

Analisi Soal Sumatif Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan Kelas XII Fase F Di Sekolah Menengah Atas Negeri 1 Lengayang Tahun 2024

Deri Aldewita Husna¹, Khairuddin², Damrah³, Dessi Novita Sari⁴

Pendidikan Olahraga, Fakultas Ilmu Keolahragaan, Universitas Negeri Padang, Indonesia

derialdewita1201@gmail.com, khairuddin@fik.unp.ac.id, damrahburhan@yahoo.co.id,

dessinovita10@fik.unp.ac.id

Doi JPDO: <https://doi.org/10.24036/JPDO.8.12.2025.7>

Kata Kunci : Analisis Soal, Asesmen Sumatif, Validitas, Reliabilitas, Daya Beda, Tingkat Kesukaran, Efektivitas Pengecoh.

Abstrak : Penelitian ini bertujuan untuk menganalisis kualitas butir soal objektif sumatif mata pelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan (PJOK) kelas XII Fase F di SMA Negeri 1 Lengayang tahun ajaran 2024/2025. Latar belakang penelitian ini adalah pentingnya penyusunan soal yang validitas, reliabilitas, Tingkat Kesukaran, daya pembeda dan efektivitas pengecoh dalam mengukur pencapaian kompetensi siswa. Metode yang digunakan adalah kuantitatif dengan pendekatan *ex post facto*. Data diperoleh dari 50 butir soal pilihan ganda yang dikerjakan oleh 35 siswa. Analisis data dilakukan secara manual menggunakan Microsoft Excel dengan indikator validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Hasil penelitian menunjukkan bahwa hanya 34% soal yang valid reliabilitas berada pada kategori belum tinggi ($r = 0,64$), tingkat kesukaran menunjukkan 52% mudah dan 48% sedang, daya pembeda soal tergolong sangat jelek hingga sangat baik, dan efektivitas pengecoh hanya 50% yang berfungsi baik. Temuan ini mengindikasikan perlunya peningkatan kualitas penyusunan soal, dengan demikian, diperlukan perbaikan dalam penyusunan soal agar evaluasi pembelajaran lebih akurat dan adil.

Keywords : *Item Analysis, Summative Assessment, Validity, Reliability, Discriminating Power, Difficulty Level, Distractor Effectiveness.*

Abstract : *This study aims to analyze the quality of objective summative test items for the subject of Physical Education, Sports, and Health (PJOK) for Grade XII Phase F at SMA Negeri 1 Lengayang in the 2024/2025 academic year. The background of this research is the importance of constructing test items that meet the criteria of validity, reliability, difficulty level, discrimination index, and distractor effectiveness to accurately assess students' competency achievements. This research uses a quantitative approach with an *ex post facto* design. Data were obtained from 50 multiple-choice questions answered by 35 students. Data analysis was conducted manually using Microsoft Excel based on indicators of validity, reliability, difficulty level, discrimination index, and distractor effectiveness. The results show that only 34% of the items are valid, reliability is in the low category ($r = 0.64$), the difficulty level shows that 52% of the items are easy and 48% are moderate, the discrimination index ranges from very poor to very good, and only 50% of the distractors function effectively. These findings indicate the need to improve the quality of test item construction so that learning evaluations can be more accurate and fair.*

PENDAHULUAN

Setiap warga negara Indonesia berhak atas pendidikan, sesuai dengan Undang-Undang Dasar 1945. Negara berkewajiban menjamin terwujudnya hak tersebut. Hal ini mencakup penyediaan sumber daya dan fasilitas pendidikan yang memadai, memastikan semua warga negara memperoleh akses ke sekolah, memastikan mutu pendidikan yang diberikan, dan memberikan rasa aman dan nyaman selama proses pendidikan. (Jasmine, 2024:46).

Peraturan Pemerintah Republik Indonesia Nomor 32 Tahun 2013 tentang Standar Nasional Pendidikan Pasal 22 Ayat (1): Menyatakan bahwa evaluasi pendidikan dilakukan untuk mengukur pencapaian standar kompetensi peserta didik dalam berbagai aspek pembelajaran, termasuk dalam Pendidikan Jasmani, Olahraga, dan Kesehatan. Pasal 22 Ayat (2): Evaluasi hasil pembelajaran dilakukan dengan menggunakan instrumen yang beragam untuk menilai kompetensi yang ditargetkan dalam kurikulum.

Pendidikan adalah akumulasi semua ilmu pengetahuan yang diperoleh selama hidup dalam semua konteks dan lingkungan yang berkontribusi terhadap perkembangan jiwa manusia. Dalam arti luas, mengajar juga merupakan proses pelaksanaan kegiatan pembelajaran, dan pembelajaran dapat dilaksanakan kapan saja dan di mana saja. (Ujud et al., 2023:7912)

Dalam Undang-Undang No. 20/2003 tentang Sistem Pendidikan Nasional Pendidikan dapat kita didefinisikan sebagai sebuah “usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi pada dirinya sehingga memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian,

kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa, dan negara” (Makkawaru, 2019:116) .

Pendidikan jasmani merupakan bagian integral dari Pendidikan holistic dan aktifitas fisik diajarkan di sekolah untuk mempromosikan gaya hidup aktif gaya hidup aktif dalam meningkatkan kebugaran fisik (Khairuddin et al., 2023: 241-253).

Pelaksanaan pembelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan (PJOK) diartikan sebagai implementasi nyata dari rencana pembelajaran yang telah disusun sebelumnya. Pelaksanaan ini mencakup kegiatan mengajar yang dilakukan oleh guru, intraksi dengan peserta didik, penggunaan sarana dan prasarana, serta evaluasi terhadap proses dan hasil belajar (Arie Asnaldi & Arie Prima Richi 2017:92-100)

Mulai tahun 2022–2023, pemerintah menetapkan kurikulum mandiri yang fleksibel untuk pendidikan dasar dan menengah, yang berpusat pada kompetensi, karakter, dan inovasi. Semua lembaga pendidikan di bawah kendali Negara Kesatuan Republik Indonesia (NKRI) diharapkan untuk menggunakan kurikulum mandiri (Mulyasa, 2023:174).

Dalam konteks perubahan kebijakan pendidikan antara tahun 2020 hingga 2024, terdapat penyesuaian dalam pelaksanaan evaluasi akhir di tingkat SMA. Salah satunya adalah penggantian Ujian Sekolah Berstandar Nasional (USBN) dengan asesmen yang diselenggarakan oleh sekolah secara mandiri. Perubahan ini memberikan keleluasaan bagi sekolah untuk merancang dan melaksanakan evaluasi yang lebih sesuai dengan konteks dan kebutuhan siswa, serta mendorong penilaian yang lebih komprehensif dan relevan dengan kompetensi yang diharapkan.

Menurut Parma, R., Hidayani, F., & Asnaldi, A. (2022). Evaluasi proses dilakukan untuk membantu mengimplementasi keputusan sampai sejauh mana rencana telah diterapkan. Evaluasi ini berusaha mencari jawaban atas pertanyaan apakah program sedang dilaksanakan.

Penilaian sumatif adalah penilaian yang menggunakan skor atau angka yang kemudian digunakan untuk menentukan prestasi belajar siswa. Penilaian sumatif dirancang untuk mencatat pencapaian siswa secara keseluruhan dengan cara tertentu. Evaluasi sumatif umumnya dilakukan ketika berakhirnya suatu pokok bahasan, tengah semester, dan akhir semester atau kenaikan kelas (Nurhayati & , Langlang Handayani, 2020:131)

Soal ujian objektif merupakan alat evaluasi yang penting dalam mengukur pencapaian belajar peserta didik. Soal objektif harus memenuhi beberapa syarat agar dapat digunakan secara efektif, yaitu validitas, reliabilitas, dan memiliki daya pembeda serta tingkat kesukaran yang sesuai. (Sudjana 2005:34)

Soal objektif juga harus mempertimbangkan aspek teknis penulisan, seperti penggunaan tata bahasa yang benar, kejelasan petunjuk, serta konsistensi dalam penyajian format soal. Oleh karena itu, penyusunan soal objektif harus dilakukan dengan teliti dan sistematis agar dapat memberikan gambaran hasil belajar yang akurat dan dapat dipercaya. (Zainal Arifin 2012:228).

Dalam konteks pendidikan menengah atas, khususnya pada mata pelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan (PJOK), penyusunan soal asesmen tidak hanya dilakukan oleh guru mata pelajaran di satu sekolah, tetapi juga

melibatkan kolaborasi guru melalui wadah Musyawarah Guru Mata Pelajaran (MGMP). MGMP PJOK tingkat kabupaten/kota berperan sebagai forum profesional yang menyusun soal secara kolektif dan sistematis berdasarkan kurikulum dan capaian pembelajaran yang berlaku.

Oleh karena itu dalam penelitian ini, analisis soal objektif merupakan proses metodis yang mengamati respons siswa untuk menilai kualitas soal ujian pilihan ganda, untuk itu penulis tertarik untuk meneliti " Analisis Soal Objektif Ujian Akhir Sekolah Tahun 2025 Pada Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan Kelas XII Fase F Di SMA Negeri 1 Lengayang".

METODE

Penelitian ini menggunakan pendekatan kuantitatif dengan metode *ex post facto*. Instrumen yang dianalisis adalah 50 butir soal pilihan ganda dari asesmen sumatif akhir yang telah dikerjakan oleh siswa. (Arikunto 2012)

Teknik analisis data dilakukan secara manual dengan bantuan Microsoft Excel, mengacu pada indikator kualitas soal), yaitu validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh.

Penyusunan soal tersebut dilakukan oleh Musyawarah Guru Mata Pelajaran (MGMP) PJOK tingkat kabupaten yang terdiri dari para guru PJOK dari berbagai satuan pendidikan. MGMP memiliki peran penting dalam menjamin kualitas, kesesuaian, dan kesetaraan instrumen asesmen yang digunakan secara kolektif di tingkat wilayah.

Subjek dalam penelitian ini adalah Soal Ujian atau kumpulan soal objektif (pilihan ganda) yang digunakan dalam asesmen sumatif akhir pada mata pelajaran Pendidikan Jasmani, Olahraga, dan Kesehatan (PJOK) kelas XII fase F di SMA

Negeri X. Soal-soal ini terdiri dari 50 butir pilihan ganda yang telah digunakan secara resmi dalam pelaksanaan evaluasi akhir semester genap tahun ajaran 2024/2025.

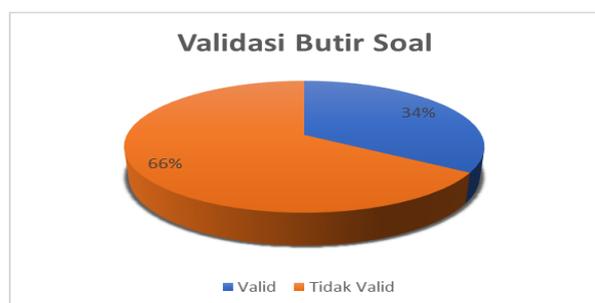
HASIL

Penelitian ini dilakukan untuk mengetahui kualitas Soal Objektif Sumatif Mata Pelajaran Pendidikan Jasmani dan Kesehatan Kelas XII Fase F di Sekolah Menengah Atas Negeri 1 Lembang Tahun Ajaran 2024/2025. Beberapa hal yang ingin diketahui dari analisis ini antara lain Validitas, Reabilitas, Daya pembeda, Tingkat Kesukaran, Epektifitas Pengecoh, yang terdiri dari 50 butir soal dengan 34 siswa dari seluruh siswa kelas XII Fase F1 Sekolah Menengah Atas 1 Lembang

Data diperoleh dengan menggunakan metode survei. Survei dilakukan terhadap lembar jawaban Ujian Sumatif Mata Pelajaran Pendidikan Jasmani Olahraga dan Kesehatan Kelas XII Fase F. Selanjutnya lembar Jawaban tersebut dianalisis dengan manual menggunakan Microsoft Excel untuk menghitung analisis butir soal pilihan ganda ditinjau dengan Validitas, Reabilitas, Daya pembeda, Tingkat Kesukaran, Efektivitas Pengecoh.

Berdasarkan Hasil penelitian disajikan dalam bentuk grafik, tabel atau deskriptif. Analisis dan interpretasi hasil ini diperlukan sebelum dibahas.

1. Validitas



Gambar 2. Diagram Validitas

Hasil dari perhitungan soal akan dikonsultasikan dengan r_{tabel} dengan taraf signifikansi α 0,05(5%). Jumlah subjek penelitian adalah 35 peserta didik, sehingga taraf signifikansi 5% dan $n = 35$ diperoleh r_{tabel} 0,3338. Apabila $r_{hitung} \geq r_{tabel}$ maka butir soal tersebut valid, sebaliknya jika $r_{hitung} < r_{tabel}$ maka soal tersebut tidak valid. Berdasarkan hasil analisis Butir soal pilihan ganda ujian akhir semester ganjil kelas XII mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024, dapat diketahui bahwa butir soal yang valid berjumlah 17 butir 34% dan butir soal yang tidak valid berjumlah 33 butir 66%.

Berdasarkan hasil analisis Butir soal pilihan ganda ujian akhir semester ganjil kelas XII mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024, memiliki kualitas yang kurang baik dari segi validitasnya. Butir soal yang tidak valid perlu diperbaiki dengan menyesuaikan indikator dan meningkatkan penguasaan teknik tentang penyusunan butir soal.

2. Reliabilitas

Pengujian reliabilitas XII mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024 dilakukan dengan bantuan Excel menggunakan Sumber dari Arikunto suharsimi (2013:100-102).

$$r_{11} = \left(\frac{k}{k-1} \right) \left(\frac{SD_t - \sum pq}{SD_t} \right)$$

Dimana :

r_{11} : Reliabilitas Instruments

k : Jumlah Sampel

SD_t : Varians Total

p : Proporsi Subjek item menjawab benar

q : Proporsi Subjek item menjawab salah

$$r_{11} = \left(\frac{35}{34}\right) \left(\frac{23,41 - 8,78}{23,41}\right) = 10294 \times 0,6249 = 0,64$$

- Apabila r_{11} sama dengan atau lebih besar daripada 0,7 berarti tes belajar yang sedang di uji reliabilitasnya dinyatakan telah memiliki reliabilitas yang tinggi.
- Apabila r_{11} lebih kecil daripada 0,7 berarti tes hasil berarti yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (un-reliable).

Berdasarkan hasil perhitungan Reliabilitas menunjukkan jika soal pilihan ganda ujian akhir semester kelas XII Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024 memiliki koefisien Reliabilitas < 0,70. Dengan nilai 0,64 yang berarti belum termasuk ke dalam Reliabilitas tinggi (un-reliable).

3. Tingkat Kesukaran



Gambar 3: Diagram Tingkat Kesukaran

Berdasarkan hasil analisis tingkat kesukaran soal pilihan ganda pada Ujian Akhir Semester Ganjil tahun 2024 untuk mata pelajaran Pendidikan Jasmani Olahraga dan

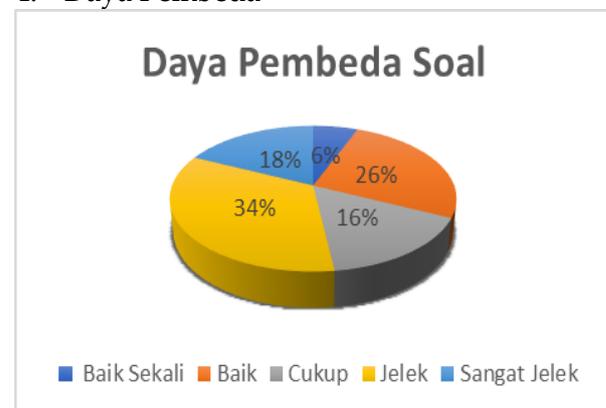
Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024, diketahui bahwa dari total butir soal yang dianalisis, tidak terdapat satupun soal yang tergolong dalam kategori sukar 0%. Sebaliknya, sebanyak 48% soal termasuk dalam kategori sedang, dan 52% termasuk dalam kategori mudah.

Analisis ini dilakukan dengan bantuan program Excel dengan jumlah peserta didik sebanyak 35 orang, menggunakan klasifikasi tingkat kesukaran berdasarkan Arikunto (2009:210), yaitu :

Tingkat Kesukaran	Kriteria
0,00 – 0,30	Sukar
0,31 – 0,70	Sedang
0,71 – 1,00	Mudah

Sumber : Arikunto(2009:210)

4. Daya Pembeda



Gambar 4. Diagram Daya Pembeda Soal

Berdasarkan hasil analisis daya pembeda butir soal pilihan ganda pada Ujian Akhir Semester Ganjil tahun 2024 untuk mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lembang 2024, diketahui bahwa kualitas soal masih bervariasi. Analisis ini dilakukan menggunakan bantuan Microsoft Excel dan mengacu pada klasifikasi menurut Fernandes

(2012), yang membagi daya pembeda ke dalam lima kategori, yaitu :

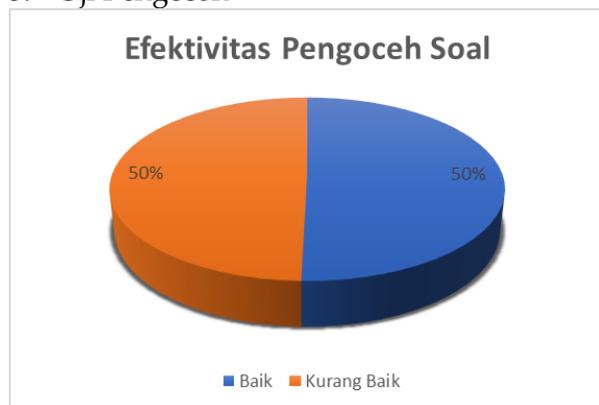
Indek Beda (D)	Kriteria
0,71 – 1,00	Baik Sekali
0,41 – 0,70	Baik
0,21 – 0,40	Cukup
0,00 – 0,20	Jelek
D = Negatif	Sangat jelek

Sumber : Fernandes dalam Kartowagiran(2012:23)

Berdasarkan hasil perhitungan, diketahui bahwa sebanyak 9 soal 18% tergolong dalam kategori sangat jelek, 17 soal 34% termasuk kategori jelek, 8 soal 16% berada dalam kategori cukup, 13 soal 26% termasuk kategori baik, dan hanya 3 soal 6% yang memiliki daya pembeda sangat baik. Dengan demikian, sebagian besar soal masih belum mampu membedakan secara optimal antara peserta didik yang berkemampuan tinggi dan rendah.

Oleh karena itu, diperlukan evaluasi dan perbaikan terhadap soal-soal yang berada dalam kategori sangat jelek dan jelek agar kualitas soal sebagai alat ukur semakin baik dan adil dalam menilai pencapaian peserta didik.

5. Uji Pengecoh



Gambar 5. Diagram Efektivitas Pengecoh Soal

Berdasarkan hasil analisis butir soal pilihan ganda pada Ujian Akhir Semester Ganjil tahun 2024 untuk kelas XII mata pelajaran Pendidikan Jasmani Olahraga dan Kesehatan di Sekolah Menengah Atas Negeri 1 Lengayang 2024, diketahui bahwa dari keseluruhan butir soal yang dianalisis, sebanyak 50% termasuk dalam kategori pengecoh yang baik, sedangkan 50% sisanya termasuk dalam kategori pengecoh yang kurang baik.

Hal ini menunjukkan bahwa separuh dari pengecoh yang digunakan dalam soal belum berfungsi secara optimal dalam mengelabui peserta didik yang kurang memahami materi. Pengecoh yang baik idealnya dipilih oleh sebagian siswa yang tidak memahami materi sebagai alternatif jawaban, sehingga berkontribusi terhadap validitas soal. Oleh karena itu, diperlukan perbaikan atau revisi pada pengecoh yang kurang baik agar daya beda soal meningkat dan kualitas soal secara keseluruhan menjadi lebih baik.

PEMBAHASAN

1. Validasi

Validitas butir soal merupakan aspek krusial yang menunjukkan sejauh mana soal mampu mengukur apa yang harus diukur. Dari hasil analisis, hanya 17 butir soal 34% yang dinyatakan valid, sedangkan 33 butir lainnya 66% tidak valid. Artinya, sebagian soal tidak memiliki keterkaitan yang kuat dengan indikator pembelajaran atau tujuan intruksional yang ingin di capai.

Hal ini dapat disebabkan oleh kurang tepatnya pemilihan materi, perumusan soal yang ambigu, atau kurangnya kesesuaian antara soal dan kompetensi dasar yang hendak diuji. Soal yang tidak valid beresiko menimbulkan interpretasi ganda dan tidak mencerminkan kemampuan sebenarnya dari peserta didik.

Dalam konteks evaluasi pembelajaran, validitas sangat menjadi dasar dalam menentukan ketepatan keputusan yang diambil berdasarkan hasil tes. Ketika soal tidak valid, hasil tes menjadi tidak akurat dan dapat menyesatkan dalam pengambilan keputusan pendidikan, seperti penentuan kelulusan atau evaluasi keberhasilan pembelajaran. Soal yang valid dirancang untuk mengukur kompetensi yang benar-benar sesuai dengan tujuan pembelajaran, sehingga informasi yang diperoleh mencerminkan kemampuan siswa secara nyata (Arikunto, 2012:25).

Validitas juga berperan dalam menjaga mutu instrumen evaluasi, karena hanya instrumen yang sah dan teruji yang dapat menghasilkan data yang konsisten dan dapat dipercaya (Gronlund, 2006:47).

Selain itu, validitas mencegah kesalahan interpretasi terhadap hasil tes, yang seringkali terjadi ketika soal tidak sesuai dengan materi atau tingkat kognitif yang diharapkan (Gay, Mills, & Airasian, 2012:120).

Ketepatan soal dalam mengukur kemampuan juga berdampak pada meningkatnya objektivitas dalam proses penilaian, karena guru tidak lagi menilai berdasarkan dugaan atau persepsi semata, melainkan berdasarkan data yang sah (Mertler & Reinhart, 2017:45).

Data yang valid inilah yang akan mendukung pengambilan keputusan yang tepat terkait intervensi pembelajaran, pengembangan program, hingga penentuan keberhasilan belajar siswa (Brown, 2004:22).

Ketidakvalidan soal dalam evaluasi pembelajaran dapat menimbulkan berbagai kelemahan serius yang berdampak langsung pada kualitas penilaian dan proses pendidikan secara keseluruhan. Soal yang tidak valid cenderung tidak mengukur

kompetensi yang dituju, sehingga hasil yang diperoleh tidak merefleksikan kemampuan siswa sesuai dengan tujuan pembelajaran (Arikunto, 2012:25).

Kondisi ini tidak hanya merugikan siswa secara akademik, tetapi juga berpotensi menurunkan kredibilitas guru dan sistem evaluasi itu sendiri, sebab hasil penilaian yang tidak valid akan diragukan oleh siswa, orang tua, dan pemangku kepentingan lainnya. Soal yang tidak valid juga menghambat perbaikan proses pembelajaran, karena evaluasi yang tidak tepat tidak dapat memberikan umpan balik yang akurat untuk mengidentifikasi kelemahan dalam pengajaran dan pembelajaran (Brown, 2004:22).

2. Reliabilitas

Reliabilitas mengukur sejauh mana alat tes dapat memberikan hasil yang konsisten ketika digunakan berulang kali dalam kondisi yang sama. Dalam penelitian ini, diperoleh koefisien reliabilitas sebesar 0,64, yang berarti masih berada dibawah ambang batas reliabilitas tinggi ($\geq 0,70$). Ini mengindikasikan bahwa soal-soal tersebut belum cukup konsisten dalam mengukur kemampuan peserta didik secara menyeluruh.

Faktor-faktor seperti variasi kualitas butir soal, ketidakseimbangan tingkat kesukaran, atau daya pembeda yang lemah dapat menjadi penyebab rendahnya reliabilitas.

Rendahnya reliabilitas juga dapat memengaruhi kepercayaan terhadap hasil ujian. Jika soal tidak reliabel, maka hasil yang diperoleh siswa belum tentu mencerminkan kemampuan sebenarnya, dan mungkin akan berbeda jika tes yang sama diberikan kembali dalam waktu dekat.

Oleh karena itu, Reliabilitas butir soal menjadi syarat mutlak dalam pelaksanaan

evaluasi pembelajaran yang kredibel. Soal yang reliabel tidak hanya menjamin konsistensi hasil, tetapi juga memperkuat integritas seluruh proses penilaian. Instrumen yang menghasilkan skor serupa dalam kondisi yang setara mencerminkan tingkat kestabilan yang tinggi, menjadikan data evaluasi dapat dipercaya sebagai representasi objektif dari kemampuan peserta didik (Arikunto, 2013:100-102).

Reliabilitas yang rendah juga berdampak pada menurunnya kepercayaan terhadap hasil evaluasi; guru, siswa, maupun orang tua bisa meragukan nilai yang dihasilkan dan mempertanyakan validitas proses penilaian itu sendiri. Instrumen yang tidak reliabel juga menghambat proses diagnosis pembelajaran karena guru tidak dapat mengidentifikasi kekuatan dan kelemahan siswa secara tepat, yang pada akhirnya mengganggu perencanaan pembelajaran selanjutnya (Mertler & Reinhart, 2017).

Oleh karena itu, soal yang tidak memiliki reliabilitas tinggi tidak hanya mengurangi efektivitas evaluasi, tetapi juga dapat menurunkan kredibilitas keputusan pendidikan yang didasarkan pada hasil evaluasi tersebut.

3. Tingkat Kesukaran

Tingkat kesukaran suatu butir soal dipengaruhi oleh kemampuan dan keseriusan peserta didik yang melaksanakan ujian. Oleh karena itu, Hasil tingkat kesukaran menunjukkan sejauh mana suatu soal dapat dijawab dengan benar oleh siswa. Berdasarkan hasil analisis, mayoritas soal termasuk dalam kategori mudah 52% dan sedang 48%, sementara tidak ada satupun soal yang tergolong sukar.

Ketidakseimbangan ini menunjukkan bahwa sebagian besar soal kurang menantang dan berpotensi tidak mampu

menggambarkan variasi kemampuan peserta didik. Soal yang terlalu mudah membuat hampir semua siswa dapat menjawab benar, sehingga mengurangi kemampuan soal untuk menyeleksi dan menilai dengan akurat.

Arikunto (2012:25) menegaskan bahwa analisis tingkat kesukaran membantu menciptakan distribusi soal yang seimbang antara kategori mudah, sedang, dan sukar, sehingga seluruh rentang kemampuan siswa dapat terakomodasi secara optimal.

Tingkat kesukaran yang tepat, terutama pada kategori sedang, juga berkaitan langsung dengan daya pembeda soal yang tinggi, karena dapat membedakan siswa berdasarkan tingkat penguasaan materi secara lebih akurat. Selanjutnya, hasil analisis ini menjadi dasar dalam membangun bank soal yang terstruktur dan berkualitas, yang dapat dimanfaatkan untuk berbagai jenis asesmen secara berkelanjutan (Purwanto, 2013:137).

4. Daya Pembeda

Daya pembeda adalah kemampuan suatu soal dalam membedakan peserta didik yang kemampuan tinggi dan rendah. Dari hasil analisis diketahui bahwa sebagian besar soal berada pada kategori sangat jelek 18% dan jelek 34%, sedangkan hanya 6% soal yang dikategori sangat baik.

Ini menunjukkan bahwa mayoritas soal tidak mampu secara efektif menyeleksi siswa yang benar-benar memahami materi dengan yang tidak. Soal dengan daya pembeda rendah atau bahkan negatif, cenderung dijawab benar oleh siswa dengan kemampuan rendah dan salah oleh siswa dengan kemampuan tinggi, sehingga justru menyesatkan.

Daya pembeda menunjukkan kemampuan soal dalam memisahkan siswa berkemampuan tinggi dari yang

berkemampuan rendah, yang merupakan inti dari proses evaluasi itu sendiri Arikunto (2012:226).

Soal yang memiliki daya pembeda tinggi cenderung dikategorikan sebagai soal yang berkualitas, sedangkan soal yang memiliki daya pembeda negatif harus diperbaiki atau dibuang karena justru membingungkan siswa yang cerdas dan menguntungkan yang kurang memahami (Sudijono, 2015:370).

Selain itu, analisis daya pembeda sangat penting dalam proses revisi dan penyempurnaan instrumen evaluasi, karena dapat membantu guru mengevaluasi efektivitas soal secara menyeluruh. Soal dengan daya pembeda tinggi juga meningkatkan akurasi hasil penilaian, sehingga prestasi siswa dapat tercermin secara objektif dan adil (Purwanto, 2013:180).

Daya pembeda berfungsi sebagai dasar dalam pengembangan bank soal yang berkualitas, karena hanya soal-soal yang mampu membedakan kemampuan siswa secara jelas yang layak disimpan dan digunakan kembali. Dengan demikian, analisis daya pembeda bukan sekadar prosedur teknis, tetapi langkah esensial dalam menjamin mutu evaluasi yang akurat, adil, dan berdaya guna.

5. Efektivitas Pengecoh

Pengecoh adalah alternatif jawaban salah dalam soal pilihan ganda yang berfungsi untuk menjebak peserta didik yang kurang memahami materi. Hasil penelitian menunjukkan bahwa 50% pengecoh yang berkerja secara efektif, sedangkan sisanya tergolong kurang baik.

Pengecoh yang tidak efektif umumnya tidak dipilih sama sekali atau sangat jarang dipilih oleh siswa, sehingga keberadaan tidak memberikan kontribusi terhadap validitas atau daya pembeda soal. Keberadaan

pengecoh yang buruk juga sering kali menyebabkan soal menjadi terlalu mudah.

Pengecoh yang efektif bukan sekadar jawaban salah, melainkan elemen penting yang mampu menilai ketepatan pemahaman siswa. Distraktor yang bekerja dengan baik dapat meningkatkan daya pembeda soal, karena membantu memisahkan siswa yang memahami materi dengan benar dari mereka yang hanya memiliki pemahaman dangkal (Arikunto, 2012:214-215).

Selain itu, pengecoh yang logis dan menyerupai jawaban benar akan mengurangi kemungkinan siswa menjawab secara asal atau menebak dengan benar, sehingga hasil evaluasi menjadi lebih sahih. Pengecoh yang efektif juga memiliki nilai diagnostik yang tinggi karena dapat mengungkap pola kesalahan berpikir atau miskonsepsi siswa terhadap materi tertentu, yang penting sebagai bahan perbaikan pembelajaran (Haladyna, Downing, & Rodriguez, 2002:309).

Jika pengecoh tidak dirancang dengan baik, maka hal itu dapat merusak validitas dan reliabilitas soal, sebab siswa akan lebih mudah mengenali jawaban benar tanpa berpikir kritis, yang pada akhirnya menurunkan kredibilitas alat evaluasi (Nitko & Brookhart, 2011:14).

Oleh karena itu, pengecoh dalam soal pilihan ganda harus dirancang dengan cermat agar instrumen evaluasi tidak hanya sahih dan andal, tetapi juga benar-benar mampu mengukur kemampuan siswa secara komprehensif.

KESIMPULAN

Berdasarkan hasil penelitian dan pembahasan analisis butir soal pilihan ganda yang telah dilakukan meliputi tingkat Validitas, Reabilitas, Daya pembeda, Tingkat Kesukaran, Efektivitas Pengecoh pada soal Objektif Sumatif Mata Pelajaran Pendidikan Jasmani Olahraga Dan Kesehatan Kelas XII

Fase F Di Sekolah Menengah Atas Negeri 1 Lengayang Tahun Ajaran 2024/2025, maka dapat disimpulkan dari 50 Soal Objektif setelah dilakukan analisis butir soal diperoleh hasil sebagai berikut:

1. Validitas soal menunjukkan bahwa 50 butir soal yang dianalisis, hanya 17 butir 34% yang memenuhi syarat validitas, sementara 33 butir 66% dinyatakan tidak valid. Hal ini menunjukkan bahwa sebagian besar soal belum secara optimal mengukur kompetensi yang ditargetkan.
2. Reliabilitas soal termasuk dalam kategori kurang tinggi dengan nilai koefisien reliabilitas $< 0,70$ dengan nilai 0,64. Ini berarti soal yang diujikan belum dapat menghasilkan hasil yang konsisten dan stabil, serta perlu dilakukan revisi untuk meningkatkan keandalan instrumen evaluasi.
3. Tingkat kesukaran soal didominasi oleh soal kategori mudah 52% dan sedang 48%, tanpa adanya soal kategori sukar. Hal ini menunjukkan bahwa sebagian besar soal masih terlalu mudah dan kurang menantang bagi peserta didik.
4. Daya pembeda soal menunjukkan bahwa sebagian besar soal belum mampu membedakan dengan baik antara peserta didik yang berkemampuan tinggi dan rendah. Hanya 3 butir soal 6% yang tergolong sangat baik, sedangkan 26 butir soal 52% masuk dalam kategori sangat jelek dan jelek.
5. Efektivitas pengecoh menunjukkan bahwa 50% pengecoh dalam soal berkerja dengan baik, sementara sisanya kurang efektif. Distraktor yang tidak berfungsi dengan baik perlu diperbaiki agar tidak mengurangi kualitas soal dan daya bedanya.

Secara umum, instrumen evaluasi yang digunakan dalam ujian ini masih perlu

ditingkatkan dari segi validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Upaya perbaikan dan pengembangan soal yang berkelanjutan sangat penting untuk meningkatkan kualitas evaluasi hasil belajar.

DAFTAR PUSTAKA

- Arikunto, S. (2013). *Prosedur Penelitian: Suatu Pendekatan Praktik* (Edisi Revisi). Jakarta: Rineka Cipta, 100-102
- Arikunto, S. (2012). *Dasar-dasar evaluasi pendidikan* (hal. 25). Jakarta: Bumi Aksara.
- Asnaldi, A., Sepriani, R., & Syamsuar, S. (2025). Tinjauan Sarana dan Prasarana Pendidikan Jasmani Olahraga dan Kesehatan (PJOK) Di Sekolah Menengah Pertama (SMP) Se-Kecamatan Simpang Alahan Mati (Simpati) Kabupaten Pasaman. *Jurnal JPDO*, 8(2), 261-271.
- Richi, A. P., & Asnaldi, A. (2017). Studi Pelaksanaan Proses Belajar Mengajar Pendidikan Jasmani Olahraga Kesehatan. *Sport Science*, 17(2), 92-100.
- Brown, H. D. (2004). *Language Assessment: Principles and Classroom Practices* (hal. 22). White Plains, NY: Pearson Education
- Gay, L. R., Mills, G. E., & Airasian, P. W. (2012). *Educational Research: Competencies for Analysis and Applications* (10th ed., hal. 102). Boston, MA: Pearson(hlm120).
- Gronlund, N. E. (2006). *Assessment of Student Achievement* (8th ed., hal. 47). Boston, MA: Allyn & Bacon.
- Jasmine, K. (2024). tanggung jawab negara dalam pemenuhan hak atas pendidikan

- menurut UUD 1945. *Mediation : Journal of Law* Volume 3, Nomor 2, Juni 2024, 3, 44–51.
- Khairuddin, K., Masrun, M., Bakhtiar, S., &, Syahrudin, S, (2023). An analysis of the learning implementation of physical education in junior high schools. *Jurnal Cakrawala Pendidikan*, 42(1), 241-253. doi: <https://doi.org/10.21831/cp.v42i1.54605>
- Makkawaru, M. (2019). Pentingnya Pendidikan Bagi Kehidupan dan Pendidikan Karakter dalam Dunia Pendidikan. *Jurnal Konsepsi*, 8(3), 1–4.
- Magdalena, I., Ridwanita, A., & Aulia, B. (2020). Evaluasi belajar peserta didik. *Jurnal Pendidikan Dan Dakwah*, 2(1), 117–127. <https://ejournal.stitpn.ac.id/index.php/pandawa>
- Mertler, C. A., & Reinhart, R. V. (2017). *Advanced and multivariate statistical methods* (hal. 45). New York, NY: Routledge.
- Marzuki, I. (2024). Implementasi Prinsip-Prinsip Evaluasi Pembelajaran Pada Mata Pelajaran Pendidikan Agama Islam. *Tadarus Tarbawiy*, 6(1), 91–97.
- Nurhayati, H., & Langlang Handayani, N. W. (2020). Jurnal basicedu. *Jurnal Basicedu*, 5(5), 3(2), 524–532. <https://journal.uin.ac.id/article/view/971>
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational Assessment of Students* (6th ed., hal. 37). Boston, MA: Pearson.
- OECD. (2021). *21st-Century Readers: Developing Literacy Skills in a Digital World*. OECD Publishing. <https://doi.org/10.1787/a83d84cb-en>
- Pengertian.id. (2023). *Ujian Sumatif Akhir Jenjang SMA*. Diakses pada 17 Februari 2025, dari <https://pengertian.id/ujian-sumatif-akhir-jenjang-sma/>
- Sukardi, H. M. (2009). *Evaluasi pendidikan: Prinsip dan operasionalnya* (hal. 81). Jakarta: Bumi Aksara.
- Ujud, S., Nur, T. D., Yusuf, Y., Saibi, N., & Ramli, M. R. (2023). Penerapan Model Pembelajaran Discovery Learning Untuk Meningkatkan Hasil Belajar Siswa Sma Negeri 10 Kota Ternate Kelas X Pada Materi Pencemaran Lingkungan. *Jurnal Bioedukasi*, 6(2), 337–347. <https://doi.org/10.33387/bioedu.v6i2.7305>
- Undang-Undang. Undang-undang Republik Indonesia nomor 20 tahun 2003 tentang sistem pendidikan nasional, Pub. L. No. 20 (2003). Indonesia. Retrieved from https://peraturan.bkpm.go.id/jdih/userfiles/batang/UU_20_2003.pdf