

Petunjuk Penulisan Naskah Jurnal Pengabdian Masyarakat dan Riset Pendidikan (Versi Template 2019)

Tri Warsih^{1*}, Amrazi Zakso², Astrini Eka Putri³

^{1,2,3}Pendidikan Sejarah, Universitas Tanjungpura, Jl. Prof. Dr. H.Jl. Profesor Dokter H. Hadari Nawawi, Bansir Laut, Kec. Pontianak Tenggara, Kota Pontianak, Kalimantan Barat 78124, Indonesia

E-mail: Triwarsihwarsih@gmail.com

*Corresponding Author



<https://doi.org/10.31004/jerkin.v4i2.3456>

ARTICLE INFO

Article history

Received: 20 October 2025

Revised: 25 October 2025

Accepted: 24 November 2025

Kata kunci:

analisis butir soal, kualitas instrumen, evaluasi pembelajaran, Sejarah Indonesia, Anates V4

Keywords:

Problem item analysis, instrument quality, learning evaluation, Indonesian history, Anates V4



ABSTRACT

Penelitian ini bertujuan untuk menganalisis kualitas butir soal Ulangan Akhir Semester Ganjil mata pelajaran Sejarah Indonesia kelas XI IPS di SMA Negeri 1 Sungai Kuyit tahun ajaran 2021/2022. Metode penelitian yang digunakan adalah kuantitatif deskriptif dengan menganalisis 35 butir soal pilihan ganda yang diujikan kepada 61 siswa sebagai sampel. Data dikumpulkan melalui studi dokumentasi terhadap lembar jawaban siswa dan wawancara dengan guru. Analisis data dilakukan menggunakan aplikasi Anates V4 untuk mengukur validitas, reliabilitas, tingkat kesukaran, daya beda, dan efektivitas pengecoh. Hasil penelitian menunjukkan bahwa dari segi validitas, hanya 49% butir soal yang valid pada taraf signifikansi 5%. Reliabilitas instrumen sebesar 0,65 termasuk dalam kategori cukup. Sebanyak 51% soal memiliki tingkat kesukaran sedang, 37% soal memiliki daya beda yang tidak baik, dan 69% soal memiliki pengecoh yang tidak berfungsi dengan baik. Temuan ini mengindikasikan bahwa kualitas butir soal secara keseluruhan belum memenuhi standar yang optimal dan perlu perbaikan untuk meningkatkan kualitas evaluasi pembelajaran. Penelitian ini merekomendasikan revisi terhadap butir soal yang tidak memenuhi kriteria kualitas serta pelatihan bagi guru dalam pengembangan instrumen evaluasi yang valid dan reliabel.

This study aims to analyze the quality of the end-of-Semester quiz question on odd subjects of Indonesian history class XI IPS at SMA Negeri 1 Sungai Kuyit for the 2021/2022 academic year. The research method used is descriptive quantitative by analyzing 35 items of multiple choice questions tested to 61 students as a sample. The Data was collected through documentation studies on student answer sheets and interviews with teachers. Data analysis was conducted using Anates V4 application to measure validity, reliability, difficulty, difference power, and effectiveness of the deceiver. The results showed that in terms of validity, only 49% of the questions were valid at a significance level of 5%. Instrument reliability of 0.65 is included in the sufficient category. A total of 51% of the questions have a medium difficulty level, 37% of the questions have a difference power that is not good, and 69% of the questions have an impostor that does not work properly. This finding indicates that the overall quality of the subject matter has not met optimal standards and needs improvement to improve the quality of learning evaluation. This study recommends revisions to the items that do not meet the quality criteria as well as training for teachers in the development of valid and reliable evaluation instruments.



This is an open access article under the [CC-BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.

How to Cite: Tri Warsih, et al (2025). Petunjuk Penulisan Naskah Jurnal Pengabdian Masyarakat dan Riset Pendidikan (Versi Template 2019) 4(2) 10776- 10784 <https://doi.org/10.31004/jerkin.v4i2.3456>

PENDAHULUAN

Menurut Permendikbud Tahun 2016 No. 23 Standar Penilaian Pendidikan, pembelajaran terjadi ketika siswa terlibat satu sama lain, dengan guru, dan dengan materi pembelajaran di lingkungan kelas. Menurut W. S. Winkel in Susanto (2016: "Belajar adalah proses mental yang terjadi selama kontak aktif antara individu dan lingkungannya dan menghasilkan perubahan yang cukup berkelanjutan dan dapat dilacak dalam pengetahuan, pemahaman, kemampuan, dan sikap apresiatif" . Setelah menyelesaikan proses pembelajaran, siswa mengalami setidaknya tiga perubahan: adaptasi kognitif, afektif, dan psikomotorik. Permendikbud Tahun 2016 Nomor 23 tentang Standar Penilaian Pendidikan memuat tiga butir yang tercantum di atas. Perubahan-perubahan inilah yang kemudian menjadi dasar seorang pendidik menilai hasil belajar siswa. Menurut Permendikbud Tahun 2016 Nomor 23 Tentang Standar Penilaian Pendidikan,

Proses pengumpulan dan analisis data untuk mengukur seberapa baik siswa memenuhi tujuan pembelajaran disebut penilaian. Penilaian dalam pembelajaran termasuk ke dalam rangkaian kegiatan yang dinamakan evaluasi. Menurut Sunal dalam Susanto (2016:5) mengatakan bahwa,

" Evaluasi adalah proses menganalisis data untuk menilai seberapa baik suatu program memenuhi kebutuhan siswa. Selanjutnya, ujian atau penilaian tersebut dapat digunakan untuk memberikan umpan balik kepada siswa, serta untuk menilai tingkat kompetensinya. Untuk itu, evaluasi dalam pembelajaran menjadi sangat penting bagi kedua pihak yaitu pendidik dan peserta didik guna untuk memperbaiki dan meningkatkan kualitas dari pembelajaran yang telah dilaksanakan sebelumnya."

Data dari kegiatan pengukuran diperlukan untuk kegiatan evaluasi hasil pembelajaran, klaim Mardapi (2012: 9). Prosedur pengukuran memerlukan penggunaan alat ukur atau alat yang mampu menghasilkan data yang tepat dan dapat dipercaya. Selain itu, Mardapi menyatakan bahwa evaluasi dapat diselesaikan sebagai pekerjaan rumah, kuis, ujian tengah semester, atau ujian akhir. Hal ini sejalan dengan Permendikbud Tahun 2016 Nomor 23 tentang Standar Penilaian Pendidikan, yang menetapkan bahwa guru dapat memanfaatkan tes, observasi, tugas, atau bentuk lain untuk mengevaluasi hasil belajar siswa.

Sejarah merupakan bagian penting dalam kehidupan manusia. Mempelajari sejarah secara tidak langsung menggiring kita untuk bertindak lebih bijaksana daripada peradaban atau kebudayaan sebelumnya. Menurut Susanto (2014 : 7) mengatakan bahwa,

"Sejarah adalah upaya manusia untuk mengilustrasikan semua aktivitasnya dalam urutan kronologis yang terorganisir secara ilmiah, dengan analisis dan interpretasi kritis agar mudah dipahami" Sejarah dapat melukiskan gambaran tentang perilaku dan aktivitas manusia dalam semua variasinya. Pergeseran ini menjadi subjek analisis sejarah.

Dari penjelasan diatas, dapat dikatakan bahwa pembelajaran sejarah memang sangat penting. Terutama dalam mewujudkan pendidikan yang berbasis pendidikan karakter, contohnya saja dalam sejarah diajarkan tentang bagaimana para pejuang daerah maupun nasional dalam mendapatkan dan mempertahankan kemerdekaan indonesia, mereka berani bertaruh segalanya. Dalam hal ini, pendidikan karakter yang bisa diajarkan adalah nasionalisme.

Evaluasi dalam mata pelajaran Sejarah oleh Guru Sejarah juga tidak lepas dari penilaian yang meliputi kognitif, afektif, dan psikomotorik peserta didik. Guru sejarah seringkali memberikan peserta didik tes yang menonjolkan hasil berupa kognitif atau pengetahuan saja. Selain itu dalam pelaksanaannya seringkali guru sejarah mengabaikan hal-hal penting dalam memberikan tes, Membuat pertanyaan atau yang disebut alat evaluasi lainnya adalah salah satu contohnya. Profesor sejarah ceroboh saat membuat soal ujian untuk siswa karena ingin mengurangi waktu. Namun, dalam praktiknya, masalah yang dibuat dengan mempertimbangkan elemen penting seperti ketergantungan, validitas, tingkat kesulitan, dan kekuatan pembeda akan jauh lebih unggul.

Oleh karena itu, penting untuk memeriksa alat evaluasi, termasuk dengan memeriksa pertanyaan yang dibuat. Aspek-aspek seperti tingkat kerumitan pertanyaan, kapasitasnya untuk dibedakan satu sama lain, kemanjuran, validitas, dan reliabilitas si pengecoh semuanya dapat dimasukkan dalam analisis masalah. Faktor-faktor di atas harus diperhitungkan saat menciptakan masalah; tujuannya adalah agar guru dapat mengukur seberapa baik siswanya memahami konten yang telah diberikan kepada mereka. Dengan cara ini, tujuan pembelajaran yang ditentukan guru dapat diimplementasikan secara efektif. Terkait dengan hal ini, peneliti merasa perlu adanya dilakukan pengecekan tes hasil pembelajaran

melalui analisis butir soal guna perbaikan jika didalamnya terdapat kekurangan.

Berdasarkan dari hasil riset dilapangan dalam bentuk wawancara dengan Guru Sejarah di SMAN 1 Sungai Kunyit bahwa untuk mata pelajaran Sejarah Indonesia digunakan teknik penilaian berupa tes tertulis. Tes tertulis digunakan sesuai dengan ketetapan sekolah agar penilaiain dapat dipertanggung jawabkan. Lebih lanjut guru tersebut menyatakan sudah memperhatikan aspek-aspek yang telah disebutkan diatas. Meskipun demikian, guru tidak pernah memeriksa kualitas instrumen soal atau item soal itu sendiri. Hal ini karena tidak ada waktu khusus untuk memeriksa item masalah, dan keterbatasan penggunaan analisis item masalah juga menyulitkan guru untuk memeriksa validitas, reliabilitas, tingkat kesulitan, perbedaan kekuatan, dan kemandirian si pengecoh. Terkait hal tersebut guru memiliki minat untuk ikut bersama peneliti dalam menganalisis butir soal yang telah dibuat dan diujikan kepada siswa. Adapun mekanisme pembuatan soal yang diterapkan oleh guru bisa dikatakan masih secara sederhana. Beliau mengatakan soal yang dibuat disusun dengan menggunakan konsep cara berpikir sejarah yaitu kronologis. Hal ini menunjukkan bahwa untuk membantu siswa memetakan isi pembelajaran secara efektif, pertanyaan disusun dari bagian pertama hingga bagian terakhir dari proses pembelajaran.

Dalam penelitian ini, Peneliti yang tertarik untuk berbicara tentang analisis kualitas item masalah untuk mengidentifikasi masalah yang baik, buruk, dan buruk. Para peneliti melakukan penelitian dengan judul tersebut untuk mendapatkan data yang lebih tepat. " Analisis Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran Sejarah Indonesia Kelas XI IPS SMA Negeri 1 Sungai Kunyit Tahun Ajaran 2021/2022".

METODE

Penelitian ini berfokus pada jenis penelitian deskriptif kuantitatif tertentu. Fokus penelitian kuantitatif semacam ini adalah pada analisis data dengan perhitungan statistik atau numerik. Hal ini sejalan dengan pernyataan Sugiyono (2012:8) bahwa prosedur penelitian kuantitatif digunakan untuk mengeksplorasi suatu populasi atau sampel dengan menggunakan peralatan penelitian untuk pengumpulan data dan pengolahan data kuantitatif/statistik, dengan tujuan untuk memverifikasi hipotesis yang telah ditetapkan. Darmawan (2013: 49), bertujuan untuk mengkarakterisasi objek atau kegiatan yang diminati peneliti yaitu temuan kajian soal ujian akhir untuk topik semester ganjil. Menggunakan Anates V4, kelas XI sejarah Indonesia di IPS SMA Negeri 1 Sungai Kunyit.

Penelitian ini dilakukan di SMA Negeri 1 Sungai Kunyit yang beralamat di Jl. Raya Sungai Kunyit, Sei Limau, Kec. Sungai Kunyit, Kab. Mempawah, Prov. Kalimantan Barat. Adapun waktu penelitian dilaksanakan pada tanggal 28-29 November 2023.

Dalam penelitian ini, populasi merupakan subjek umum yang diteliti. (Syahrudin dan Salim, 2014: 113). Berdasarkan pendapat tersebut, yang menjadi populasi penelitian adalah semua siswa kelas XI IPS di SMA Negeri 1 Sungai Kunyit yang terdiri dari kelas XI IPS 1 berjumlah 31 siswa dan kelas XI IPS 2 berjumlah 30 siswa. Dengan demikian, jumlah populasi adalah 61 siswa.

Suharsimi Arikunto, yang menasihati peneliti dalam pengambilan sampel, mengatakan lebih baik mengambil semua subjek jika jumlahnya kurang dari 100. Karena hanya ada 61 murid di seluruh populasi (kurang dari 100), penelitian ini merupakan studi populasi, artinya setiap anggota populasi dijadikan sampel. Sampel yang digunakan berjumlah 61 siswa yang terbagi dalam 2 kelas.

Pendekatan pembelajaran diawali dengan hubungan langsung dengan guru sejarah Indonesia untuk mendapatkan informasi detail mengenai evaluasi pembelajaran. Selanjutnya, data dokumen dikumpulkan dalam bentuk lembar soal ujian tengah semester, lembar jawaban mahasiswa, kunci jawaban, dan kisi-kisi soal. Data yang diperoleh kemudian dievaluasi dengan alat Anates V4 untuk mengetahui kualitas soal pilihan ganda. Langkah- langkah analisis menggunakan Anates V4 adalah sebagai berikut:

1. Luncurkan aplikasi dan pilih "Run Anates" dari layar utama..
2. Klik "Buat File" untuk membuat file baru, lalu masukkan jumlah mata pelajaran dan pertanyaan, lalu tekan " OK.".
3. Menulis nama peserta tes (atau menggunakan nomor urut) dan skor ideal untuk setiap nomor soal.
4. Simpan file dengan kembali ke menu utama, pilih "Simpan" , lalu beri nama file data.
5. Lengkapi halaman file data dengan memasukkan respon masing-masing peserta terhadap setiap item dan kunci jawabannya..

6. Melakukan analisis butir soal dengan klik „Olah Semua“ untuk analisis otomatis seluruhnya atau memilih perintah tertentu pada menu utama untuk analisis satu per satu.

Data dalam penelitian ini terdiri dari dua jenis:

1. Data kualitatif yang di dapatkan berdasarkan wawancara dengan guru.
2. Data kuantitatif berupa dokumen arsip lembar jawaban siswa (LJS) ulangan tengah semester ganjil mata pelajaran Sejarah Indonesia kelas XI IPS tahun ajaran 2021/2022.

Instrumen yang digunakan untuk mengumpulkan data adalah:

1. Panduan wawancara, yang berfungsi sebagai pedoman wawancara guru sejarah Indonesia.
2. Dokumen arsip, berupa kisi-kisi soal, lembar jawaban siswa, kunci jawaban, dan lembar soal kuis tengah semester.

Strategi pengumpulan data yang digunakan dalam penyelidikan ini adalah:

1. Komunikasi langsung untuk mendapatkan detail tentang proses evaluasi pembelajaran melalui wawancara dengan guru sejarah Indonesia.
2. Dokumentasi, yaitu pengambilan data dari arsip dokumen yang relevan dengan penelitian untuk memperkuat hasil penilaian.

Penelitian ini menggunakan pendekatan analisis data statistik deskriptif. Data dianalisis menggunakan alat anates V4 untuk memeriksa butir soal pilihan ganda. Aspek-aspek yang dianalisis untuk menentukan kualitas butir soal adalah:

1. Validitas, mengukur ketepatan tes dalam menjalankan fungsinya.
2. Reliabilitas, mengukur konsistensi dari tes.
3. Tingkat kesulitan dan persentase siswa yang menjawab setiap pertanyaan dengan benar.
4. Kekuatan diferensial adalah kapasitas masalah untuk membedakan antara siswa dengan tingkat kemampuan tinggi dan rendah.
5. Efektivitas pengecoh, untuk menilai fungsi pengecoh pada soal pilihan ganda.

Data yang telah dianalisis kemudian dideskripsikan untuk menjawab permasalahan dan tujuan penelitian mengenai kualitas instrumen dan butir soal ulangan tersebut.

HASIL DAN PEMBAHASAN

Hasil

Dari data penelitian diperoleh melalui pemeriksaan Soal mata pelajaran ganjil mata pelajaran sejarah Indonesia akhir semester kelas XI IPS di SMAN 1 Sungai Kuyit tahun ajaran 2021/2022 menggunakan aplikasi Anates V4.

1. Kualitas Butir Soal
 - a. Daya Pembeda

Kriteria perbedaan daya menyatakan bahwa perbedaan daya sebesar 0,00-0,20 tidak diinginkan, sedangkan nilai negatif (-) menunjukkan tidak ada perbedaan daya. Perbedaan daya 0,21 hingga 0,30 tidak terlalu baik. Perbedaan daya yang sangat baik berada di antara 0,51 dan 1,00, perbedaan daya yang baik antara 0,31 dan 0,40, dan perbedaan daya yang baik antara 0,41 dan 0,50. Terbukti dari analisis 35 soal pilihan ganda dari semester terakhir Ulangan sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit menggunakan aplikasi Analyze V4 bahwa 4 dari 35 soal masuk dalam kategori "sangat bagus", Meskipun lima dari tiga puluh lima pertanyaan termasuk dalam kategori "baik", sebelas dari tiga puluh lima pertanyaan memiliki kekuatan "cukup baik." Ada dua pertanyaan dalam kategori "kurang baik". Mengenai pertanyaan tersebut, dari 35 item tersebut, 13 ditandai sebagai "tidak bagus" yang memiliki kekuatan yang bervariasi.

Deskripsi hasil berdasarkan beberapa Indeks Daya adalah sebagai berikut:

Table 1. Hasil Analisis Daya Pembeda Butir Soal

No butir	Kelompok atas	Kelompok bawah	beda	Indeks daya beda	Daya pembeda
1	1	10	11	-6.25	FALSE
2	2	10	4	37.5	CUKUP BAIK
3	3	11	3	50	BAIK
4	4	6	1	31.25	CUKUP BAIK

5	5	13	2	68.75	AMAT BAIK
6	6	8	2	37.5	CUKUP BAIK
7	7	4	3	6.25	TIDAK BAIK
8	8	7	1	37.5	CUKUP BAIK
9	9	6	1	31.25	CUKUP BAIK
10	10	4	2	12.5	TIDAK BAIK
11	11	3	1	12.5	TIDAK BAIK
12	12	12	5	43.75	BAIK
13	13	8	4	25	KURANG BAIK
14	14	9	5	25	KURANG BAIK
15	15	15	6	56.25	AMAT BAIK
16	16	8	3	31.25	CUKUP BAIK
17	17	9	6	18.75	TIDAK BAIK
18	18	10	7	18.75	TIDAK BAIK
19	19	14	8	37.5	CUKUP BAIK
20	20	6	0	37.5	CUKUP BAIK
21	21	7	2	31.25	CUKUP BAIK
22	22	2	4	-12.5	FALSE
23	23	4	3	6.25	TIDAK BAIK
24	24	14	9	31.25	CUKUP BAIK
25	25	3	1	12.5	TIDAK BAIK
26	26	13	5	50	BAIK
27	27	14	5	56.25	AMAT BAIK
28	28	8	7	6.25	TIDAK BAIK
29	29	13	5	50	BAIK
30	30	10	9	6.25	TIDAK BAIK
31	31	14	5	56.25	AMATBAIK
32	32	14	8	37.5	CUKUP BAIK
33	33	10	2	50	BAIK
34	34	9	6	18.75	TIDAK BAIK
35	35	1	2	-6.25	FALSE

Adapun persentase pada daya beda soal bentuk pilihan ganda dapat dilihat pada table berikut:

Table 2. Distribusi Soal Pilihan Ganda Menggunakan Indeks Daya Pembeda

Kategori	Jumlah soal	Nomor soal	Persentase (%)
Amat baik	4	5, 12, 27, 31	11%
Baik	5	3, 12, 26, 29,33	14%
Cukup baik	11	2, 4, 6, 8, 9, 16, 19, 20, 21, 25, 32	31%
Kurang baik	2	13, 14	6%
Tidak baik	13	1, 7, 10, 11, 17, 18, 22, 23, 25, 28, 30, 34, 35	37%
Jumlah	35	35	100%

b. Tingkat kesukaran

Tingkat kesulitan mengukur seberapa menantang suatu masalah bagi siswa. Pertanyaan bagus tentang perbandingan seimbang, dengan 30% mengacu pada kategori mudah, 40% untuk kelompok menengah, dan 30% untuk kategori sulit. Pertanyaan item tidak boleh terlalu sederhana atau terlalu rumit. Informasi berikut dihasilkan oleh perangkat lunak V4 Anates berdasarkan temuan evaluasi tingkat kesulitan ujian akhir mata pelajaran semester ganjil mata pelajaran sejarah Indonesia kelas XI IPS SMA NEGERI 1 Sungai Kunyit :

Table 3. Hasil Analisis Tingkat Kesukaran Butir Soal

No	Tingkat Kesukaran	Kriteria
----	-------------------	----------

1	77.05	Mudah
2	39.34	Sedang
3	36.07	Sedang
4	14.75	Sangat Sukar
5	47.54	Sedang
6	26.23	Sukar
7	21.31	Sukar
8	18.03	Sukar
9	16.39	Sukar
10	18.03	Sukar
11	9.84	Sangat Sukar
12	37.7	Sedang
13	29.51	Sukar
14	44.26	Sedang
15	59.02	Sedang
16	31.15	Sedang
17	39.34	Sedang
18	57.38	Sedang
19	63.93	Sedang
20	18.03	Sukar
21	27.87	Sukar
22	19.67	Sukar
23	29.51	Sukar
24	75.41	Mudah
25	9.84	Sangat Sukar
26	59.02	Sedang
27	59.02	Sedang
28	50.82	Sedang
29	72.13	Mudah
30	60.66	Sedang
31	68.85	Sedang
32	63.93	Sedang
33	40.98	Sedang
34	39.34	Sedang
35	9.84	Sangat Sukar

Tabel tersebut menunjukkan bahwa jumlah soal pilihan ganda yang digolongkan sebagai "sangat sulit" adalah 4 dari 35, sedangkan yang dinilai "sulit" adalah 10 dari 35. Sementara soal digolongkan sebagai "sedang", berjumlah menjawab 18 dari 35 soal. Tiga dari 35 pertanyaan dinilai sebagai "mudah."

Table 4. Distribusi Butir Soal Pilihan Ganda Berdasarkan Indeks Tingkat Kesukaran

No	Indeks kesukaran	Butir soal	jumlah	Persentase
1	0,00 - 0,10 (sangat sukar)	4, 11, 25, 35	4	11%
2	0,11 - 0,30 (sukar)	6, 7, 8, 9, 10, 13, 20, 21, 22, 23,	10	29%
3	0,31 - 0,70 (sedang)	2, 3, 5, 12, 14, 15, 16, 17, 18, 19, 26, 27, 28, 30, 31, 32, 33, 34	18	51%
4	0,71 – 0,90 (mudah)	1, 24, 29	3	9%
5	0,91 – 1,00 (sangat mudah)	-	-	-

c. Efektifitas Pengecoh

Kemampuan pengecoh diukur untuk melihat seberapa banyak yang dapat bekerja sebagaimana dimaksud. Salah satu opsi respons dalam pertanyaan pilihan ganda adalah pengecoh. Para siswa akan memilih mereka yang mendapat nilai bagus secara setara. Namun,

karena terlalu mencolok atau menipu, si Pengecoh, item dari pertanyaan yang mengerikan, dipilih secara tidak merata. Ketika setidaknya 5% dari semua siswa memilih pengecoh, itu dianggap sebagai pengecoh yang sangat baik. Dalam hal ini, 3.05. mewakili 5% dari 61 siswa. Oleh karena itu, setidaknya empat siswa harus memilih si pengecoh untuk penelitian ini. 35 item dalam penelitian ini adalah pertanyaan pilihan ganda dengan lima kemungkinan tanggapan untuk setiap pertanyaan., yaitu A,B,C,D, dan E.

Berikut ini adalah pemeriksaan kemampuan issue solver menggunakan software Anates V4.

Table 5. Pola Sebaran Jawaban Butir Soal dan Hasil Analisis Efektivitas Pengecoh

No	Jawaban taste yang memilih jawaban					Pengecoh yang berfungsi	Pengecoh yang tidak berfungsi
	A	B	C	D	E		
1	47*	2	0	10	2	D	B,C,E
2	20	24*	2	9	6	A,D,E	C
3	22*	14	5	18	2	B,C D	E
4	20	6	9*	19	7	A,B,D,E	-
5	9	8	11	4	29*	A,B,C,D	-
6	24	5	16*	12	4	A,B,D,E	-
7	19	15	13*	6	8	A,B,D,E	-
8	35	6	2	7	11*	A,B,D	C
9	17	3	22	10*	9	A,C,E	B
10	21	5	13	11	11*	A,B,C,D	-
11	15	22	6*	15	3	A,B,D	E
12	26	5	7	23*	0	A,B,C	E
13	18*	4	25	6	8	B,C,D,E	-
14	25	27*	1	5	3	A,D	C,E
15	7	36*	5	5	8	A,C,D,E	-
16	12	13	19*	8	9	A,B,D,E	-
17	17	13	24*	6	1	A,B,D	E
18	13	5	2	6	35*	A,B,D	C
19	39*	12	0	4	6	B,D,E	C
20	26	8	11*	8	8	A,B,D,E	-
21	11	4	8	17*	21	A,B,C,D	-
22	17	10	2	12*	20	A,B,E	C
23	7	6	18*	2	28	A,B,E	D
24	2	3	3	7	46*	D	A,B,E
25	6*	25	3	24	3	B,D	C,E
26	3	2	36*	16	4	D,E	A,B
27	6	1	6	36*	12	A,C,E	B
28	6	11	11	31*	2	A,B,C	E
29	44*	4	6	5	2	B,C,D	E
30	5	7	37*	5	7	A,B,D,E	-
31	12	42*	2	4	1	A,D	C,E
32	5	3	8	39*	5	A,C,E	B
33	10	2	12	12	25*	A,C,D	B
34	2	6	9	20	24*	B,C,D	A
35	41	8	6	0	6*	A,B,C	D

Menurut bagan di atas, 11 pertanyaan menampilkan pengecoh yang berhasil berfungsi karena lebih dari 5% peserta tes memilih solusi alternatif, dengan total 61. Pawang yang berfungsi penuh dengan 11 komponen bernomor: 4, 5, 6, 7, 10, 13, 15, 16, 20, 21, dan 30. Sementara gandum bermasalah dengan pengecoh yang tidak berfungsi dengan baik, angka 24 muncul pada Gambar 1, 2, 3, 8, 9, 11, 12, 14, 17, 18, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, dari 35.

Adapun persentase pada efektifitas pengecoh soal bentuk pilihan ganda Mata Pelajaran Sejarah Indonesia pada Ulangan akhir semester ganjil kelas XI IPS SMA Negeri 1 Sungai Kunyit dapat dilihat

pada tabel berikut:

Table 6. Distribusi Soal Pilihan Ganda Berdasarkan Indeks Efektifitas Pengecoh

Kategori	Jumlah Soal	Nomor Soal	Persentase(%)
Berfungsi	11	4, 5, 6, 7, 10, 13, 15, 20, 21, 30	31%
Tidak berfungsi	24	1, 2, 3, 8, 9, 11, 12, 14, 17, 18, 19, 22, 23, 24, 25, 26, 27, 28, 29, 31, 32, 33, 34, 35	69%

2. Kualitas Instrumen

a. Validitas

Uji validitas dibagi menjadi dua langkah: menilai validitas isi dan menguji validitas tes dengan menggunakan soal butir; Namun dalam penelitian ini, peneliti hanya akan mengevaluasi validitasnya secara garis besar. Para peneliti menggunakan perangkat lunak anates V4 untuk memeriksa validitas masalah butir, yang menghasilkan hasil sebagai berikut :

Tabel 7. Hasil Analisis Validitas Butir Soal

No	korelasi	r-tabel		Kriteria
		5%	1%	
1	-0.001	0,291	0,376	Invalid
2	0.33	0,291	0,376	Valid
3	0.305	0,291	0,376	Valid paada signifikansi 5% tapi invalid pada signifikansi 1%
4	0.319	0,291	0,376	Valid pada signifikansi 5% tapi invalid pada signifikansi 1%
5	0.42	0,291	0,376	Valid
6	0.198	0,291	0,376	Invalid
7	0.041	0,291	0,376	Invalid
8	0.424	0,291	0,376	Valid
9	0.252	0,291	0,376	Invalid
10	0.073	0,291	0,376	Invalid
11	0.226	0,291	0,376	Invalid
12	0.395	0,291	0,376	Valid
13	0.228	0,291	0,376	Invalid
14	0.281	0,291	0,376	Invalid
15	0.456	0,291	0,376	Valid
16	0.314	0,291	0,376	Valid
17	0.184	0,291	0,376	Invalid
18	0.23	0,291	0,376	Invalid
19	0.359	0,291	0,376	Valid paada signifikansi 5% tapi invalid pada signifikansi 1%
20	0.424	0,291	0,376	Valid
21	0.231	0,291	0,376	Invalid
22	-0.062	0,291	0,376	Invalid
23	0.049	0,291	0,376	Invalid
24	0.311	0,291	0,376	Valid paada signifikansi 5% tapi invalid pada signifikansi 1%
25	0.075	0,291	0,376	Invalid
26	0.473	0,291	0,376	Valid
27	0.465	0,291	0,376	Valid
28	0.07	0,291	0,376	Invalid
29	0.406	0,291	0,376	Valid
30	0.017	0,291	0,376	Invalid

31	0.392	0,291	0,376	Valid
32	0.333	0,291	0,376	Valid pada signifikansi 5% tapi invalid pada signifikansi 1%
33	0.391	0,291	0,376	Valid
34	0.259	0,291	0,376	Invalid
35	-0.062	0,291	0,376	Invalid

Pada tabel di atas, analisis validitas dibandingkan dengan tabel-R dengan tingkat signifikansi 5%. Jika R dihitung dalam tabel, maka item tersebut sah. Namun, menghitung tabel-R membuat masalah menjadi tidak valid. Berdasarkan perhitungan, 35 pertanyaan pada tingkat signifikansi 5% menghasilkan 17 pertanyaan sah dan 18 pertanyaan tidak sah. Pada tingkat signifikansi 1%, terdapat 12 pertanyaan sah dan 23 pertanyaan tidak sah. Berikut ini adalah pembagian soal pilihan ganda menurut indeks validitasnya:

Table 8. Distribusi Soal Pilihan Ganda Berdasarkan Indeks Validitas

Signifikansi	kategori	Jumlah soal	Nomor soal	persentase
1%	Valid	12	2, 5, 8, 12, 15, 16, 20, 26, 27, 29, 31, 33	34%
	Invalid	23	1, 3, 4, 6, 7, 9, 10, 11, 13, 14, 17, 18, 19, 21, 22, 23, 24, 25, 28, 30, 32, 34, 35	66%
5%	Valid	17	2, 3, 4, 5, 8, 12, 15, 16, 19, 20, 24, 26, 27, 29, 31, 32, 33	49%
	invalid	18	1, 6, 7, 9, 10, 11, 13, 14, 17, 18, 21, 22, 23, 25, 28, 30, 34, 35	51%

b. Reliabilitas

Reliabilitas terkait dengan kepercayaan. Keyakinan tinggi didefinisikan sebagai kemampuan tes untuk menghasilkan temuan yang konsisten. Peneliti memanfaatkan program V4 Anates untuk menguji reliabilitas mata ujian semester akhir mata pelajaran ganjil mata pelajaran sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit, dan diperoleh hasil :

Table 9. Hasil Analisis Reliabilitas Butir Soal

Reliabilitas	Jumlah butir soal
0,65	35

Perhitungan ini menunjukkan bahwa koefisien ketergantungan adalah 0,65. Hal ini menandakan bahwa mata kuliah yang terkait dengan ujian akhir mata kuliah sejarah Indonesia semester ganjil kelas XI IPS SMA Negeri 1 Sungai Kuyit dapat dikatakan “cukup reliabel”.

Pembahasan

1. Kualitas Butir Soal Ulangan Akhir Semester Ganjil Mata Pelajaran Sejarah Indonesia Kelas XI IPS SMA Negeri 1 Sungai Kuyit

Berdasarkan hasil analisis soal ulangan akhir semester ganjil kelas XI IPS SMA Negeri 1 Sungai Kuyit tahun ajaran 2021/2022 yang terdiri dari 35 soal pilihan ganda dan sampel yang diambil dari 61 siswa di masing-masing kelas, hasil analisisnya adalah sebagai berikut:.

a. Daya Pembeda

Sejauh mana item pertanyaan dapat mengidentifikasi siswa yang belum menguasai keterampilan tertentu berdasarkan kriteria tertentu dikenal sebagai kekuatan pembedanya. Tes ini dianggap memiliki sedikit kekuatan pembeda jika menghasilkan temuan yang rendah bila diberikan kepada anak-anak yang berprestasi baik tetapi hasilnya lebih tinggi bila diberikan kepada anak-anak yang berkinerja buruk.

Berdasarkan temuan analisis daya pembeda kuis akhir semester mata pelajaran sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit, masing-masing 23 butir (37%), 2 butir (3%), 11 butir (31%), 5 butir (14%), dan 4 butir (11%), mengklasifikasikan masalah sebagai kurang baik, cukup baik, atau sangat baik.

untuk itu, yang dapat disimpan di bank soal dan memiliki daya pembeda yang tinggi dapat digunakan kembali dalam penilaian selanjutnya. Dimungkinkan untuk meningkatkan item materi dengan kekuatan pembeda yang lebih sedikit untuk menjadikannya materi dengan tingkat

kekuatan pembeda yang baik. Lebih baik membuang pertanyaan dengan kekuatan pembeda yang buruk sehingga tidak akan digunakan dalam evaluasi selanjutnya. Hal ini disebabkan ketidakmampuan masalah untuk membedakan antara siswa dengan kemampuan tinggi dan buruk.

b. Tingkat Kesukaran

Rasio pertanyaan yang berhasil dijawab terhadap peserta ujian dikenal sebagai tingkat kesulitan. Perhitungan tingkat kesulitan menunjukkan tingkat kompleksitas suatu masalah. Suatu masalah dapat dianggap baik jika tingkat kesulitannya proporsional dan seimbang. Pernyataan yang baik adalah pernyataan yang tidak terlalu mendasar atau terlalu sulit.

Program Anates V4 digunakan untuk menilai tingkat kesulitan ujian akhir mata kuliah sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit semester ganjil. Kesimpulan berikut diambil dari analisis tingkat kesulitan pada 35 soal pilihan ganda: soal sulit berjumlah 10 butir (29%), soal mudah berjumlah 3 butir (9%), soal sulit berjumlah 18 butir (51%), dan soal sangat sulit berjumlah 4 butir (11%).

Dari segi apa yang dapat dilakukan berdasarkan hasil analisis tingkat kesulitan soal, butir soal yang termasuk dalam kategori sedang dapat terus diupdate dengan menambahkannya ke bank soal agar dapat dimanfaatkan di evaluasi selanjutnya. Mereka dapat dengan mudah teralihkan.

c. Efektifitas Pengecoh

Siswa yang memberikan jawaban salah akan diberikan kesempatan yang sama untuk mendapatkan nilai bagus. Namun, dalam situasi yang kurang menguntungkan, pengecoh akan dipilih dengan cara yang tidak setara. Analisis penalaran yang efektif terbatas pada pertanyaan pilihan ganda (MCQ) dengan tanggapan alternatif.. Salah satu dari lima kemungkinan jawaban (A, B, C, D, dan E) pada ujian akhir mata kuliah semester ganjil IPS kelas XI IPS SMA Negeri 1 Sungai Kuyit adalah jawaban atau kunci jawaban yang benar. Di sisi lain, respons alternatifnya salah atau curang. Distraktor dianggap efektif jika setidaknya 5% peserta tes memilihnya, menurut Suharsimi Arikunto.

Berdasarkan hasil evaluasi kemampuan ujian ulangan mata pelajaran odd mata pelajaran sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit pada akhir semester, dari 35 mata pelajaran, ditemukan alternatif atau jawaban efektif sebesar 31%, sedangkan mata pelajaran tidak efektif sebesar 69%.

Akibatnya, item pertanyaan dengan pengecoh fungsional dapat ditambahkan ke bank soal dan digunakan dalam pengujian lebih lanjut. Sebelum digunakan kembali, item yang pengisi dayanya tidak berfungsi dapat diperbaiki.

2. Kualitas Instrumen Soal Ulangan Akhir Semester Ganjil Mata Pelajaran Sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit

Kualitas instrumen sangat perlu untuk di perhatikan karena dapat menunjukkan suatu tes yang digunakan sudah baik atau belum. Kualitas instrument dalam penelitian ini meliputi validitas butir soal dan reliabilitas butir soal.

a. Validitas

Ujian akhir mata kuliah semester ganjil mata kuliah sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit didasarkan pada pemeriksaan keabsahan mata kuliah tersebut. Temuan tersebut kemudian dibandingkan dengan tabel R pada level signifikan 1% dan 5%. Tabel tersebut sah jika R sama dengan R. Namun, item dalam terbitan dianggap tidak valid jika R dihitung. Ada hingga 61 contoh peserta didik dalam evaluasi ini. sehingga nilai tabel R pada tingkat signifikansi 1% adalah 0,376. Namun, ambang signifikansi tabel R 5% adalah 0,291. Dua belas soal (34%) dianggap sah dan dua puluh tiga soal lainnya (66%) dianggap tidak sah, berdasarkan temuan studi validitas bentuk pilihan ganda mata pelajaran sejarah Indonesia Kelas XI IPS SMA Negeri 1 Sungai Kuyit pada tingkat signifikansi sebesar 1%. Pada tingkat signifikansi 5%, 17 pertanyaan (49%) dianggap sah sedangkan 18 pertanyaan lainnya (51%) dianggap tidak sah.

Berdasarkan data tersebut, soal ujian semester akhir mata pelajaran ganjil mata pelajaran sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit memiliki kualitas yang buruk karena terdapat 66% dan 51% item yang tidak valid pada tingkat signifikansi masing-masing sebesar 1% dan 5%. Ini menunjukkan bahwa ada lebih banyak pertanyaan yang tidak valid.

Akibatnya, bahan tidak mengukur apa yang seharusnya. Berdasarkan temuan studi validitas item ini, tindakan berikut dapat diambil:

- 1) Item yang divalidasi disimpan dan dapat digunakan kembali dalam evaluasi berikutnya.
 - 2) Setiap barang yang tidak valid harus diperbaiki atau dibuang.
- Ada beberapa alasan mengapa item tersebut mungkin tidak valid.

Hal ini sesuai dengan usulan Gronlund bahwa tiga elemen instrumen evaluasi, administrasi dan penilaian, dan tanggapan peserta didik mempengaruhi validitas temuan pengujian.

d. Reliabilitas

Tingkat konsistensi suatu instrumen dikenal sebagai keandalannya. Pertanyaan apakah suatu pengujian bersifat komprehensif dan dapat diandalkan sesuai dengan standar yang telah ditentukan adalah fokus dari ketergantungan pengujian. Anates V4 digunakan untuk menguji reliabilitas soal pilihan ganda kelas sejarah bahasa Indonesia Kelas XI IPS SMA Negeri 1 Sungai Kuyit. Kriteria nilai reliabilitas kemudian digunakan untuk membandingkan hasil. Reliabilitas sangat tinggi ditunjukkan dengan nilai reliabilitas antara 0,800 dan 1,000, reliabilitas tinggi dengan nilai reliabilitas antara 0,600 dan 0,800, reliabilitas cukup dengan nilai reliabilitas antara 0,400 dan 0,600, reliabilitas rendah dengan nilai reliabilitas antara 0,200 dan 0,400, dan reliabilitas sangat rendah dengan nilai reliabilitas 0,200.

Berdasarkan temuan kajian reliabilitas soal ujian akhir pilihan ganda mata kuliah semester ganjil mata kuliah sejarah Indonesia kelas XI IPS SMA Negeri 1 Sungai Kuyit, reliabilitas soal masuk dalam kategori adequate, dengan skor 0,65. Ini menunjukkan bahwa keandalan item harus ditingkatkan sekali lagi, karena ini adalah salah satu persyaratan penting untuk pengujian kualitas.

Sejumlah variabel berkontribusi pada masalah yang tidak dapat diandalkan, termasuk (1) masalah ujian itu sendiri, yaitu durasi ujian dan kualitas pertanyaan. (2) item yang berkaitan dengan pengujian itu sendiri (3) item yang berkaitan dengan pelaksanaan pengujian. Ketergantungan mata

pelajaran ganjil mata pelajaran sejarah Indonesia aneh kelas XI IPS SMA Negeri 1 Sungai Kuyit, pada akhir semester, dapat ditingkatkan dengan memasukkan topik-topik dengan pertanyaan-pertanyaan yang tulus dan meningkatkan tingkat kerumitan soal. Hal ini memvalidasi anggapan Suharsimi Arikunto bahwa ketergantungan yang tinggi dan rendah merupakan indikator validitas yang lemah. Selain itu, hambatan berdampak pada ketergantungan, karena banyak keadaan sederhana dan rumit memiliki keandalan yang rendah.

SIMPULAN

Untuk kelas sejarah Indonesia XI IPS SMA Negeri 1 Sungai Kuyit tahun ajaran 2021/2022, dapat ditarik kesimpulan dari penelitian dan pembahasan analisis soal butir: kekuatan pembeda, tingkat kesulitan, efektivitas pengecoh, validitas, dan reliabilitas pada soal ujian semester ganjil akhir semester:

Dari 35 soal pilihan ganda, empat bagus, lima bagus, 14 cukup bagus, dan dua tidak. Dan 13 pertanyaan dalam kategori itu tidak terlalu bagus. Tingginya sekitar tiga soal pada kategori mudah, delapan belas soal pada kategori sedang, sepuluh soal pada kategori sulit, dan empat soal pada kategori sangat sulit. Terakhir, berdasarkan efisiensi pengecoh, penyelidikan ini menemukan bahwa sebanyak 11 pertanyaan berfungsi dan 24 tidak berfungsi.

Dengan tingkat signifikansi 5%, 17 item soal pilihan ganda tabel R dinyatakan valid, sedangkan 18 dianggap tidak valid. Tabel R mengakui 1%. Dua belas kueri dianggap asli, sedangkan 23 dianggap tidak valid. Dalam hal ketergantungan, koefisien 0,65 dianggap cukup dapat diandalkan.

REFERENSI

- Arifin, Zainal. 2016. *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya. Arikunto, Suharsimi (2018) *Dasar-Dasar Evaluasi Pendidikan*) Jakarta: Bumi Aksara
- Azwar, Saifuddin (2016) *Tes Prestasi Fungsi Pengembangan Pengukuran Prestasi Belajar*. Yogyakarta: PUSTAKA PELAJAR.
- Daryanto (2014) *Evaluasi Pendidikan* (Cetakan ke-8) Jakarta RINEKA CIPTA Dirman & Juarsih, Cicih. (2014) *Penilaian dan Evaluasi*. Jakarta Rincka Cipta.
- Kunandar (2015) *Penilaian Autentik: Penilaian Hasil Belajar Peserta Didik* PERSADA Mardapi, Djemari. (2012) *Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Yogyakarta: Nuha Medika

- Nugrahani, Farida (2014). *Metode Penelitian Kualitatif dalam Penelitian Pendidikan Bahasa*. Surakarta: Deepublish.
- Nurkencana, Wayan dan Sunartana. 1992. *Evaluasi Hasil Belajar*. Surabaya: Usaha Nasional. Permendikbud No. 23 Tahun 2016 Tentang Standar Penilaian Pendidikan.
- Purwanto. (2016). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar. Ratnawulan & Rusdiana. *Evaluasi Pembelajaran*. Bandung: Pustaka Setia. 2015.
- Sugiyono. (2018) *Metode Penelitian Kuantitatif, Kualitatif, Dan R&D*. Bandung ALFABETA.
- Susanto, Ahmad. (2016) *Teori Belajar dan Pembelajaran Di Sekolah Dasar Edisi Pertama* (Cetakan ke-4). Jakarta: PRENADAMEDIA GROUP
- Susanto, Heri. (2014). *Seputar Pembelajaran Sejarah*. Yogyakarta: Aswaja Pressindo. Syahrums & Salim. (2016). *Metodologi Penelitian Kuantitatif*. Bandung Citapustaka Media.
- Tim (2019) *Pedoman Penulisan Karya Ilmiah Fakultas Keguruan dan Ilmu Pendidikan Universitas Tanjungpura*. Edisi Kedelapan (Revisi). Pontianak: Edukasi Press FKIP Untan
- Widoyoko, Eko Putra. (2016). *Penilaian Hasil Pembelajaran Di Sekolah* (Cetakan Kedua). Yogyakarta: Pustaka Pelajar