JURNAL RISET DAN EVALUASI PENDIDIKAN
Vol. 2 No. 1 (2025) pp 438-449

Analisis Kualitas Butir Soal Penilaian Akhir Semester Mata Pelajaran Seni Budaya di SD Shibghah Akhlak Qur'an Bekasi

Received: 24/01/2025
Accepted: 15/02/2025
Published: 01/03/2025

[1]Hilmiyah Ridwan, [2]Sigit Muryono
[1,2] Jurusan Penelitian dan Evaluasi Pendidikan Program Magister Universitas Prof. Dr. Hamka, Indonesia

[1]hilmiyah.hr@gmail.com
[2]msigituryono60@gmail.com

Abstract
Evaluation ensures that the teaching and learning processes align with the desired objectives by using test items (instruments) to measure students' competency progress. Initial observations indicate that the test items in the semester-end exams have not undergone thorough analysis, resulting in unknown quality. This study examines the quality of the semester-end test items for the Art and Culture subject for 11th-grade students at SD Shibghah Akhlak Qur'an Bekasi in the 2024/2025 academic year, considering difficulty level, discrimination power, and distractor effectiveness. The researchers employed a descriptive quantitative method to analyze the test items' quality. The sample included 72 students, with data collected through documentation and interviews, and analyzed using ANATES V4 and Microsoft Excel. The results showed that out of 40 test items, 10% were of very high quality and retained, 17.5% were good, 45% were satisfactory and archived for revision, 12.5% were poor, and 12.5% were very poor and thus deleted or replaced. Overall, the analysis of the test items indicates many items are not optimal, necessitating improvements and revisions in test items to ensure more accurate and meaningful evaluation outcomes.

Keywords: *Item Analysis, Validity, Reliability*

Abstrak
Evaluasi diterapkan untuk memastikan bahwa proses pengajaran dan pembelajaran sesuai dengan tujuan yang diinginkan dengan menggunakan butir soal (instrumen) sebagai pengukur kemajuan kompetensi peserta didik. Observasi awal menunjukkan bahwa butir soal dalam Penilaian akhir semester belum dianalisis secara menyeluruh, sehingga kualitasnya tidak dapat diketahui. Penelitian ini menggambarkan kualitas butir soal Penilaian Akhir Semester Gasal Mata Pelajaran Seni Budaya Kelas XI SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025 dengan mempertimbangkan tingkat kesukaran, daya beda, dan efektivitas pengecoh. Metode deskriptif kuantitatif digunakan untuk menganalisis kualitas daripada butir soal. Adapun sampel yang digunakan sebanyak 72 peserta didik. Data dikumpulkan melalui dokumentasi dan wawancara, dianalisis menggunakan ANATES V4 dan *Microsoft Excel*. Hasil penelitian menunjukkan dari 40 butir soal, 10% dengan kualitas sangat baik disimpan, 17,5% baik, 45% cukup baik diarsipkan untuk revisi, 12,5% tidak baik, dan 12,5% sangat tidak baik dihapus atau diganti. Keseluruhan analisis butir soal menunjukkan banyaknya butir soal belum optimal, sehingga proses pembelajaran dan evaluasi membutuhkan perbaikan dan revisi terkait butir soal Penilaian agar hasil evaluasi menjadi lebih akurat dan bermakna.

Kata kunci: Analisis Butir Soal, Validitas, Reliabilitas

Pendahuluan

Pendidikan adalah proses bimbingan atau bantuan yang diberikan oleh orang dewasa kepada anak-anak untuk membantu mereka mencapai kematangan, sehingga anak-anak dapat mengatasi tugas-tugas hidup mereka secara mandiri tanpa memerlukan bantuan orang lain. Sesuai dengan Fungsi dan Tujuan Pendidikan Nasional No. 20 Tahun 2003 tentang Sistem Pendidikan Nasional yang menyatakan bahwa pendidikan berfungsi mengembangkan kemampuan dan membentuk watak serta peradaban bangsa yang dalam rangka mencerdaskan kehidupan bangsa, bertujuan untuk mengembangkan potensi peserta didik agar menjadi manusia yang beriman dan bertakwa kepada Tuhan Yang Maha Esa, berakhlak mulia, sehat, berilmu, cakap, kreatif, mandiri, dan warga menjadi negara yang demokratis serta bertanggung jawab. Pendidikan menjadi upaya prioritas negara Indonesia dalam mengentaskan bangsa dari keterbelakangan dan ketertinggalan dari bangsa lain.

Proses belajar mengajar yang efektif merupakan kunci utama dalam mencapai tujuan pendidikan dan mengantarkan peserta didik menuju keberhasilan. Diperlukan visi, misi, dan tujuan yang jelas dan terukur untuk melaksanakan proses pembelajaran yang terarah dan sistematis. Untuk itu pelaksanaan evaluasi diperlukan sebagai tolok ukur untuk mengukur sejauh mana pemahaman materi yang telah dicapai oleh peserta didik. Evaluasi pendidikan adalah proses identifikasi untuk menentukan apakah program yang direncanakan telah mencapai tujuannya, apakah memiliki nilai atau tidak, dan untuk menilai tingkat efisiensi pelaksanaannya (Riinawati, 2021, p. 15). Evaluasi dapat diartikan sebagai proses menilai dengan melakukan pengukuran terlebih dahulu. Dengan kata lain, evaluasi dalam pembelajaran terdiri dari dua kegiatan yaitu pengukuran dan penilaian yang digunakan untuk mendapatkan hasil dari evaluasi dalam pembelajaran. Pengukuran bertujuan untuk mengumpulkan data atau informasi dan penilaian digunakan untuk menganalisis data atau informasi yang telah dikumpulkan dan memberikan penilaian terhadap pencapaian peserta didik. Dengan demikian, evaluasi tidak hanya bertujuan untuk mengukur efektivitas peserta didik dalam perjalanan pembelajarannya, namun juga berperan sebagai alat untuk mengambil keputusan tentang bagaimana pembelajaran itu berlangsung, baik yang sedang berlangsung maupun yang akan datang.

Permendikbud No. 69 (2013) tentang struktur kurikulum SD/MA, seni budaya termasuk salah satu mata pelajaran wajib yang harus diajarkan di sekolah. Menurut Monica (2019, p. 2), keberadaan mata pelajaran seni budaya disekolah dapat menanamkan nilai-nilai pendidikan dan sikap budaya, yaitu sikap saling menghargai, dan kerjasama dalam menghadapi kehidupan bermasyarakat. Selain itu, pendidikan seni berfungsi sebagai salah satu media untuk mengembangkan kemampuan kognitif, afektif, dan psikomotorik peserta didik di bidang seni (Putra et al., 2023, p. 1020). Dalam pelaksanaan proses pembelajarannya, mata pelajaran ini erat hubungannya dengan kegiatan apresiasi dan ekspresi yang berfungsi sebagai cara untuk mengembangkan pengetahuan dan kemampuan peserta didik dalam seni. Setelah serangkaian proses pembelajaran dilakukan, selanjutnya pendidik melakukan penilaian untuk mengukur

mana peserta didik mampu mengembangkan kemampuannya sesuai dengan apa yang telah dipelajari (Warju et al., 2020, p. 95). Penilaian pendidikan adalah proses mengumpulkan dan mengolah informasi untuk menentukan pencapaian hasil belajar peserta didik (Qodir, 2017, p. 5). Dalam kaitannya dengan proses belajar mengajar, terdapat dua jenis metode yang digunakan untuk menilai hasil belajar peserta didik di sekolah: tes dan non-tes. Tes merupakan kumpulan informasi yang bersifat resmi dibandingkan dengan instrumen lainnya karena adanya batasan-batasan tertentu (Afifah, 2020, p. 2). Tes biasanya digunakan untuk menilai dan mengukur prestasi belajar peserta didik, terutama dalam hal hasil pembelajaran kognitif seperti penguasaan mata pelajaran atau persiapan untuk pelajaran yang diajarkan. Salah satu metode evaluasi yang digunakan untuk memancarkan keterampilan peserta didik adalah ujian akhir semester, yang mencakup kedua komponen tertulis seperti pertanyaan dan esai pilihan ganda, serta penilaian lisan. Instrumen penilaian pilihan ganda digunakan secara luas karena mencakup berbagai materi dan merangsang kemampuan kognitif peserta didik, termasuk mengingat, memahami, menerapkan, menganalisis, mensintesis, dan memutar (Komarudin & Sarkadi, 2017, p. 50).

Evaluasi harus disusun untuk menilai dengan jelas kemampuan atau keterampilan spesifik, konten yang akan dievaluasi, alat evaluasi, dan interpretasi hasil evaluasi (Ropii & Fahrurrozi, 2017, p. 16). Instrumen penilaian atau dalam hal ini butir soal, perlu untuk dijelaskan kualitasnya agar dapat memberikan hasil evaluasi yang akurat. Analisis butir soal menurut Muluki (2020, p. 42) diartikan sebagai proses metode yang akan menghasilkan informasi yang sangat rinci tentang butir soal. Sehingga dapat kita pahami dengan jelas bah wa analisis butir soal adalah identifikasi terhadap pertanyaan yang disiapkan dan jawaban peserta didik yang memberikan informasi tentang hal itu (Safira, 2016, p. 14). Menurut Suharsimi (dalam Rahmasari & Ismiyati, 2016, p. 318) ” tujuan analisis soal dalam tes guru adalah untuk mengidentifikasi soal-soal yang baik, kurang baik, dan jelek. Analisis ini juga dapat memberikan "petunjuk" untuk perbaikan”. Kualitas butir soal dinyatakan baik apabila memenuhi sejumlah karakteristik yang meliputi validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh.

Berdasarkan pengamatan di SD Shibghah Akhlak Qur’an Bekasi, ditemukan bahwa pelaksanaan UAS mata pelajaran seni budaya dilakukan dengan menggunakan butir soal pilihan ganda sebanyak 40 butir yang disusun oleh guru, terdapat ketimpangan yang terlihat karena hasil UAS menunjukkan sebagian peserta didik tidak mencapai kriteria Ketuntasan Minimal (KKM) meskipun kegiatan belajar mengajar telah dilaksanakan dengan baik, serta ditemukan beberapa butir soal yang kurang sesuai untuk mencakup tujuan pembelajaran. Selain itu, guru mata pelajaran Seni Budaya di SD Shibghah Akhlak Qur’an Bekasi kelas XIknown bukan merupakan lulusan pendidikan seni dan belum pernah melakukan analisis kualitas pada butir soal UAS yang telah disusun dikarenakan: a) guru belum mengetahui bagaimana cara menganalisis butir soal, dan b) guru belum pernah mengikuti pelatihan tentang analisis

butir soal. Hal ini menyoroti pentingnya evaluasi lebih lanjut terhadap kualitas butir soal yang digunakan. Selama ini, guru mata pelajaran Seni Budaya di SD Shibghah Akhlak Qur'an Bekasi hanya menilai kualitas butir soal berdasarkan jawaban peserta dan tingkat kesulitan soal berdasarkan jumlah peserta yang menjawab dengan benar. Hal ini dapat menyebabkan panas mengenai kualitas perangkat tes sebagai alat evaluasi peserta didik. Adapun rumusan masalah dalam penelitian ini adalah bagaimana kualitas butir soal Penilaian Akhir Semester Gasal mata pelajaran Seni Budaya kelas XI SD Shibghah Akhlak Qur'an Bekasi tahun ajaran 2024/2025?. Penelitian ini bertujuan untuk mendeskripsikan kualitas butir soal Penilaian Akhir Semester Gasal mata pelajaran Seni Budaya kelas XI SD Shibghah Akhlak Qur'an Bekasi tahun ajaran 2024/2025.

Penelitian ini berupaya mengisi yang ada dalam evaluasi pendidikan, khususnya dalam mata pelajaran seni budaya yang belum banyak dibahas penelitian sebelumnya. Banyak kajian yang fokus pada evaluasi mata pelajaran ilmu eksakta dan linguistik namun jarang yang mendalami kualitas instrumen evaluasi dalam mata pelajaran seni budaya. Evaluasi dalam pembelajaran seni menjadi hal yang sangat penting karena tidak hanya meningkatkan kemampuan peserta didik di bidang seni, tetapi juga dapat meningkatkan sikap melalui nilai-nilai seni tradisional yang mereka pelajari (Sagala et al., 2023). Selain itu, penelitian ini penting dilakukan karena memberikan kontribusi baru dan manfaat yang signifikan, seperti menyediakan analisis terperinci tentang kualitas butir soal pilihan ganda untuk mata pelajaran seni budaya di SD Shibghah Akhlak Qur'an Bekasi. Hasil penelitian ini dapat memberikan rekomendasi konkrit untuk perbaikan dan penyusunan soal butir, serta meningkatkan kompetensi guru dalam menyusun dan menganalisis soal butir. Dengan demikian, penelitian ini membantu menciptakan instrumen evaluasi yang lebih akurat dan representatif terhadap kemampuan peserta didik, yang pada akhirnya dapat meningkatkan kualitas pembelajaran seni budaya.

Metode Penelitian

Penelitian ini merupakan penelitian kuantitatif. Penelitian kuantitatif dapat diartikan sebagai proses penemuan pengetahuan yang menggunakan data dalam bentuk angka untuk menghasilkan informasi mengenai topik yang diteliti (Darmawan, 2013, hal. 37). Metode kuantitatif yang digunakan bersifat deskriptif, yang berarti hasil analisis dapat memberikan gambaran tentang gejala atau fenomena yang diamati, bukan angka. Tujuan utama penelitian deskriptif bukan untuk menguji hipotesis, melainkan untuk menggambarkan dan memahami suatu fenomena atau keadaan secara apa adanya. Peneliti berusaha mengumpulkan data yang objektif dan akurat tentang hal-hal yang menjadi fokus penelitian. Penelitian ini mengikuti langkah-langkah sebagai berikut:

(1) mengumpulkan data,

(2) melakukan pengecekan dan input data,

(3) menganalisis data,

(4) menyusun simpulan dari hasil analisis.

Penelitian ini dilaksanakan di SD Shibghah Akhlak Qur'an Bekasi. Adapun waktu penelitian ini dilaksanakan pada bulan Januari-Maret 2024. Populasi dalam penelitian ini adalah butir soal UAS Gasal Tahun Ajaran 2024/2025 mata pelajaran Seni Budaya kelas XI di SD Shibghah Akhlak Qur'an Bekasi dengan sampel yang dipakai dalam penelitian ini adalah sampel populasi sebanyak 72 peserta didik. Penelitian ini menggunakan teknik dokumentasi dan wawancara untuk mengumpulkan data. Dokumentasi dalam penelitian ini berupa naskah soal, lembar jawaban peserta didik, dan kunci jawaban PAS Gasal Tahun Ajaran 2024/2025 mata pelajaran Seni Budaya kelas XI SD Shibghah Akhlak Qur'an Bekasi. Teknik analisis data kuantitatif dilakukan untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh dengan aplikasi ANATES V4 dan Microsoft excel.

Hasil Penelitian

Hasil

Validitas menunjukkan seberapa jauh kemampuan suatu tes mengukur apa yang hendak diukurnya (Zainuri et al., 2021, p. 156). Oleh karena itu, tes yang sah adalah tes yang benar-benar mampu mengukur apa yang ingin diukur. Dalam penelitian ini, peneliti menganalisis validitas berdasarkan validitas item. Validitas item Merujuk pada kesesuaian butir soal dengan keseluruhan tes. Butir soal dianggap valid jika skor yang diperoleh dari jawaban pada butir tersebut memiliki korelasi positif yang signifikan dengan skor keseluruhan tes (Komarudin & Sarkadi, 2017, p. 134). Untuk mengetahui validitas tiap butir soal, peneliti menggunakan aplikasi program Microsoft excel dengan perhitungan rumus Pearson Product Moment. Pada penelitian ini hasil uji validitas dikonsultasikan ke $rtabel$ pada taraf signifikansi 5% dengan jumlah sampel sebanyak 72 peserta didik atau N = 72. Dengan memperhatikan adanya Degree of Freedom (DF) atau derajat kebebasan pada
$rtabel$ yang memiliki ketentuan DF = N-2, maka nilai $rtabel$ yang digunakan adalah 0,239. Apabila $r11 \geq rtabel$ maka butir soal tersebut dinyatakan valid. namun sebaliknya, jika $r11 < rtabel$ maka butir soal tersebut dinyatakan tidak valid. Adapun distribusi butir soal yang bagiannya berdasarkan indeks validitas dapat dilihat pada Tabel 1 berikut.

Tabel 1. Distribusi Hasil Uji Validitas Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025

Indeks Validitas	Data Soal	Jumlah	Persentase
≥ 0,239	3, 4, 6, 7, 13, 17, 18, 22, 23, 24, 26, 27, 29, 32, 34, 37, 38, dan 39	18	45% (Valid)
< 0,239	1, 2, 5, 8, 9, 10, 11, 12, 14, 15, 16, 19, 20, 21, 25, 28, 30, 31, 33, 35, 36, dan 40	22	55% (Tidak Valid)

Reliabilitas

Reabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen pengukuran, yang mengukur sejauh mana instrumen tersebut dapat diandalkan untuk

menghasilkan hasil yang konsisten (Murtafiah, 2018, p. 62). Dalam penelitian ini, reliabilitas butir soal diuji menggunakan metode konsistensi internal, yang instrumen mana diuji coba kan sekali saja, dan kemudian data yang diperoleh dianalisis menggunakan teknik tertentu untuk menilai sejauh mana butir soal konsistensi tersebut dalam mengukur apa yang diukurnya. Salah satunya menggunakan rumus Spearman Brown dengan patokan apabila r_{11} ≥ 0,70 maka soal yang diujikan memilikI reliabilitas yang tinggi (reliable), tetapi apabila r_{11}< 0,70 maka soal yang diujikan memiliki reliabilitas yang rendah atau tidak reliabilitas (tidak dapat diandalkan) (Jumardin, 2020, p. 37). Reliabilitas menurut Basuki dan Hariyanto (dalam Yuniasari, 2016, p. 69) berdasarkan nilai koefisiennya memenuhi menjadi lima kategori sebagaimana tercantum pada Tabel 2 berikut.

Tabel 2. Koefisien Reliabilitas

Rentang Koefisien Reliabilitas	Kategori
0,000 ≤ r ≥ 0,199	Sangat Rendah
0,200 ≤ r ≥ 0,399	Rendah
0,400 ≤ r ≥ 0,699	Cukup
0,700 ≤ r ≥ 0,899	Tinggi
0,900 ≤ r ≥ 1,000	Sangat Tinggi

Tabel 3. Hasil Uji Reliabilitas Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025

Reliabilitas	Jumlah Butir Soal
0,78	40

Tingkat Kesukaran

Tingkat kesulitan soal adalah ukuran peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu, biasanya dinyatakan dalam bentuk indeks numerik. Menurut Dewi (2021, p. 24) tingkat kesukaran dapat diartikan sebagai angka yang menunjukkan proporsi peserta tes yang mampu menjawab butir soal yang diberikan dengan benar. Tingkat kesukaran pada setiap butir soal dipahami menjadi tiga kelompok, yaitu mudah, sedang, dan sukar (Komarudin & Sarkadi, 2017, p. 171). Berdasarkan pengujian terhadap tingkat kesulitan butir soal, setelah disarankan berdasarkan indeks kesulitan masing-masing kelompok, dapat melihat persentase tingkat kesulitan pada setiap kelompok. Distribusi butir soal berdasarkan indeks kesulitan dapat dilihat pada Tabel 4 berikut.

Tabel 4. Distribusi Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025 Berdasarkan Tingkat Kesukaran

Indeks Kesukaran	Butir Soal	Jumlah	Persentase
0,00 – 0,30 (Sukar)	14, 33, dan 40	3 soal	7,5%
0,31 – 0,70 (Sedang)	2, 5, 11, 16, 19, 21, 22, 24, 25, 28, 31, 36, dan 38	13 soal	32,5%
0,71 – 1,00 (Mudah)	1, 3, 4, 6, 7, 8, 9, 10, 12, 13, 15, 17, 18, 20, 23, 26, 27, 29, 30, 32, 34, 35, 37, dan 39	24 soal	60%

Daya Beda

Daya beda merupakan kemampuan instrumen dalam membedakan antara audiens yang pandai (berkemampuan tinggi) dengan audiens yang tidak pandai (berkemampuan rendah) (Febriana, 2019, p. 128). Daya beda dapat diukur dengan angka indeks daya beda. Untuk mengetahui besar kecilnya nilai daya beda digunakan aplikasi ANATES V4 untuk memberikan hasil dalam bentuk persen kemudian dikelompokkan dalam beberapa kategori. Menurut Zainuri (2021, p. 103) kemampuan daya beda terbagi dalam lima kategori, yaitu sangat tidak memuaskan, tidak memuaskan, memuaskan, dan sangat memuaskan. Distribusi soal berdasarkan indeks daya beda dapat dilihat pada Tabel 5 berikut.

Tabel 5. Distribusi Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025 Berdasarkan Daya beda

Indeks Daya beda	Butir Soal	Jumlah	Persentase
0,00 – 0,19 (Sangat Tidak Memuaskan)	1, 2, 5, 9, 10, 15, 16, 17, 19, 20, 25, 30, 35, dan 40	14 soal	35%
0,20 – 0,29 (Tidak Memuaskan)	3, 7, 8, 12, 13, 24, 26, 28, 29, 33, 34, dan 36	12 soal	30%
0,30 – 0,39 (Memuaskan)	14, 18, 21, 31, 32, dan 37	6 soal	15%
0,40 – 1,00 (Sangat Memuaskan)	4, 6, 11, 22, 23, 27, 38, dan 39	8 soal	20%

Efektivitas Pengecoh

Butir soal yang baik akan memiliki pilihan jawaban yang dipilih secara merata oleh peserta didik. Sebaliknya, butir soal yang buruk akan memiliki pilihan jawaban yang dipilih secara tidak merata karena terlalu mencolok atau berputar. Menurut Amelia (2017, p. 36) efektivitas pengecoh diukur berdasarkan seberapa baik ia dapat mengirimkan peserta tes yang tidak mengetahui jawaban yang benar. Semakin banyak peserta tes yang tertipu oleh pengecoh tersebut, semakin efektif pula fungsinya. Butir

soal pengecoh dapat dikatakan termasuk dalam kategori baik jika pengecoh tersebut dipilih setidaknya 5% dari seluruh peserta didik, maka pengecoh dianggap efektif (Jumardin, 2020, hal. 7). Dengan jumlah 72 peserta didik dan asumsi 5% dari mereka dapat diterima oleh pengecoh, maka setidaknya 4 peserta didik harus memilih pengecoh tersebut. Dalam penelitian ini, pertanyaan terdiri dari 40 soal pilihan ganda, masing-masing dengan empat jawaban potensial (A, B, C, D, dan E). Efektivitas pengecoh pada setiap pertanyaan dihitung menggunakan kriteria analisis dari skala Likert dan diklasifikasikan menjadi lima kategori: tidak baik, kurang baik, cukup baik, baik, dan sangat baik. Adapun distribusi butir soal hasil berdasarkan efektivitas pengecohnya disajikan pada Tabel 6 berikut.

Tabel 6. Distribusi Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025 Berdasarkan Efektivitas Pengecoh

Efektivitas Pengecoh	Butir Soal	Jumlah	Persentase
Sangat Baik	-	0 soal	0%
Baik	2, 7, 19, 22, 26, 33, dan 40	7 soal	17,5%
Cukup Baik	1, 3, 5, 6, 8, 11, 13, 18, 21, 24, 25, 27, 28, 29, 36, dan 39	16 soal	40%
Kurang Baik	4, 9, 12, 14, 15, 16, 17, 23, 30, 31, 34, 35, 37, dan 38	14 soal	35%
Tidak Baik	10, 20, dan 32	3 soal	7,5%

Analisis Butir Soal ditinjau dari Validitas, Tingkat Kesukaran, Daya beda, dan Efektivitas Pengecoh

Skala Likert digunakan untuk menganalisis hasil uji validitas, tingkat kesulitan, daya beda, dan efektivitas pengujian guna menentukan kualitas keseluruhan subjek. Skala Likert adalah instrumen kuantitatif yang digunakan untuk memancarkan pandangan, pemahaman, atau sikap individu atau kelompok terhadap suatu subjek atau topik tertentu (Sugiyono, 2013, p. 93). Kriteria kualitas butir soal mengungkapkan menjadi lima kriteria yang disajikan pada Tabel 7 berikut.

Tabel 7. Kriteria Kualitas Butir Soal

Kriteria Butir Soal	Jumlah kriteria yang terpenuhi (validitas, daya beda, tingkat kesukaran, dan efektivitas pengecoh)	Revisi	Masuk Bank Soal
Sangat Baik	4	0	✔
Baik	3	1	Perbaikan
Sedang	2	2	Perbaikan
Tidak Baik	1	3	✗
Sangat Tidak Baik	0	4	✗

Tabel 8. Distribusi Butir Soal PAS Gasal Mata Pelajaran Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Berdasarkan Analisis Butir Soal Secara Keseluruhan

Kriteria	Butir Soal	Jumlah	Persentase
Sangat Baik	6, 18, 22, dan 27	4 soal	10%
Baik	3, 7, 21, 24, 26, 29, dan 39	7 soal	17,5%
Cukup Baik	1, 2, 4, 5, 8, 11, 12, 13, 14, 17, 19, 23, 25, 28, 31, 32, 36, 37, dan 38	19 soal	47,5%
Tidak Baik	9, 16, 33, 34, dan 40	5 soal	12,5%
Sangat Tidak Baik	10, 15, 20, 30, dan 35	5 soal	12,5%

Tabel 9. Indikasi Penyebab Kegagalan Butir Soal PAS Gasal Seni Budaya Kelas XI di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025

Kriteria	Butir Soal	Jumlah	Persentase
Validitas (Tidak Valid)	1, 2, 5, 8, 9, 10, 11, 12, 14, 15, 16, 19, 20, 21, 25, 28, 30, 31, 33, 35, 36, dan 40	22 soal	29%
Tingkat Kesukaran (Sukar dan Mudah)	4, 10, 14, 15, 20, 23, 30, 32, 33, 34, 35, 37, dan 40	10 soal	13%
Daya beda (Tidak Memuaskan dan Sangat Tidak Memuaskan)	1, 2, 3, 5, 7, 8, 9, 10, 12, 13, 15, 16, 17, 19, 20, 24, 25, 26, 28, 29, 30, 33, 34, 35, 36 dan 40	26 soal	35%
Efektivitas Pengecoh (Kurang Baik dan Tidak Baik)	4, 9, 10, 12, 14, 15, 16, 17, 20, 23, 30, 31, 32, 34, 35, 37, dan 38	17 soal	23%

Pembahasan

Analisis komprehensif terhadap butir soal UAS Seni Budaya kelas XI Gasal di SD Shibghah Akhlak Qur'an Bekasi untuk tahun ajaran 2024/2025 mengungkapkan temuan-temuan yang bervariasi dalam aspek validitas, reliabilitas (keandalan), tingkat kesukaran, daya beda, dan efektivitas pengecoh. Ditinjau dari hasil uji validitas, dapat diketahui sebanyak 45% butir soal berkategori valid dan 55% butir soal berkategori tidak valid. Sebagaimana yang dinyatakan oleh Zainuri (2021, p. 37) bahwa validitas mengacu pada sejauh mana suatu alat pengukuran dapat mengukur atau mencerminkan secara akurat konsep atau variabel yang sedang diukur, dengan benar mencerminkan keadaan yang sebenarnya dari variabel atau konsep yang diukur. Selain itu, Asrul (2014, p. 36) juga menjelaskan bahwa instrumen tes dikatakan valid jika mampu mengukur secara tepat, benar, dan shahih hasil belajar yang telah dicapai

peserta didik setelah menempuh proses pembelajaran dalam rentang waktu tertentu. Dengan mempertimbangkan beberapa pendapat dan hasil uji validitas, 40 soal pilihan ganda dianggap cukup valid. Uji reliabilitas menunjukkan bahwa butir soal mencapai koefisien keandalan sebesar 0,78. Menurut berbagai perspektif dan penilaian reliabilitas, UAS Seni Budaya kelas XI Gasal di SD Shibghah Akhlak Qur'an Bekasi Tahun Ajaran 2024/2025 dianggap sangat dapat diandalkan atau termasuk dalam kategori tinggi.

Setelah melakukan prosedur perhitungan dengan ANATES V4, ditemukan bahwa analisis tingkat kesukaran untuk butir soal UAS gasal Seni Budaya kelas XI di SD Shibghah Akhlak Qur'an Bekasi tidak sesuai dengan proporsi tingkat kesulitan yang ditentukan. Untuk dapat dikatakan baik, tingkat kesukaran dianjurkan sesuai dengan kriteria berikut: 25% harus diklasifikasikan sebagai sulit, 50% sebagai sedang, dan 25% sebagai mudah (Rasmuin & Luddin, 2022, p. 35). Selain itu, dari segi daya beda ditemukan 14 butir soal yang dianggap sangat tidak memuaskan, 12 butir soal tidak memuaskan, 6 butir soal memuaskan, 8 butir soal sangat tidak memuaskan, dan terdapat 4 butir soal dengan indeks daya beda negatif. Ketika indeks meningkat, butir soal menjadi semakin efektif dalam membedakan antara peserta didik yang kompeten dan tidak kompeten. Sidabutar, Putrayasa, dan Martha (2017, p. 8) melalui penelitian mereka menyoroti tiga faktor utama yang berkontribusi pada kualitas daya beda suatu pertanyaan, sebagai berikut: (1) pertanyaan dengan kekuatan pembeda yang baik dirancang sedemikian rupa sehingga tidak terlalu menantang, memungkinkan peserta didik dari kelompok atas maupun bawah untuk menjawab dengan benar; (2) pertanyaan disajikan sesuai dengan materi yang diajarkan oleh guru, memungkinkan peserta didik memahami dan merespons dengan tepat; dan (3) kesiapan dan motivasi peserta didik untuk ujian memiliki dampak pada kemampuan mereka untuk menjawab pertanyaan dengan benar, baik dari kelompok atas maupun bawah.. Dari 40 butir soal yang dianalisis, sekitar 57,5% memiliki pengecoh yang efektif, sementara sekitar 42,5% memiliki pengecoh yang kurang efektif. Butir soal dengan pengecoh yang efektif dapat dimasukkan ke dalam bank soal dan digunakan untuk penilaian berikutnya. Namun, butir soal yang memiliki pengecoh yang kurang efektif perlu diperbaiki sebelum digunakan kembali.

Berdasarkan analisis secara keseluruhan jelas bahwa butir soal PAS gasal Seni Budaya tahun 2024/2025, terdapat 4 butir soal (10%) yang diklasifikasikan sebagai sangat baik, 7 butir soal (17,5%) yang memenuhi kriteria untuk menjadi baik, 19 butir soal (45%) yang dianggap berkualitas baik, 5 butir soal (12,5%) yang dianggap buruk, dan 5 butir soal (12,5%) yang berkualitas sangat buruk. Beberapa faktor berkontribusi pada kegagalan yang mencolok. Pertama, daya beda menunjukkan bahwa butir soal gagal membedakan antara peserta didik yang memahami materi pelajaran dan peserta didik yang tidak. Kedua, validitas menunjukkan ketidaksesuaian antara butir soal yang diberikan frngan apa yang ingin diukur. Ketiga, efektivitas pengecoh dikompromikan ketika butir soal kekurangan pilihan jawaban yang dapat menyesatkan peserta tes secara efektif. Terakhir, tingkat kesukaran menandakan bahwa butir soal belum mampu

untuk menilai kemampuan peserta didik dalam memberikan respons yang tepat dan akurat terhadap pertanyaan.

Kesimpulan

Berdasarkan analisis butir soal PAS gasal mata pelajaran Seni Budaya tahun ajaran 2024/2025, disimpulkan bahwa kualitas soal butir soal tergolong cukup baik dengan jumlah butir soal berkualitas baik lebih banyak dari jumlah soal yang belum berkualitas baik, yaitu terdapat 4 butir soal (10%) yang berkualitas sangat baik, sebanyak 7 butir soal (17,5%) berkualitas baik, 19 butir soal (47,5%) termasuk dalam soal yang memiliki kualitas cukup baik, soal yang berkualitas tidak baik berjumlah 5 butir soal (12,5%), dan terdapat 5 butir soal (12,5%) yang memiliki kualitas sangat tidak baik. Adapun saran yang dapat diajukan dari penelitian yang telah dilakukan, antara lain: (1) Diperlukan perbaikan pada distribusi soal berdasarkan tingkat kesukaran agar lebih proporsional. Soal-soal yang dianggap mudah perlu ditingkatkan kompleksitasnya, sementara soal yang dianggap sukar perlu disederhanakan untuk memudahkan peserta didik untuk memahaminya. (2) Meskipun sebagian besar butir soal memiliki daya beda yang baik, perlu dilakukan perbaikan agar soal lebih efektif dalam membedakan antara kelompok atas dan bawah. Butir soal dengan daya beda negatif atau nol perlu direvisi untuk meningkatkan efektivitasnya. (3) Sebagian besar soal memiliki pengecoh yang kurang efektif. Oleh karena itu, perlu dilakukan revisi pada pengecoh yang tidak berfungsi agar dapat meningkatkan efektivitas soal dalam membingungkan peserta tes. (4) Soal-soal yang berkualitas sangat baik dapat disimpan di bank soal untuk digunakan kembali. Namun, soal-soal yang berkualitas baik atau cukup baik perlu direvisi sesuai dengan indikator kegagalan agar dapat digunakan kembali. Sementara itu, soal-soal yang berkualitas tidak baik atau sangat tidak baik sebaiknya dibuang atau diganti dengan butir soal yang lebih baik.

Referensi

Afifah, N. (2020). Analisis Kualitas Butir Soal Penilaian Akhir Semester Gasal Mata Pelajaran Matematika Kelas IX SMP Negeri 3 Ajibarang Kabupaten Banyumas. In *Core.Ac.Uk*. https://core.ac.uk/download/pdf/295326217.pdf

Amelia, M. A. (2017). Analisis Soal Tes Hasil Belajar High Order Thinking Skills (HOTS) Matematika Materi Pecahan Untuk Kelas 5 Sekolah Dasar. *Jurnal Penelitian*, *20*, 123–131.

Asrul, Ananda, R., & Rosnita. (2014). *Evaluasi Pembelajaran* (2nd ed.). Citapustaka Media. Darmawan, D. (2013). *Metode Penelitian Kuantitatif*. PT Remaja Rosdakarya Bandung.

Dewi, R. L. (2021). *Analisis Tingkat Kesukaran Soal Penilaian Akhir Semester Ganjil pada PelajaranPendidikan JaSDni Olahraga dan Kesehatan Kelas V di SD Negeri Pangengudang Kabupaten Purworejo*.

Febriana, R. (2019). *Evaluasi Pembelajaran* (B. Sari (ed.)). Bumi Aksara.

Jumardin. (2020). *Analisis Butir Soal Ujian Akhir Semester Ganjil Mata Pelajaran IPS Terpadu Kelas VIII di MTs Badrussalam NW Sekarbela Tahun Pelajaran 2019/2020*. http://etheses.uinmataram.ac.id/id/eprint/2179

Komarudin, & Sarkadi. (2017). *Evaluasi Pembelajaran* (2nd ed.). Laboratorium Sosial Politik Press.

Monica, S., & Sudarman, Y. (2019). *Analisis Butir Soal Ujian Tengah Semester Ganjil Seni Budaya Kelas VII Di SMPN 29 Sijunjung*. *7*(3), 1–23.

Muluki, A. (2020). *Analisis Kualitas Butir Tes Semester Ganjil Mata Pelajaran Ilmu Pengetahuan Alam Kelas IV MI Radhiatul Adawiyah*.

Murtafiah, N. H. (2018). Evaluasi pendidikan. In *Jakarta: Rineka Cipta*.

Putra, Z. A. W., Satrianingsih, A. R. O., & Medilianasari, R. (2023). *Arts Education's View on Artistic Research*.

12(4), 1017–1024.

Qodir, A. (2017). *Evaluasi dan Penilaian Pembelajaran* (Ngalimun (ed.)). Penerbit K-Media.

Rahmasari, D., & Ismiyati. (2016). Analisis Butir Soal Mata Pelajaran Pengantar Administrasi Perkantoran.

Economic Education Analysis Journal, 317–330. http://journal.unnes.ac.id/sju/index.php/eeaj

Rasmuin, & Luddin, S. (2022). Validitas Tes Buatan Guru Bidang Studi Matematika Menurut Teori Tes Klasik Tingkat SMP di Kota Baubau. *Jurnal Akademik Pendidikan Matematika*, *8*, 135–141. https://doi.org/10.55340/japm.v8i2.876

Riinawati. (2021). *Pengantar Evaluasi Pendidikan*. Thema Publishing.

Ropii, M., & Fahrurrozi, M. (2017). Evaluasi Hasil Belajar. Evaluasi Hasil Belajar. in *Yogyakarta: Pustaka Pelajar*.

Safira, P. (2016). *Analisis Butir Soal Penilaian Akhir Semester Gasal Mata Pelajaran Bahasa Indonesia SMP Negeri 2 Magelang Tahun Pelajaran 2015/2016* (Issue June).

Sagala, M. D., Ghozali, I., Putra, Z. A. W., Satrianingsih, A. R. O., & Aditya, M. C. P. (2023). Pelatihan Evaluasi Pembelajaran Seni Berbasis Website Bagi Mahasiswa Prodi Pendidikan Seni Pertunjukan Fkip Universitas Tanjungpura. *Jurnal Pengabdian Kepada Masyarakat (JPKM) - Aphelion*, *5*(JUNI), 207–212. http://jurnal.globalhealthsciencegroup.com/index.php/JPM

Sidabutar, G. D. U., Putrayasa, I. B., & Martha, I. N. (2017). Kualitas Butir Soal Penilaian Akhir Semester Ganjil Bahasa Indonesia Kelas XI SD Negeri 2 Singaraja Tahun Pelajaran 2015/2016 dari Segi Taraf Kesukaran, Daya Pembeda, dan Fungsi Pengecoh. *E-Journal PBSI*, *7*(2), 1–11. https://ejournal.undiksha.ac.id/index.php/JJPBS/article/view/7202

Sugiyono. (2013). *Metode Penelitian Kuantitatif, Kualitatif, dan R&D* (19th ed.). CV Alfabeta.

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

Yuniasari, T. (2016). *Analisis butir soal pilihan ganda Penilaian akhir semester genap tahun pelajaran 2014/2015 mata pelajaran Bahasa Indonesia kelas IV SD di Kecamatan Depok*.

Zainuri, A., Aquami, & AnNur, S. (2021). *Evaluasi Pendidikan (Kajian Teoritik) | 1*.