangan:

P = indeks kesukaran

B = jumlah siswa yang menjawab benar

N = jumlah seluruh siswa

2. Kriteria Tingkat Kesukaran

$P > 0,70$ → soal mudah

$0,30 \leq P \leq 0,70$ soal sedang

$P < 0,30$ → soal sulit

3. Penjelasan

a. Soal mudah: hampir semua siswa dapat menjawab dengan benar, sehingga kurang mampu membedakan kemampuan siswa.

b. Soal sedang: sebagian siswa dapat menjawab dengan benar,

sehingga dianggap paling ideal untuk digunakan dalam tes.

- c. Soal sulit: hanya sedikit siswa yang dapat menjawab, biasanya karena materi belum dikuasai atau soal terlalu kompleks (Bungai, 2025: 131).

Tingkat kesukaran soal sangat berpengaruh terhadap kualitas tes. Soal yang baik adalah soal dengan tingkat kesukaran sedang, karena mampu mengukur kemampuan siswa secara lebih akurat dan membedakan antara siswa yang memahami materi dan yang belum (Sumardi, 2021: 59). Oleh karena itu, dalam penyusunan tes perlu adanya keseimbangan antara soal mudah, sedang, dan sulit.

Kualitas Tes yang Baik

Kualitas tes yang baik merupakan gambaran bahwa suatu tes telah memenuhi kriteria sebagai alat evaluasi yang efektif dan dapat dipercaya. Tes yang berkualitas tidak hanya mampu mengukur hasil belajar siswa, tetapi juga memberikan informasi yang akurat untuk pengambilan keputusan dalam proses pembelajaran (Maulani et al., 2024: 129). Suatu tes dikatakan berkualitas apabila memenuhi beberapa syarat utama, yaitu sebagai berikut:

1. Valid

Tes harus mampu mengukur apa yang seharusnya diukur sesuai dengan tujuan pembelajaran. Soal yang diberikan harus relevan dengan materi yang diajarkan dan indikator yang telah ditetapkan, sehingga hasilnya benar-benar mencerminkan kemampuan siswa.

2. Reliabel

Tes harus memiliki tingkat konsistensi yang tinggi. Artinya, jika tes dilakukan berulang kali dalam kondisi yang sama, maka hasilnya relatif tetap. Hal ini menunjukkan bahwa tes tersebut dapat dipercaya.

3. Objektif

Tes harus bebas dari unsur subjektivitas penilai. Penilaian harus didasarkan pada kriteria yang jelas, sehingga siapapun yang menilai akan menghasilkan skor yang sama.

4. Tingkat Kesukaran yang Seimbang

Tes yang baik memiliki variasi tingkat kesukaran, yaitu soal mudah, sedang, dan sulit. Namun, proporsi soal sedang sebaiknya lebih banyak karena mampu mengukur kemampuan siswa secara optimal.

5. Daya Pembeda yang Baik

Tes harus mampu membedakan antara siswa yang memiliki kemampuan tinggi dan rendah. Soal yang baik akan dijawab benar oleh siswa yang pandai dan salah oleh siswa yang kurang mampu.

6. Praktis dan Efisien

Tes mudah digunakan, tidak memerlukan waktu yang terlalu lama, serta mudah dalam proses pemeriksaan dan pengolahan hasilnya (Juliandi et al., 2020: 39).

Kualitas tes yang baik sangat menentukan keakuratan hasil evaluasi pembelajaran. Tes yang memenuhi kriteria valid, reliabel, objektif, serta memiliki tingkat kesukaran dan daya pembeda yang baik akan menghasilkan penilaian yang adil dan dapat dipercaya. Oleh karena itu, guru perlu memperhatikan kualitas tes dalam setiap proses evaluasi agar tujuan pembelajaran dapat tercapai secara maksimal.

D. Kesimpulan

Berdasarkan pembahasan, dapat disimpulkan bahwa analisis butir soal merupakan langkah penting dalam menjamin kualitas tes sebagai alat evaluasi pembelajaran. Melalui analisis validitas, reliabilitas, dan

tingkat kesukaran, guru dapat mengetahui kualitas setiap butir soal sehingga mampu menyusun instrumen evaluasi yang lebih akurat, objektif, dan dapat dipercaya. Tes yang berkualitas harus memenuhi kriteria valid, reliabel, memiliki tingkat kesukaran yang seimbang, daya pembeda yang baik, serta praktis dalam penggunaannya. Oleh karena itu, guru disarankan untuk melakukan analisis butir soal secara berkala sebelum dan sesudah tes digunakan, meningkatkan kompetensi dalam penyusunan instrumen evaluasi, serta memperhatikan aspek kognitif, afektif, dan psikomotorik dalam penyusunan soal. Selain itu, pihak sekolah perlu mendukung peningkatan kemampuan guru melalui pelatihan dan pendampingan terkait evaluasi pembelajaran agar kualitas penilaian dan mutu pembelajaran dapat terus ditingkatkan.

DAFTAR PUSTAKA

Abdullah, I., Listyotami, M. K., Ferita, R. A., Maulida, D., Yanti, W., & Hidayat, F. (2025). Analisis Kualitas Butir Soal Asesmen Sumatif Matematika Pilihan Ganda Semester Genap Kelas VIII MTs Baitul Ma' mur Tahun Ajaran 2023 / 2024 Analysis of the Quality of Multiple-Choice

- Mathematics Summative Assessment Items for the Even Semester of Grade V. *NUES: Nusantara Journal of Education and Social Science*, 2(1), 27–32. <https://doi.org/2.https://doi.org/10.69959/nujess.v2n1.49>.
- Abdullah, G. (2025). *Peran evaluasi dalam meningkatkan kualitas pembelajaran*. CV Rey Media Grafika.
- Amran Hapsan, N. D. A. T. (2024). *Prinsip-prinsip pengukuran dan evaluasi pendidikan: Disertai dengan contoh kasus*. CV. Ruang Tentor.
- Basri, Y. F., & Karima, E. M. (2023). Kualitas Butir Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah Terhadap Hasil Belajar Kognitif Siswa di Fase E SMK Negeri 1 Painan Program Studi Pendidikan Sejarah , Universitas Negeri Padang. *Jurnal Pendidikan Tambusai*, 7, 21160–21171. <https://doi.org/https://doi.org/10.31004/jptam.v7i3.9639>.
- Bungai, J., et al. (2025). *Evaluasi pembelajaran di sekolah dasar*. CV Eureka Media Aksara.
- Forester, B. J., et al. (2024). Penelitian kuantitatif: Uji reliabilitas. *Edu Society: Jurnal Pendidikan, Ilmu Sosial dan Pengabdian kepada Masyarakat*, 4(3), 1820. <https://doi.org/10.56832/edu.v4i3.577>.
- Jannah, R. (2024). Analisis Kualitas Butir Soal dan Abilitas Siswa pada Asesmen Madrasah Ibtidaiyah dengan Rasch Model. *Jurnal Pendidikan Dasar*, 8(1), 331–342. <https://doi.org/10.26811/didaktika.v8i1.1234>.
- Juliandi, A., et al. (2020). *Metodologi penelitian bisnis, konsep dan aplikasi: Sukses menulis skripsi & tesis mandiri*. UMSU Press.
- Ketaren, M. A., et al. (2024). Uji validitas dan uji daya beda soal buatan pilihan ganda dengan tes sumatif siswa kelas IV UPT SD Negeri 065013 Medan. *Jurnal Pengabdian Masyarakat Bangsa*, 1(11), 3283. <https://doi.org/10.59837/jpmba.v1i11.706>
- Magdalena, I. (2022). *Teori dan praktik evaluasi pembelajaran SD*. CV Jejak.
- Maulani, G., et al. (2024). *Evaluasi pembelajaran*. Sada Kurnia Pustaka.
- Rahmat. (2019). *Evaluasi pembelajaran pendidikan agama Islam*. Bening Pustaka.
- Sumardi. (2021). *Teknik pengukuran dan penilaian hasil belajar*. Deepublish.
- Tarumasely, Y. (2024). *Buku ajar strategi pembelajaran*. Academia Publication.