

Kualitas Butir Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah Terhadap Hasil Belajar Kognitif Siswa di Fase E SMK Negeri 1 Painan

Yuna Febriya Basri¹, Elfa Michellia Karima²

^{1,2}Program Studi Pendidikan Sejarah, Universitas Negeri Padang
e-mail: yunafebriyabasri09@gmail.com

Abstrak

Tes merupakan salah satu instrumen dalam melakukan evaluasi keberhasilan belajar. Pentingnya tes dalam penilaian pembelajaran maka dibutuhkan instrumen butir soal yang berkualitas agar menjamin kualitas tes yang disajikan kepada peserta didik. Penelitian ini bertujuan untuk mendeskripsikan kualitas butir soal penilaian tengah semester ganjil mata pelajaran sejarah fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023. Metode yang digunakan pada penelitian ini adalah metode deskriptif kuantitatif. Data dianalisis dengan menggunakan program Anates ver.4.0.9. Berdasarkan hasil penelitian analisis 30 butir soal pilihan ganda, ditinjau dari segi validitas empiris terdapat 96,7% butir soal yang termasuk kategori valid. Ditinjau dari segi reliabilitas, soal termasuk memiliki kategori reliabilitas tinggi yaitu 0,74. Ditinjau dari segi daya pembeda, terdapat 93,3% butir soal yang memiliki daya pembeda baik. Ditinjau dari segi tingkat kesukaran, terdapat 63,3% butir soal yang memiliki tingkat kesukaran baik. Ditinjau dari segi efektivitas pengecoh, terdapat 96,7% butir soal yang termasuk dalam kategori opsi pengecoh berkualitas.

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

Abstract

Tests are one of the instruments in evaluating learning success. The importance of tests in learning assessment means that quality instrument instruments are needed to ensure the quality of the tests presented to students. This research aims to describe the quality of the odd mid-semester assessment items in the history subject phase E at SMK Negeri 1 Painan for the 2022/2023 academic year. The method used in this research is a quantitative descriptive method. Data were analyzed using the Anates program ver.4.0.9. Based on the results of research analysis of 30 multiple choice questions, in terms of empirical validity, there were 96.7% of the questions included in the valid category. In terms of reliability, the questions have a high reliability category, namely 0.74. In terms of discriminating power, there are 93.3% of questions that have good discriminating power. In terms of level of difficulty, there are 63.3% of questions

that have a good level of difficulty. Judging from the effectiveness of distractors, there are 96.7% of questions that fall into the category of quality distractor options.

Keywords : *Analysis of Questions Items, Anates, History*

PENDAHULUAN

Pendidikan adalah salah satu sektor yang sangat penting dalam pembangunan di setiap negara. Menurut Syafril dan Zelhendri Zen (2017: 42) pendidikan memegang peranan penting bagi penyiapan peserta didik untuk menghadapi lingkungan hidup yang mengalami perubahan yang pesat, sehingga mendorong untuk terlibat dalam proses mengubah kehidupan kearah yang lebih baik. Evaluasi pembelajaran adalah suatu proses merencanakan, memperoleh, dan menyediakan informasi yang sangat diperlukan untuk membuat alternatif-alternatif keputusan sesuai dengan kemajuan hasil belajar yang ditunjukkan oleh peserta didik. Proses evaluasi harus diarahkan ke tujuan untuk mendapatkan berbagai jawaban tentang bagaimana memperbaiki pembelajaran. Dengan demikian evaluasi merupakan proses yang berhubungan dengan pengumpulan informasi yang memungkinkan guru untuk menentukan tingkat kemajuan pembelajaran dan menentukan pembelajaran ke depan agar lebih baik.

Menurut Nengsi (2018: 119) keputusan pendidikan umumnya diperoleh dari kegiatan pengukuran dan penilaian pendidikan. Pengukuran adalah kegiatan yang dilakukan untuk memberi angka terhadap obyek pengukuran seperti prestasi atau hasil belajar peserta didik. Kegiatan pengukuran ini biasanya melalui tes, secara spesifik diperoleh dari tes hasil belajar. Sedangkan penilaian adalah penerapan berbagai cara dan penggunaan berbagai alat penilaian untuk memperoleh informasi tentang sejauh mana hasil belajar peserta didik atau ketercapaian pada tujuan pembelajaran. Penilaian tidak hanya untuk mengukur capaian peserta didik terhadap kompetensi yang telah ditetapkan namun juga memungkinkan pendidik menggunakan informasi kondisi peserta didik untuk memperbaiki pembelajaran. Menurut Nugraha & Kuswono (2019: 54) pelaksanaan pembelajaran di sekolah tidak hanya berupa proses belajar yang dilakukan guru dan peserta didik tetapi terdapat proses penilaian terhadap hasil belajar yang telah dilakukan. Penilaian ini menjawab pertanyaan tentang sebaik apa hasil atau prestasi belajar peserta didik sehingga dapat meningkatkan efektivitas dalam pengajaran (Yusrizal, 2015: 14).

Penilaian hasil belajar peserta didik diukur dengan menggunakan tes, jika dilihat dari aspek standarisasi ada dua macam tes yaitu tes baku dan tes non baku (tes buatan guru). Tes baku adalah tes yang sudah diuji di lapangan dengan maksud mendapatkan data tentang keterandalan (reliability) dan kesahihan (validity) pengukuran serta standar normatif yang dipakai untuk menaksir skor tes. Sedangkan tes non baku atau tes buatan guru ialah tes hasil belajar yang disusun oleh guru itu sendiri untuk mengukur dan menilai hasil belajar peserta didik, baik pada setiap penyajian satu satuan pembelajaran, maupun pada ujian-ujian formatif serta sumatif (Toijo & Hulukati, 2013: 134). Butir soal tes pilihan ganda yang terstandar harus

memenuhi kriteria karakteristik butir soal yang baik. Parameter karakteristik butir soal pilihan ganda yang baik yaitu valid, reliabel, tingkat kesukaran, daya beda dan keberfungsian pengecoh ((Wibawa, 2019: 89).

Tes memiliki peranan penting untuk mengukur keberhasilan peserta didik dalam pembelajaran sejarah. Mengingat pentingnya tes dalam penilaian pembelajaran maka dibutuhkan instrumen butir soal yang berkualitas sehingga dapat menjamin kualitas tes yang disajikan kepada peserta didik. Upaya untuk mengetahui apakah soal yang dibuat oleh guru sudah tergolong layak dan baik, serta memberikan hasil yang maksimal dengan mengukur dan meningkatkan tingkat pemahaman peserta didik, maka dapat dilakukan analisis pada setiap butir soal pilihan ganda (*Multiple Choices Questions*). Guru harus menganalisis soal pada setiap butirnya guna mengetahui kualitas soal tersebut. Hal ini bertujuan agar dapat meningkatkan kualitas tes melalui revisi atau membuang soal yang tidak efektif, serta untuk mengetahui informasi peserta didik telah menguasai materi yang diajarkan oleh guru. Analisis butir soal merupakan suatu proses untuk mengetahui kualitas pada setiap butir soal. Analisis butir soal bertujuan untuk mengidentifikasi soal yang baik dan soal yang buruk untuk mengukur hasil belajar kognitif peserta didik. Oleh karena itu untuk mengetahui butir soal yang berkualitas dan baik sebagai alat ukur hendaklah dilakukan suatu analisis butir soal sehingga dapat diketahui butir soal mana yang harus direvisi atau bahkan dibuang.

Sebuah tes disebut valid apabila tes itu dapat tepat mengukur apa yang hendak diukur. Jika data yang dihasilkan dari sebuah instrumen valid maka dapat dikatakan bahwa instrumen tersebut valid, karena dapat memberikan gambaran tentang data secara benar sesuai dengan kenyataan atau keadaan sesungguhnya. Butir soal yang memiliki kualitas baik dapat digunakan sebagai alat ukur yang dapat dipercaya, sedangkan butir soal yang memiliki kualitas buruk dapat mempengaruhi hasil penilaian. Ketepatan penilaian akan mempengaruhi upaya peningkatan kualitas pendidikan, oleh karena itu dapat dikatakan bahwa hasil penilaian merupakan salah satu hal yang mempengaruhi kualitas pendidikan. Untuk mengetahui intensitas sebuah soal dalam hal kesukaran dibutuhkan sebuah daya pembeda, yaitu kemampuan antara butir soal dapat membedakan antara peserta didik berkemampuan tinggi atau yang memiliki kemampuan menguasai materi diujikan dan peserta didik yang berkemampuan rendah atau yang belum menguasai materi yang diujikan. Butir soal yang terlalu mudah tidak merangsang peserta didik untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan peserta didik menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya (Arikunto, 2013: 222). Sedangkan pengecoh (*distractor*) yang baik, pengecoh yang dapat dihindari oleh peserta didik yang berkemampuan tinggi dan terpilih oleh peserta didik yang berkemampuan rendah. Bertujuan untuk mengecohkan peserta didik yang kurang mampu atau tidak tahu untuk dapat dibedakan dengan yang mampu.

Berdasarkan observasi di lapangan ditemukan bahwa soal yang dibuat oleh guru mata pelajaran, pada kualitas soal belum diketahui karena belum dilakukan uji coba dan tidak dilakukan analisis setelahnya sehingga tidak ada *feedback* untuk

pelaksanaan penilaian selanjutnya. Guru hanya melakukan penskoran terhadap soal penilaian tengah semester pada tingkat ketuntasan peserta didik untuk melakukan remedial. Oleh karena itu penting dilakukan uji coba pada soal penilaian tengah semester pembelajaran sejarah untuk mengetahui apakah soal tersebut berkualitas dan memenuhi standar.

Peneliti mengangkat masalah ini untuk mengetahui kualitas butir soal pilihan ganda melalui analisis butir soal menggunakan program Anates ver.4.0.9 untuk mengetahui seberapa baik kualitas butir soal penilaian tengah semester ganjil mata pelajaran sejarah kelas X SMK Negeri 1 Painan tahun pelajaran 2022/2023 untuk mengukur hasil belajar kognitif peserta didik ditinjau dari tingkat validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh. Manfaat dari penelitian ini yaitu sebagai harapan bahwa instrumen penilaian tes objektif bentuk pilihan ganda yang dianalisis diketahui validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan kualitas pengecoh butir soal. Soal yang memiliki kualitas baik untuk mengukur hasil belajar kognitif peserta didik dapat disimpan sebagai bank soal sejarah.

METODE

Jenis penelitian ini adalah penelitian deskriptif kuantitatif. Penelitian ini menggunakan pendekatan kuantitatif, yaitu pendekatan yang memandang tingkah laku manusia dapat diramal dengan realitas sosial, objektif, dan dapat diukur. Oleh karena itu, penggunaan penelitian kuantitatif dengan instrumen yang valid dan reliabel serta analisis statistik yang sesuai dan tepat menyebabkan hasil penelitian yang dicapai tidak menyimpang dari kondisi yang sesungguhnya (Nurlan, 2019: 13-14). Metode yang digunakan dalam penelitian ini adalah metode deskriptif yaitu, metode untuk menggambarkan suatu hasil penelitian. Bertujuan untuk memberikan deskripsi, penjelasan, juga validasi mengenai fenomena yang tengah diteliti (Ramdhan, 2021: 7-8).

Sesuai dengan masalah yang diteliti, maka populasi dari penelitian ini yaitu peserta didik kelas X atau fase E di SMK Negeri 1 Painan yang belajar sejarah. Alasan peneliti memilih kelas X karena kelas tersebut adalah kelas yang sudah menerapkan Kurikulum Merdeka pada pembelajaran sejarah di SMK Negeri 1 Painan tahun pelajaran 2022/2023 sehingga soal yang diberikan harus soal yang berkualitas. Kelas X di SMK Negeri 1 Painan tahun pelajaran 2022/2023 terdiri dari 12 kelas, yaitu kelas X AKL 1 berjumlah 29 siswa, X AKL 2 berjumlah 29 siswa, X MPLB 1 berjumlah 34 siswa, X MPLB 2 berjumlah 33 siswa, X BDP 1 berjumlah 32 siswa, X BDB 2 berjumlah 32 siswa, X TJKT 1 berjumlah 30 siswa, X TJKT 2 berjumlah 33 siswa, X PHT 1 berjumlah 35 siswa, X PHT 2 berjumlah 33 siswa, X KLN berjumlah 33 siswa, X TE berjumlah 30 siswa. Sehingga total populasi pada penelitian ini adalah 383 siswa.

Dalam penelitian ini teknik pengambilan sampel yang digunakan adalah menggunakan teknik pengambilan sampel *purposive sampling*, artinya semua penentuan sampel dimana sampel diambil berdasarkan kriteria tertentu. Menurut Mamik (2015: 50) *purposive sampling* adalah unit sampel yang disesuaikan dengan kriteria-kriteria tertentu yang diterapkan berdasarkan tujuan penelitian. Dalam

penelitian ini supaya peneliti dapat dengan mudah melakukan penelitian maka sampel yang digunakan adalah peserta didik yang mengerjakan soal pilihan ganda (*Multiple Choice Questions*) pada ujian tengah semester ganjil yang berjumlah 30 butir yang terdiri dari 5 *option* pada penilaian tengah semester mata pelajaran sejarah tahun pelajaran 2022/2023 di kelas X AKL 1, X KLN, X MPLB 1, X MPLB 2, X PHT 1.

Tabel 1. Jumlah Peserta Didik Fase E SMK Negeri 1 Painan yang Mengikuti Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah Tahun Pelajaran 2022/2023

No.	Kelas	Jumlah
1.	X AKL 1	27
2.	X KLN	28
3.	X MPLB 1	31
4.	X MPLB 2	33
5.	X PHT 1	26
Jumlah		145 orang

Sumber: Data Observasi Peneliti

Variabel dalam penelitian ini adalah kualitas soal ujian tengah semester ganjil mata pelajaran sejarah kelas X SMK Negeri 1 Painan tahun pelajaran 2022/2023 meliputi validitas, reliabilitas, daya beda, tingkat kesukaran, dan efektivitas pengecoh.

Jenis data dalam penelitian ini adalah data primer, yaitu data yang didapatkan langsung dari set soal. Sumber data dalam penelitian ini yaitu soal ujian tengah semester ganjil mata pelajaran sejarah kelas X tahun pelajaran 2022/2023 bentuk pilihan ganda berjumlah 30 soal, kunci jawaban, dan lembaran jawaban ujian peserta didik yang digunakan sebagai alat untuk menganalisis secara kuantitatif kualitas butir soal. Teknik pengumpulan data pada penelitian ini adalah menggunakan dokumen. Data Dokumen yang dikumpulkan untuk menganalisis butir soal sejarah pada penelitian ini adalah soal, kunci jawaban, dan lembaran jawaban peserta didik kelas X pada penilaian tengah semester ganjil mata pelajaran sejarah di SMK Negeri 1 Painan tahun pelajaran 2022/2023. Analisis data untuk penelitian ini menggunakan analisis kuantitatif, dilakukan dengan menggunakan program aplikasi Anates.4.0.9.

HASIL DAN PEMBAHASAN

Validitas Empiris

Untuk menentukan apakah suatu tes merupakan alat ukur yang baik, maka perlu dilakukan analisis validitas. Validitas adalah ketepatan mengukur sebuah butir soal dalam suatu tes hasil belajar. Berdasarkan analisis validitas dikonsultasikan dengan r_{tabel} pada taraf signifikan 5%. Apabila $r_{hitung} > r_{tabel}$ maka butir soal tersebut dikategorikan valid (signifikan/sangat signifikan). Pada taraf 5% ($N = 145$ menunjukkan nilai r_{tabel} sebesar 0,163.) Berdasarkan dari hasil analisis pada 30 butir soal menunjukkan butir soal yang valid (signifikan/sangat signifikan) sejumlah 29 butir soal

(96,7%), dan butir soal ujian yang tidak valid adalah 1 butir soal, berarti dari seluruh soal yang dianalisis (3,3%) soal tidak valid.

Tabel 2. Distribusi Validitas Empiris Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

No	Kriteria Soal	No Butir	Jumlah	Presentase
1.	Valid	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	29 Soal	96,7%
2.	Tidak Valid	17	1 Soal	3,3%

Sumber: Diolah Dari Data Primer (2023)

Butir soal dapat dikatakan mempunyai validitas yang tinggi (valid), jika dapat sepenuhnya mengukur ketercapaian tujuan pembelajaran peserta didik terhadap kompetensi yang telah ditetapkan. Skor pada butir soal dapat menyebabkan skor total menjadi tinggi atau rendah, karena itu sebuah item yang memiliki validitas tinggi jika skor pada item mempunyai kesejajaran dengan skor total. Nilai validitas pada butir soal mencapai (96,7%) dari total keseluruhan soal, hal ini menandakan bahwa terdapat banyak butir soal penilaian tengah semester ganjil pembelajaran sejarah di fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023 yang mampu mengukur capaian belajar peserta didik. Pada hasil perhitungan kategori validitas empirik butir soal yang termasuk kategori valid yaitu (96,7%), berarti butir soal tersebut sudah dapat menjalankan fungsinya yaitu dapat mengukur apa yang seharusnya diukur. Sedangkan validitas empirik pada butir soal yang termasuk kategori tidak valid terdapat (3%).

Berdasarkan analisis validitas empirik dapat disimpulkan bahwa soal ini berada di kategori baik untuk mengukur hasil belajar kognitif peserta didik. Butir soal yang baik disimpan sebagai bank soal agar dapat digunakan untuk penilaian tes selanjutnya. Sedangkan sebaliknya, butir soal yang tidak valid perlu diperbaiki dengan menyesuaikan pada indikator pembelajaran agar dapat mengukur capaian pembelajaran peserta didik.

Reliabilitas

Tes dikatakan reliabel yang artinya dipercaya jika suatu tes dapat mengukur secara konsisten sesuatu yang akan diukur dari waktu ke waktu. Kepercayaan berhubungan dengan ketepatan dan konsistensi. Tes dikatakan dapat dipercaya apabila memberikan hasil pengukuran hasil belajar yang relatif tetap secara konsisten (Yusrizal, 2016: 112). Hasil perhitungan diinterpretasikan dengan menggunakan koefisien reliabilitas, jika $r_{11} \geq 0,70$ maka reliabilitasnya dikatakan tinggi. Berdasarkan hasil analisis reliabilitas, nilai reliabilitas yang dimiliki soal penilaian tengah semester ganjil mata pelajaran sejarah fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023 yang diperoleh adalah 0,74. Angka reliabilitas ini menunjukkan soal tergolong pada kategori tinggi dan soal sudah reliabel atau perangkat tes tersebut sudah baik. Nilai reliabilitas berada pada rentangan 0,600 - 0,799 artinya soal memiliki reliabilitas yang tinggi. Soal yang memiliki reliabilitas yang tinggi berarti dapat dikatakan bahwa reliabilitas tes ini telah memenuhi persyaratan reliabilitas.

Tabel 3. Reliabilitas Tes Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

No	Rata-Rata	Simpang Baku	KorelasiXY	Reliabilitas Tes
1	13,14	4,75	0,58	0,74

Sumber: Diolah Dari Data Primer (2023)

Daya Beda

Daya pembeda soal adalah bentuk analisis untuk membedakan antara peserta didik yang berkemampuan tinggi dengan peserta didik yang berkemampuan rendah. Semakin tinggi nilai beda soal (bernilai positif maka semakin baik soal tersebut). Hasil analisis pada soal penilaian tengah semester ganjil pembelajaran sejarah fase E di SMK Negeri 1 Painan tahun pelajaran 2022/2023 terdapat butir soal yang tergolong jelek berjumlah 2 soal (6,7%), cukup berjumlah 12 soal (40%), baik berjumlah 16 soal (53,3%), dan baik sekali berjumlah 0 butir (0%).

Tabel 4. Distribusi Daya Pembeda Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

No	Kategori Soal	No Butir	Jumlah	Presentase
1.	0,00 – 0,20 (Jelek)	17, 22	2 Soal	6,7%
2.	0,21 – 0,40 (Cukup)	1, 4, 7, 8, 9, 13, 14, 15, 21, 26, 29, 30	12 Soal	40%
3.	0,41 – 0,70 (Baik)	2, 3, 5, 6, 10, 11, 12, 16, 18, 19, 20, 23, 24, 25, 27, 28	16 Soal	53,3%
4.	0,71 – 1,00 (Baik Sekali)	-	-	0%

Sumber: Diolah Dari Data Primer (2023)

Soal yang memiliki daya pembeda jelek, artinya butir soal tidak bisa membedakan peserta didik yang berkemampuan tinggi dan peserta didik yang berkemampuan rendah. Soal yang memiliki nilai daya beda cukup, artinya butir soal tersebut hampir bisa membedakan berkemampuan tinggi dan peserta didik yang berkemampuan rendah. Soal yang memiliki daya pembeda baik adalah butir soal yang sedikit bisa membedakan peserta didik yang berkemampuan tinggi dan peserta didik yang berkemampuan rendah. Sedangkan soal yang memiliki daya pembeda baik sekali, artinya butir soal yang mampu membedakan peserta didik yang berkemampuan tinggi dan peserta didik yang berkemampuan rendah. Soal yang sudah memiliki daya beda cukup dan baik berarti soal ini sudah bisa digunakan untuk tes kemampuan peserta didik, sedangkan soal yang memiliki daya pembeda yang jelek sebaiknya direvisi atau dihilangkan.

Berdasarkan analisis data, dari 30 butir soal yang tergolong jelek berjumlah 2 soal (6,7%). Soal yang memiliki daya pembeda jelek artinya soal tersebut banyak dijawab benar oleh peserta didik yang berkemampuan rendah dan banyak dijawab salah oleh peserta didik yang berkemampuan tinggi. Butir soal jelek menunjukkan peserta didik yang berkemampuan tinggi tidak dapat menjawab soal sedangkan peserta didik yang berkemampuan rendah dapat menjawab soal. Faktor daya

pembeda soal jelek salah satunya disebabkan karena menjawab soal dengan cara menebak saja. Ini berarti bahwa untuk menjawab soal dengan benar, dapat dilakukan dengan menebak isian soal. Soal dengan kategori jelek menunjukkan soal yang diberikan terlalu mudah sehingga semua peserta didik dapat menjawab soal dengan benar atau soal tersebut sukar sehingga peserta didik tidak dapat menjawab soal dengan benar (Anita et al., 2018: 41). Soal seperti ini seharusnya tidak digunakan kembali untuk ujian selanjutnya. Apabila ingin digunakan, maka soal perlu direvisi atau diperbaiki. Ini dilakukan agar soal dapat berfungsi sesuai dengan tujuan yang diharapkan.

Berdasarkan hasil analisis tersebut maka tindak lanjut untuk butir item yang memiliki daya pembeda cukup, baik, dan baik sekali dapat dimasukkan kedalam bank soal mata pelajaran sejarah, dan dapat dikeluarkan kembali pada tes selanjutnya. Butir item yang masih rendah atau memiliki kategori jelek ada dua kemungkinan tindak lanjut yang bisa dilakukan, yaitu 1) memperbaiki soal dan setelah diperbaiki dapat digunakan kembali, 2) membuang item tersebut.

Tingkat Kesukaran

Soal yang baik adalah soal yang tidak terlalu mudah dan tidak terlalu sukar. Soal yang terlalu mudah tidak membuat peserta didik berkeinginan untuk memecahkannya, sebaliknya soal yang terlalu sukar akan menyebabkan peserta didik menjadi putus asa dan tidak bersemangat untuk mencoba lagi karena diluar kemampuannya. Butir soal yang baik adalah butir soal dengan kategori sedang. Maka dari itu, dalam menyusun suatu penilaian ujian tengah semester ganjil sebaiknya digunakan butir soal yang mempunyai tingkat kesukaran berimbang yaitu sukar 25%, sedang 50%, dan mudah 25%. Analisis tingkat kesukaran menunjukkan apakah butir soal tes itu tergolong terlalu sulit, sedang, dan mudah, atau terlalu mudah. Melalui analisis tingkat kesukaran berdasarkan butir tes yang diujikan, akan dapat diungkapkan kelayakan soal tes, baik masing-masing butir soal tes maupun keseluruhan soal tes berdasarkan kategori tingkat kesukaran.

Hasil analisis soal penilaian tengah semester ganjil pembelajaran sejarah fase E di SMK Negeri 1 Painan tahun pelajaran 2022/2023 pada tingkat kesukaran, dari 30 butir soal yang tergolong sukar berjumlah 6 soal (20%), sedang berjumlah 19 soal (63,3%), mudah berjumlah 5 (16,7%). Secara keseluruhan kualitas soal berdasarkan tingkat kesukaran adalah sedang, karena butir soal yang berinterpretasi sedang lebih banyak dari butir soal yang berinterpretasi mudah dan sukar. Butir soal dengan tingkat kesukaran baik (dalam artian cukup atau sedang) butir soal bisa dimasukkan dalam bank soal dan dikeluarkan lagi untuk tes selanjutnya. Dengan demikian soal ini belum tergolong baik karena belum memenuhi standar tingkat kesukaran butir soal yang berimbang sesuai dengan persentase tingkat kesukaran yang berimbang.

Tabel 5. Distribusi Tingkat Kesukaran Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

No	Kategori Soal	No Butir	Jumlah	Presentase
1.	0,00 – 0,30	8, 12, 19, 20, 22, 29	6 Soal	20%

	(Sukar)			
2.	0,31 – 0,70 (Sedang)	2, 3, 5, 6, 7, 9, 11, 13, 14, 15, 17, 18, 21, 23, 24, 25, 27, 28, 30	19 Soal	63,3%
3.	0,71 – 1,00 (Mudah)	1, 4, 10, 16, 26	5 Soal	16,7%

Sumber: Diolah Dari Data Primer (2023)

Efektivitas Pengecoh

Pada opsi jawaban terdapat dua jenis opsi yaitu, opsi kunci jawaban, dan opsi pengecoh/*distractor*. Opsi pengecoh bukan hanya sekedar pelengkap jawaban, tetapi memang sengaja dibuat untuk mengecoh jawaban peserta didik yang tidak paham dengan materi pembelajaran. Oleh sebab itu opsi pengecoh yang baik harus homogen dengan kunci jawaban, atau dibuat semirip mungkin dengan kunci jawaban. Butir soal perlu dilakukan analisis pengecoh untuk mengetahui efektif atau tidaknya pengecoh pada tes objektif bentuk pilihan ganda. Suatu pengecoh dapat dikatakan berfungsi dengan baik jika paling sedikit dipilih oleh 5% dari seluruh peserta didik. Adapun hasil analisis efektivitas pengecoh, dari 30 butir soal didapatkan pengecoh (*distractor*) yang berfungsi secara efektif dan tidak efektif.

Pada hasil analisis setiap *option* pada butir soal, pengecoh yang bernilai positif yaitu dengan kriteria sangat baik dan baik, hal ini berarti pengecoh tersebut tidak dapat membedakan kemampuan peserta didik, yaitu kemampuan peserta didik yang memperoleh skor tinggi menjawab salah soal tersebut. Menurut Ugi & Darma Ekawati (2015: 125) penyebabnya adalah umumnya rumusan kunci jawaban yang disediakan kurang baik serta peserta didik kurang memahami konsep yang diujikan dalam soal-soal penilaian tengah semester ganjil pembelajaran sejarah fase E di SMK Negeri 1 Painan tahun pelajaran 2022/2023. Nilai negatif pada pengecoh yaitu dengan kriteria kurang baik, buruk, dan sangat buruk menunjukkan bahwa pengecoh sudah berfungsi dengan baik dimana peserta tes yang skornya rendah memilih pengecoh sebagai jawaban yang benar.

Tabel 6. Distribusi Efektivitas Pengecoh Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

Kategori	Nomor Butir	Jumlah	Presentase
Berkualitas	1, 2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30	29 Soal	96,7%
Tidak Berkualitas	7	1 Soal	3,3%

Sumber: Diolah Dari Data Primer (2023)

Berdasarkan hasil analisis soal penilaian tengah semester ganjil pembelajaran sejarah di fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023, pada efektivitas pengecoh didapatkan kategori opsi pengecoh yang berkualitas berjumlah 29 butir soal (96,7%) dan yang tidak berkualitas berjumlah 1 butir soal (3,3%). Persentase tersebut diperoleh dengan membandingkan opsi pengecoh yang masuk kriteria berkualitas atau

tidak berkualitas dengan keseluruhan opsi pengecoh/ *distractor*. Pengecoh yang buruk adalah pengecoh yang sama sekali tidak dipilih oleh peserta didik. Hal ini dapat terjadi karena pengecoh terlalu mencolok dan dimengerti oleh semua peserta didik sebagai pengecoh jawaban. Pengecoh yang kurang baik dan tidak baik mengindikasikan bahwa pengecoh tersebut terlalu mencolok dan heterogen, sehingga pengecoh tersebut tidak memiliki daya tarik untuk dipilih peserta tes yang kurang memahami materi. Pengecoh yang tidak efektif direkomendasikan untuk diganti dengan pengecoh yang lebih baik/ direvisi. Sedangkan pengecoh yang baik adalah pengecoh yang tingkat kerumitan atau tingkat kesederhanaan serta panjang pendeknya relatif sama dengan kunci jawaban.

Dapat disimpulkan bahwa kualitas opsi pengecoh pada soal penilaian tengah semester ganjil pembelajaran sejarah di fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023 ini sudah memenuhi kriteria sangat bagus karena banyak opsi pengecoh yang berfungsi.

Kualitas Butir Soal

Analisis butir soal akan membantu guru untuk meningkatkan kualitas butir soal yang telah disusun. Soal yang baik adalah soal yang sudah memenuhi semua kriteria yang ada dari validitas, daya beda, tingkat kesukaran, dan efektivitas pengecoh butir soal. Ketika penyusunan soal sudah memenuhi kriteria maka kualitas butir soal pun akan meningkat sehingga sebuah soal tidak hanya sebagai alat ukur yang tidak mengukur apa yang hendak diukur akan tetapi setiap ranah dari peserta didik atau hal yang harus dievaluasi dari peserta didik dapat tergambar dari soal yang telah disusun (Fitrianawati, 2015: 290). Hal ini akan berdampak pada kualitas butir soal sehingga mengakibatkan hasil belajar peserta didik dapat dipercaya. Hasil analisis data soal penilaian tengah semester ganjil pembelajaran sejarah di fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023 dari keseluruhan butir soal ditinjau dari aspek validitas empiris, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh dapat dilihat pada Tabel 7.

Tabel 7. Analisis Data Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

No	Aspek yang Dianalisis	Kriteria Soal
1.	Validitas Empiris	29 Soal (Valid)
2.	Reliabilitas	0,74 (Tinggi)
3.	Daya Pembeda	28 Soal (Baik)
4.	Tingkat Kesukaran	19 Soal (Baik)
5.	Efektivitas Pengecoh	Kualitas Opsi 97% (Berkualitas)

Sumber: Diolah dari Data Primer (2023)

Kualitas butir soal diadaptasi dari skala *Likert* dibagi menjadi 5 kategori yaitu: sangat baik, baik, sedang, tidak baik, dan sangat tidak baik.

1. Butir soal dikatakan sangat baik jika memenuhi 4 kriteria soal yang baik yaitu validitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh. Butir soal ini dapat disimpan pada bank soal untuk digunakan kembali.

2. Butir soal dikatakan baik jika memenuhi 3 kriteria dari 4 kriteria yang baik. Butir soal ini belum bisa disimpan pada bank soal, dan harus direvisi agar memenuhi 4 kriteria.
3. Butir soal dikatakan sedang jika memenuhi 2 kriteria dari 4 kriteria soal yang baik. Butir soal ini belum bisa disimpan pada bank soal, dan harus direvisi agar memenuhi 4 kriteria.
4. Butir soal dikatakan tidak baik jika memenuhi 1 dari 4 kriteria soal yang baik. Butir soal ini tidak bisa disimpan pada bank soal, karena harus direvisi secara menyeluruh sehingga lebih baik dibuang.
5. Selain syarat berbasis butir, maka soal secara keseluruhan harus reliabel (Werdiningsih, 2015).

Hasil rekapitulasi semua aspek analisis butir soal penilaian tengah semester ganjil pembelajaran sejarah di fase E SMK Negeri 1 Painan tahun pelajaran 2022/2023 yang terdiri dari validitas empirik, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh didapatkan hasil soal pilihan ganda yang sangat baik berjumlah 17 butir soal (56,7%), soal yang berkualitas baik berjumlah 11 butir soal (36,6%), soal yang berkualitas sedang berjumlah 2 butir soal (6,7%), soal yang berkualitas tidak baik berjumlah 0 butir soal (0%), dan soal yang berkualitas sangat tidak baik berjumlah 0 butir soal (0%). Hasil analisis rekap soal yang dapat dipakai dapat dilihat pada tabel 8.

Tabel 8. Distribusi Efektivitas Pengecoh Soal Penilaian Tengah Semester Ganjil Mata Pelajaran Sejarah di Fase E SMK Negeri 1 Painan Tahun Pelajaran 2022/2023

Kriteria Kualitas Soal	Nomor Butir	Jumlah	Presentase
Sangat Baik	2, 3, 5, 6, 9, 11, 13, 14, 15, 18, 21, 23, 24, 25, 27, 28, 30	17 Soal	56,7%
Baik	1, 4, 7, 8, 10, 12, 16, 19, 20, 26, 29	11 Soal	36,6%
Sedang	17, 22	2 Soal	6,7%
Tidak Baik	-	-	0%
Sangat Tidak Baik	-	-	0%

Sumber: Diolah Dari Data Primer (2023)

Kualitas butir soal yang baik adalah soal yang sudah memenuhi semua kriteria yang ada, yaitu butir soal yang telah diujikan tersebut sudah memenuhi kriteria analisis soal berdasarkan analisis kuantitatif dilihat dari segi validitas, reliabilitas, daya beda, tingkat kesukaran, dan efektivitas pengecoh. Kualitas soal terbagi menjadi 5 kategori yaitu, sangat baik, baik, sedang, tidak baik, dan sangat tidak baik menurut skala likert. Butir soal dengan kualitas sangat baik bisa dimasukkan ke dalam bank soal dan menjaga kerahasiaan soal tersebut sehingga dapat digunakan lagi sebagai instrumen evaluasi (tes) yang akan datang. Butir soal dengan kualitas baik dan sedang belum bisa dimasukkan ke dalam bank soal karena harus dilakukan revisi terlebih dahulu sesuai dengan indikator kegagalannya. Sedangkan butir soal dengan kualitas tidak

baik dan sangat tidak baik lebih baik dibuang karena butir tersebut membutuhkan revisi yang signifikan.

SIMPULAN

Berdasarkan hasil penelitian analisis 30 butir soal pilihan ganda, ditinjau dari validitas empiris terdapat 96,7% butir soal yang termasuk kategori valid. Ditinjau dari reliabilitas, soal termasuk memiliki kategori reliabilitas tinggi yaitu 0,74. Ditinjau dari daya pembeda, terdapat 93,3% butir soal yang memiliki daya pembeda baik. Ditinjau dari tingkat kesukaran, terdapat 63,3% butir soal yang memiliki tingkat kesukaran baik. Ditinjau dari efektivitas pengecoh, terdapat 96,7% butir soal yang termasuk kategori opsi pengecoh berkualitas. Dapat disimpulkan bahwa soal sudah memenuhi kriteria baik untuk mengukur hasil belajar kognitif peserta didik pada mata pelajaran sejarah di fase E SMK Negeri 1 Painan.

DAFTAR PUSTAKA

- Anita, A., Tyowati, S., & Zulfadrial, Z. 2018. Analisis Kualitas Butir Soal Fisika Kelas X Sekolah Menengah Atas. *Edukasi: Jurnal Pendidikan*, 16(1), 35. <https://doi.org/10.31571/edukasi.v16i1.780>
- Arikunto, S. 2013. *Dasar-Dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.
- Fitrianawati, M. 2015. Peran Analisis Butir Soal Guna Meningkatkan Kualitas Butir Soal, Kompetensi Guru Dan Hasil Belajar Peserta Didik. *Prosiding Seminar Nasional Pendidikan PGSD UMS & HDPGSDI Wilayah Jawa*, 5(3), 282–295.
- Mamik. (2015). *Metodologi Kualitatif*. Sidoarjo: Zifatama Publisher.
- Nengsi, A. R. 2018. Pengembangan tes prestasi bentuk pilihan ganda untuk evaluasi hasil belajar geografi siswa SMA kelas IX. *UNES Journal of Education Scienties*, 2(2), 118–126. <http://lppm.ojs.unespadang.ac.id/index.php/UJES>
- Nugraha, A. S., & Kuswono. 2019. Pengembangan Alat Evaluasi Pembelajaran Sejarah Menggunakan Aplikasi Kahoot Dengan Pola Berpikir Kronologis Siswa Pada Materi Sejarah Kelas Xi IPS Di SMA Negeri 1 Kotagajah. *Jurnal Swarnadwipa*, 3(2), 54–63.
- Nurlan, F. 2019. *Metodologi Penelitian Kuantitatif*. Semarang: CV. Pilar Nusantara.
- Ramdhan, M. 2021. *Metode Penelitian*. Surabaya: Cipta Media Nusantara (CMN).
- Syafril, & Zelhendri, Z. 2017. *Dasar-Dasar Ilmu Pendidikan*. Depok: Kencana.
- Toijo, A. K. B., & Hulukati, W. 2013. Kualitas Tes Buatan Guru. In *Normalitas* (Vol. 1, Issue 1, pp. 129–141).
- Toijo, A. K. B., & Hulukati, W. 2013. Kualitas Tes Buatan Guru. In *Normalitas* (Vol. 1, Issue 1, pp. 129–141).
- Ugi, L. E., & Darma Ekawati. 2015. Kualitas Tes Pra Olimpiade Bidang Studi Matematika Tingkat SMP di Kota Baubau. *Prosiding Seminar Nasional*, 2(1), 118–127.
- Wibawa, E. A. (2019). Karakteristik Butir Soal Tes Ujian Akhir Semester Hukum Bisnis. *Jurnal Pendidikan Akuntansi Indonesia*, 17(1), 86–96. <https://doi.org/10.21831/jpai.v17i1.26339>
- Yusrizal. 2016. *Tanya Jawab Seputar Pengukuran, Penilaian, dan Evaluasi Pendidikan*. Banda Aceh: Syiah Kuala University Press.