

**Analisis Butir Soal Pilihan Ganda sebagai Alat Evaluasi Penilaian Akhir Tahun (PAT) pada Mata Pelajaran Sejarah Indonesia Kelas XI SMA Negeri 2 Pandeglang Tahun Ajaran 2021-2022**

**Muhammad Syahrone<sup>1</sup>, Ana Nurhasanah<sup>2</sup>, Arif Permana Putra<sup>3</sup>**

Universitas Sultan Ageng Tirtayasa

2288160039@untirta.ac.id., Ananur74@untirta.ac.id, arif.permana@untirta.ac.id.

**ABSTRACT**

*This study aims to determine the validity, reliability, level of difficulty, discriminating power, and the effectiveness of the multiple-choice distractor as an evaluation tool for history subjects in class XI SMAN 2 Pandeglang. This research is quantitative and descriptive. Students of class XI B and D IPS at SMAN 2 Pandeglang were used as the research sample. The documentation method is used to perform data collection techniques. Using the Microsoft Excel program, a manual calculation program is used to analyze the data obtained. Based on the results of the research on 30 year-end assessment questions (PAT) for Indonesian history class XI IPS SMAN 2 Pandeglang for the 2021/2022 academic year, there were 13 valid questions (43.3%) and 17 questions (56.7%) which invalid. The reliability score of 0.49 indicates that neither the questions nor the results are reliable. On discriminating power, 20 items (66.7%) fall into the weak category, 7 items (23.3%) fall into the sufficient category, 1 (one) item (3.3%) falls into the good category, 2 (two) items (6.7%) fall into the very weak category, and there are no items in the very good category. The index of difficulty shows that 4 (four) items or 13.3% are included in the easy category, 26 items or 86.7% are classified as difficult, and there are no items included in the very difficult category. On the effectiveness of the distractor, all items (30) have the effectiveness of the distractor with good criteria. The results of the validation by the expert lecturer on the study of the questions showed that 27 out of 30 items had fulfilled the suitability aspect of the items with indicators, 29 out of 30 items were in accordance with scientific principles, all items (30) had one answer key. Meanwhile, the results of the validation of the construction of the questions showed that all questions met the aspects of the rules for writing items. And, as many as 27 of the 30 items were stated to fulfill the language rules in writing the items.*

**Keyword:** *Item Analysis, Multiple Choice, Indonesian History.*

**ABSTRAK**

Penelitian ini bertujuan untuk mengetahui validitas, reliabilitas, tingkat kesukaran, daya pembeda, dan efektivitas pengecoh soal pilihan ganda sebagai alat evaluasi mata pelajaran sejarah di kelas XI SMAN 2 Pandeglang. Penelitian ini bersifat kuantitatif dan deskriptif. Siswa kelas XI B dan D IPS di SMAN 2 Pandeglang sebagai sampel penelitian. Metode dokumentasi digunakan untuk melakukan teknik pengumpulan data. Perhitungan manual menggunakan bantuan aplikasi Microsoft Excel digunakan untuk menganalisis data yang diperoleh. Berdasarkan hasil penelitian pada 30 butir soal Penilaian Akhir Tahun (PAT) mata

---

<sup>1</sup> Penulis 1, Mahasiswa Jurusan Pendidikan Sejarah FKIP, Universitas Sultan Ageng Tirtayasa

<sup>2</sup> Penulis 2, Dosen Jurusan Pendidikan Sejarah FKIP, Universitas Sultan Ageng Tirtayasa

<sup>3</sup> Penulis 3, Dosen Jurusan Pendidikan Sejarah FKIP, Universitas Sultan Ageng Tirtayasa

pelajaran Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun ajaran 2021/2022 terdapat 13 butir soal valid (43,3%) dan 17 butir soal (56,7%) yang tidak valid. Angka reliabilitas 0,49 menunjukkan bahwa baik soal maupun hasil tidak reliabel. Pada daya pembeda, 20 butir soal (66,7%) masuk dalam kategori lemah, 7 butir soal (23,3%) masuk dalam kategori cukup, 1 (satu) butir soal (3,3%) masuk dalam kategori baik, 2 (dua) butir soal (6,7%) masuk dalam kategori sangat lemah, dan tidak ada terdapat butir soal dengan kategori sangat baik. Indeks kesukaran menunjukkan bahwa 4 (empat) butir soal atau 13,3% termasuk dalam kategori mudah, 26 butir soal atau 86,7% yang tergolong sukar, dan tidak terdapat butir soal yang termasuk dalam kategori sangat sukar. Pada efektivitas pengecoh seluruh butir soal (30) memiliki keefektifan pengecoh dengan kriteria baik. Hasil validasi oleh dosen ahli terhadap penelaahan soal menunjukkan bahwa 27 dari 30 butir soal telah memenuhi aspek kesesuaian butir soal dengan indikator, 29 dari 30 butir soal sesuai dengan kaidah keilmuan, seluruh butir soal memiliki satu kunci jawaban. Sementara itu, hasil validasi terhadap konstruksi soal menunjukkan bahwa seluruh soal telah memenuhi aspek kaidah penulisan butir soal. Dan, sebanyak 27 dari 30 butir soal dinyatakan memenuhi kaidah bahasa dalam penulisan butir soal.

**Kata Kunci:** Analisis Butir Soal, Pilihan Ganda, Sejarah Indonesia.

## PENDAHULUAN

Idealnya, instrumen atau soal yang digunakan untuk mengukur hasil belajar siswa harus dirancang dengan cermat, disusun dengan menggunakan kaidah penulisan soal, diuji, dan dievaluasi kualitasnya. Membuat kisi-kisi soal terlebih dahulu, menyusun pedoman penskoran dan menerapkannya untuk menghindari hasil penilaian yang bias pada soal uraian, menyusun soal dan pengecoh kunci jawaban yang tepat, memperhatikan kualitas instrumen atau soal yang harus disusun agar sesuai dengan kriteria soal yang baik dan kaidah penyusunan soal, menguji soal dan menganalisisnya guna mengetahui kualitas dari setiap soal yang disusun merupakan hal-hal yang perlu diperhatikan oleh pendidik agar terhindar dari permasalahan tersebut. Diharapkan guru mengembangkan soal-soal yang berkualitas dan mampu mengukur hasil belajar siswa secara akurat untuk memastikannya selaras dengan tujuan pembelajaran (Kusnandar, 2014: 65).

Kendala pada tes saat ini adalah kumpulan soal yang digunakan sebagai alat evaluasi belum dianalisis secara menyeluruh sehingga kualitas soal masih belum jelas. Kualitas siswa dapat dipengaruhi oleh masalah yang signifikan ini. Menurut Jandaghi & Shaterian (2018), yang menerbitkan penelitian mereka dalam jurnal *Validity, Reliability, and Difficulty Indies for Instructor-Built Exam Question*, kualitas siswa adalah salah satu masalah paling signifikan dalam sistem belajar mengajar. Agar soal ujian memiliki kualitas yang lebih tinggi dari standar kelulusan yang diharapkan, maka perlu ada standar. Pertanyaan pada ujian berdampak besar pada seberapa baik siswa melakukannya. Untuk mendapatkan hasil yang baik, soal ujian harus diperiksa tingkat kesulitan, daya pembeda, validitas, dan reliabilitasnya (Jandaghi & Shaterian, 2018: 36).

Soal-soal yang telah disiapkan harus diuji terlebih dahulu. Hal ini diharapkan dapat menentukan sifat pertanyaan agar dapat mengukur kemampuan siswa. Nilai atau hasil ujian akhir siswa akan terpengaruh jika soal-soal pendidik tidak terkenal

kualitasnya (Kunandar, 2014: 167). Mengetahui karakteristik pertanyaan yang baik dapat membantu menentukan kualitasnya. Validitas, ketergantungan, tingkat kesulitan, daya pembeda, dan keefektifan distraktor untuk soal pilihan ganda merupakan komponen penting dari sebuah soal yang baik (Arikunto, 2013: 222). Pemerintah berupaya mengatasi masalah ini dengan menerbitkan buku panduan menyusun soal yang baik, menyusun soal-soal berjenis HOTS (*Higher Order Thinking Skill*), dan penilaian hasil belajar SD dan SMP, selain merevisi kurikulum 2013 untuk meningkatkan mutu pendidikan. Namun tidak semua pendidik melakukan penilaian sesuai dengan persyaratan yang ditetapkan oleh pemerintah. Mencermati pernyataan-pernyataan di atas, maka ditegaskan bahwa kebenaran sebagian besar tes yang digunakan sebagai perangkat penilaian pembelajaran belum dirinci dari atas ke bawah. Pendidik harus melakukan kegiatan analisis untuk memastikan soal tes yang diberikan kepada peserta didik berkualitas tinggi dan selaras dengan tujuan pengukuran pendidik (Arikunto, 2013: 222).

Pendidik belum melakukan analisis secara komprehensif terhadap butir-butir soal penilaian akhir tahun yang diberikan kepada siswa, sesuai observasi yang dilakukan pada tanggal 22 Januari 2021 di SMA Negeri 2 Pandeglang. Ini karena guru tidak memiliki banyak waktu, sehingga tidak mungkin melakukan analisis menyeluruh. Selama ini pendidik sudah dapat menentukan bagus atau tidaknya suatu soal berdasarkan lulus tidaknya siswa dan seberapa sulit soal berdasarkan kata kerja operasional dalam soal. Tentunya hal ini akan mengakibatkan tidak diketahuinya kualitas perangkat tes sebagai alat evaluasi dari segi validitas, dependabilitas, daya pembeda, tingkat kesukaran, dan efektivitas distraktor. Dengan menganalisis butir soal dapat ditentukan mana yang harus direvisi dan mana yang harus dihilangkan, sehingga dapat mengidentifikasi soal yang bermutu dan bermanfaat sebagai instrumen pengukuran. Guru akan dapat memperoleh hasil evaluasi yang akurat berkat kualitas soal yang tinggi. Temuan evaluasi akan mengungkapkan apa yang perlu dilakukan selanjutnya, serta informasi tentang umpan balik belajar siswa, kemajuan belajar, dan program pembelajaran. Sebagai alat evaluasi pembelajaran, rekomendasi dari kajian-kajian yang relevan untuk menjawab permasalahan tersebut di atas harus dianalisis.

Dalam mata pelajaran seperti sejarah, ada juga masalah kualitas soal. Tujuan dari penelitian ini adalah untuk meningkatkan kualitas pembelajaran sejarah di sekolah dan hasil belajar bagi siswa. Pertanyaan yang akan dianalisis adalah apa yang membedakan penelitian saat ini dari penelitian sebelumnya. Penelitian ini dilakukan pada soal pilihan ganda akhir tahun (PAT) karena ujian sumatif ini berdampak signifikan terhadap hasil belajar siswa dan menentukan apakah siswa dapat melanjutkan ke materi selanjutnya atau mendapat kenaikan nilai. Selain itu, masih menjadi tantangan bagi para peneliti untuk menemukan studi tentang subjek analisis sejarah. Mayoritas analisis dilakukan di bidang sains dan ilmu sosial, seperti akuntansi dan ekonomi. Penelitian ini diperlukan karena banyaknya permasalahan dan kurangnya penelitian tentang analisis masalah pada mata kuliah sejarah. Penelitian akan dilakukan terhadap analisis soal pilihan ganda yang berkaitan dengan

tingkat kesukaran soal, validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan keefektifan delusi berdasarkan penjelasan pendapat ahli, kenyataan, dan hasil observasi.

#### **METODE PENELITIAN**

Penelitian ini menggunakan penelitian evaluatif sebagai metodenya. Penelitian yang mensyaratkan adanya kriteria, tolok ukur, atau standar yang digunakan sebagai pembanding antara data yang diperoleh dengan data yang telah diolah merupakan kondisi sebenarnya dari objek tersebut disebut penelitian evaluasi. Selain itu, kesenjangan antara kondisi aktual dan kondisi yang diharapkan yang dituangkan dalam kriteria merupakan tujuan untuk mendapatkan gambaran umum apakah subjek penelitian memenuhi kriteria atau tidak. Menurut Arikunto, tujuan dari penelitian ini adalah membandingkan data atau informasi yang telah dikumpulkan dengan kriteria dan menarik kesimpulan (Arikunto, 2014:36).

Penelitian ini menggunakan kombinasi metode kuantitatif dan kualitatif (Creswell, 2016: 5). Pendekatan campuran adalah pendekatan di mana desain yang berbeda dan kumpulan data kuantitatif dan kualitatif digabungkan. Jika dibandingkan dengan satu pendekatan, kombinasi ini mungkin menawarkan pemahaman yang lebih komprehensif. Kualitas soal sumatif semester genap pada mata pelajaran sejarah kelas XI SMA Negeri 2 Pandeglang diukur dengan menggunakan pendekatan kuantitatif, yang akan ditunjukkan dengan menghitung angka. Sedangkan metodologi subyektif digunakan untuk memahami informasi eksplorasi secara keseluruhan. sehingga data numerik yang dihasilkan dari kegiatan analisis kuantitatif dapat dipahami secara menyeluruh dan mudah.

Penelitian ini dilaksanakan di kelas XI IPS SMA Negeri 2 Pandeglang dengan menggunakan pembelajaran UKBM (Unit Kegiatan Belajar Mandiri) yang digabung dengan program sistem kredit semester.

Pemilahan informasi direncanakan untuk memperoleh data yang relevan, tepat dan dapat dimanfaatkan sebagaimana mestinya sesuai dengan tujuan pembelajaran. Wawancara dan dokumentasi adalah metode yang digunakan dalam penelitian untuk mengumpulkan data. Wawancara adalah metode pengumpulan informasi dengan mengajukan pertanyaan yang relevan. Catatan sejarah disebut dokumentasi. Dokumen dapat berupa karya tulis, visual, atau monumental yang dilakukan oleh seorang individu (Sugiyono, 2013: 82). Seperti yang dikemukakan oleh Arikunto (2014: 158), kata "dokumen" mengacu pada bahan tertulis dan merupakan asal mula dari dokumentasi. Penelitian menggunakan metode dokumentasi untuk melihat hal-hal tertulis seperti buku, majalah, dokumen, peraturan, risalah rapat, dan sebagainya.

Wawancara terbuka digunakan untuk mengumpulkan informasi tentang penyusunan soal dan penggunaan ujian formatif dalam penelitian ini. Soal pilihan ganda juga digunakan sebagai metode dokumentasi pada siswa kelas XI B dan D IPS di SMAN 2 Pandeglang. Pendekatan dokumentasi diawali dengan pengumpulan data prestasi belajar siswa, silabus pembelajaran, makalah soal, makalah sumatif, dan

kunci jawaban soal semester genap dari guru sejarah, serta tanggapan seluruh siswa kelas XI IPS terhadap soal sumatif yang diujikan.

Teknik analisis data yang digunakan adalah *Item Response Theory* (IRT), juga dikenal sebagai teori menjawab pertanyaan, digunakan dalam proses menganalisis kualitas pertanyaan. Fungsi matematika digunakan dalam teori ini untuk membangun hubungan antara kemampuan siswa dan kemungkinan mereka akan menjawab pertanyaan dengan benar. *Item Response Theory* (IRT) adalah hubungan antara kemampuan atau tingkat pencapaian siswa dan kemungkinan menjawab item pertanyaan dengan benar. Rumus berikut akan digunakan untuk melakukan analisis kuantitatif dengan bantuan program komputer, khususnya Microsoft Office Excel 2010.

Analisis yang dilakukan meliputi uji validitas, uji reliabilitas, uji tingkat kesukaran, uji daya pembeda, dan uji efektivitas pengecoh. Serta, penarikan kesimpulan sebagai bagian akhirnya.

## **HASIL DAN PEMBAHASAN**

### **Analisis Kualitatif**

Analisis ini dilakukan sambil mempelajari soal-soal sebelum ujian. Analisis ini, juga dikenal sebagai analisis logis, memerlukan persiapan pertanyaan, pembahasan pertanyaan-pertanyaan tersebut, dan penentuan apakah fungsi masalah didasarkan pada aspek material (Sudarsono, 2012: 138). Kajian dapat memanfaatkan lembar kajian yang telah disiapkan. Tujuan dari lembar kajian ini adalah untuk mempermudah proses implementasi (Kusaeri dan Suprananto, 2012: 166).

Hasil analisis kualitatif sesuai dengan kisi-kisi soal, namun tiga soal—nomor 13, 23, dan 24—tidak sesuai dengan indikatornya. Dari segi isi materi sudah benar secara ilmiah, satu soal kurang tepat, nomor 17. Ada dua aspek penilaian yang kurang tepat dari segi konstruksi. Pilihan jawaban pertama sudah jelas tertera pada soal 13 dan 28. Angka 8 dan 14 merupakan pilihan kedua dari segi waktu dan angka. Ada dua aspek penilaian yang kurang tepat dari segi bahasa; yang pertama adalah aspek komunikatif bahasa, yang tercantum pada nomor 29. Pada nomor 6 dan 14 terdapat dua kalimat gramatikal.

### **Analisis Kuantitatif**

#### **a. Validitas**

Indikator ketepatan atau validitas instrumen adalah ukuran atau standar yang valid. Pengujian keabsahan benda pada soal Evaluasi Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun pelajaran 2021/2022 menggunakan persamaan hubungan item kedua ( $r_{xy}$ ).

Rumus korelasi *product moment* ( $r_{xy}$ ) dapat digunakan untuk menentukan validitas suatu soal. Indeks korelasi  $r_{xy}$  diperoleh dari perhitungan konsultasi dengan  $r_{tabel}$  pada taraf signifikansi 5%, tergantung dari jumlah siswa yang mengikuti tes. Jika  $r_{hitung}$  lebih besar dari  $r_{tabel}$ , maka syarat untuk melihat soal adalah benar. Di SMAN 2 Pandeglang terdapat 70 siswa yang mengikuti kelas

XI IPS B dan D. Setelah melihat  $r_{\text{tabel}}$ , didapatkan 68 peserta dengan nilai 0,2199. Jika  $r_{\text{hitung}}$  kurang dari 0,2199, maka item dianggap valid.

Hasil analisis menunjukkan bahwa soal Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI B dan D di SMAN 2 Pandeglang tahun ajaran 2021/2022 terdapat 13 soal yang dinyatakan valid atau 43,3%, dan 17 soal yang dinyatakan tidak sah atau 56,7%. Soal yang valid (43,3%) menunjukkan bahwa butir soal telah dapat menjalankan fungsinya yaitu mengukur apa yang seharusnya diukur. Lalu, ada berbagai alasan mengapa 56,7% soal tidak valid. Hal ini sesuai dengan teori Grounlund dalam Zainal Arifin (2014:247) yang menyatakan bahwa faktor administrasi dan penilaian, faktor jawaban siswa, dan faktor instrumen semuanya memiliki pengaruh terhadap validitas hasil tes. Kecenderungan siswa untuk menjawab dengan cepat dan salah dapat berdampak pada soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia tahun pelajaran 2021/2022.

Uraian di atas menunjukkan bahwa soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia Kelas XI IPS Tahun Pelajaran 2021/2022 memiliki kualitas validitas yang rendah. Bank soal dapat digunakan untuk menyimpan butir soal yang valid. Sebaliknya, item yang tidak valid harus diperbaiki dengan menyesuaikan indikator dan meningkatkan kemampuan teknis dalam penyusunan item.

#### **b. Reliabilitas**

Perhitungan reliabilitas pada soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun ajaran 2021/2022 dilakukan secara manual menggunakan bantuan program Microsoft Excel menggunakan rumus KR 20. Reliabilitas merupakan tingkat atau derajat konsistensi dari sebuah instrumen. Pengukuran yang memiliki reliabilitas tinggi jika pengukuran tersebut mampu menghasilkan data yang reliabel. Reliabilitas skor diukur dengan K-R 2, pada  $r_{11} \geq 0,70$  maka soal yang diuji memiliki reliabilitas yang tinggi, tetapi apabila  $r_{11} \leq 0,70$  maka soal yang diujikan memiliki reliabilitas yang rendah atau tidak reliabel.

Berdasarkan hasil penelitian yang dilakukan secara manual menggunakan bantuan program Microsoft Excel menunjukkan bahwa soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun ajaran 2021/2022 pada soal pilihan ganda memiliki reliabilitas sebesar  $r_{11} = 0,49$ .

Hasil perhitungan tersebut menunjukkan bahwa soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun ajaran 2021/2022 pada soal pilihan ganda memiliki tingkat reliabilitas yang rendah karena memiliki koefisien reabilitas ( $r_{11}$ )  $\leq 0,70$ . Rendahnya koefisien reliabilitas soal karena terbatasnya jumlah butir soal yang dibuat oleh guru, sehingga guru dapat menambahkan jumlah soal yang valid. Suatu instrumen tes yang memiliki validitas yang baik pada setiap butirnya juga akan memiliki

tingkat reliabilitas yang tinggi juga. Hal ini selaras dengan teori dari Suharsimi Arikunto (2013: 101) yang menyatakan bahwa tes yang terdiri dari banyak butir akan lebih valid daripada tes yang terdiri dari beberapa butir saja. Tinggi rendahnya tingkat validitas dapat menunjukkan tinggi rendahnya koefisien reliabilitas, sehingga semakin panjang tes maka reliabilitasnya semakin tinggi. Berdasarkan uraian tersebut dapat disimpulkan bahwa soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun ajaran 2021/2022 soal pilihan ganda merupakan soal yang kurang baik dilihat dari segi reliabilitasnya.

**c. Daya Pembeda**

Kemampuan suatu item untuk membedakan antara peserta dengan skor tinggi dan yang skor rendah dikenal sebagai daya pembeda. Jika pertanyaan dijawab dengan benar hanya oleh siswa yang cerdas, daya pembeda itu bermanfaat. Dengan menggunakan program Microsoft Excel, daya pembeda soal Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun pelajaran 2021/22 dihitung secara manual. Terjemahan efek samping dari perhitungan pemisahan daya menggunakan aturan yang menyertainya.

*D : 0,00 – 0,20 : jelek (poor)*

*D : 0,21 – 0,40 : cukup (satisfactory) D : 0,41 – 0,70 : baik (good)*

*D : 0,71 – 1,00 : baik sekali (excellent)*

*D : negatif, semuanya tidak baik. Jadi semua butir soal yang mempunyai nilai D negatif sebaiknya dibuang saja.*

(Suharsimi Arikunto, 2013:232)

Hasil analisis soal pilihan ganda Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS B dan D di SMAN 2 Pandeglang diketahui bahwa 20 soal (atau 66,7%) termasuk dalam kategori kurang baik, 7 soal (atau 23,3%) termasuk dalam kategori tidak baik. pada kategori cukup, 1 item (atau 3,3%) termasuk dalam kategori baik, 2 item (atau 6,7%) termasuk dalam kategori sangat lemah, dan 0 pertanyaan atau 0% memenuhi kriteria sangat tinggi. Soal dengan daya pembeda yang rendah tidak memungkinkan untuk membedakan kemampuan siswa. sehingga siswa yang berkemampuan rendah dapat menjawab soal dengan benar dan siswa yang berkemampuan tinggi dapat menjawab dengan salah.

Zainal Arifin (2014:273), perhitungan daya pembeda adalah ukuran sejauh mana suatu soal mampu membedakan siswa yang menguasai materi dengan siswa yang tidak/kurang menguasai materi berdasarkan kriteria tertentu. Anas Sudijono memberikan dukungan yang menyatakan bahwa penyusun soal harus menyadari bahwa kemampuan siswa berbeda-beda dan mengetahui daya pembeda setiap butir sangat penting karena merupakan salah satu pedoman menyusun butir soal (Anas Sudijono, 2011:386).

Soal Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun ajaran 2021/2022 memiliki daya diskriminasi yang kurang

baik karena mayoritas butir atau 66,7% masuk dalam kategori Buruk. Kesimpulan ini dapat ditarik dari pembahasan di atas. Bank soal tes hasil belajar selanjutnya harus berisi soal-soal dengan daya pembeda cukup, baik, dan sangat baik; item dengan daya pembeda yang buruk dapat diperbaiki. Sebaiknya tidak digunakan lagi untuk tes yang akan datang, terutama untuk soal-soal dengan daya pembeda negatif karena soal-soal tersebut kualitasnya sangat buruk.

**d. Tingkat Kesukaran**

Tingkat kesulitan setiap soal dapat digunakan untuk menentukan kualitasnya. Ada tiga tingkat kesulitan: mudah, sedang, dan sulit. Berdasarkan hasil temuan analisis soal pilihan ganda Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia di SMAN 2 Pandeglang, terdapat empat soal yang termasuk dalam kategori mudah (13,3%), 26 soal yang termasuk dalam kategori mudah (13,3%), kategori sedang (86,7%), dan nihil soal yang termasuk dalam kategori sulit (0%). Berikut adalah tindak lanjut yang dapat dilakukan setelah pemeriksaan tingkat kesulitan soal (Anas Sudijono). 2011: 376-378):

- a. Butir soal yang memiliki tingkat kesukaran dalam kategori baik (tingkat kesukarannya sedang), sebaiknya butir soal tersebut disimpan dalam bank soal agar dapat dikeluarkan lagi pada waktu yang akan datang.
- b. Butir soal yang tergolong dalam soal yang sukar, ada 3 kemungkinan tindak lanjut yaitu: 1) Butir soal tersebut dibuang dan tidak akan digunakan lagi pada tes selanjutnya. 2) Diteliti ulang faktor yang menyebabkan butir item yang bersangkutan sulit dijawab oleh testee. Perbaikan dapat dilakukan dengan menyederhanakan kalimat soal sehingga tidak menimbulkan multitafsir. Selanjutnya, butir soal tersebut dapat dikeluarkan lagi pada tes hasil belajar yang akan datang. 3) Butir soal tetap dipertahankan untuk digunakan lagi pada tes-tes yang sifatnya sangat ketat, dalam arti sebagian besar dari testee tidak akan diluluskan dalam tes seleksi tersebut.
- c. Butir soal yang termasuk kategori mudah, ada 3 kemungkinan Tindakan lanjutan yaitu: 1) Butir soal tersebut dibuang dan tidak akan digunakan Kembali dalam tes selanjutnya. 2) Diteliti ulang faktor yang menyebabkan butir item yang bersangkutan sulit dijawab oleh testee. Perbaikan dapat dilakukan dengan memperbaiki opsi dan membuat kalimat soal menjadi lebih kompleks. Selanjutnya, butir soal tersebut dapat dikeluarkan lagi pada tes hasil belajar yang akan datang. 3) Butir soal tetap dipertahankan untuk digunakan lagi pada tes-tes yang sifatnya longgar. Dalam kondisi seperti ini tes hanyalah formalitas saja.

**e. Efektivitas Pengecoh**

Penggunaan program Microsoft Excel digunakan untuk menghitung secara manual keefektifan distraktor pada soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia di kelas XII IPS SMAN 2 Pandeglang tahun

ajaran 2021–2022. Kriteria berikut digunakan untuk menginterpretasikan distraktor setiap item:

*Sangat baik IP = 76% - 125%*

*Baik IP = 51% - 75% atau 126% - 150%*

*Kurang Baik IP = 26% - 50% atau 151% - 175%*

*Jelek IP = 0% - 25% atau 176% - 200%*

*Sangat Jelek IP = lebih dari 200%*

(Zainal Arifin, 2014: 280)

Kemudian, dengan menggunakan standar yang diadaptasi dari skala Likert, interpretasikan keefektifan setiap item sebagai berikut:

- a. Apabila keempat jawaban pengecoh soal dapat berfungsi dengan baik, maka soal dapat dikatakan memiliki efektivitas pengecoh yang sangat baik.
- b. Apabila terdapat tiga jawaban pengecoh yang berfungsi maka soal dikatakan memiliki efektivitas pengecoh yang baik.
- c. Apabila terdapat dua jawaban pengecoh yang berfungsi maka soal dikatakan memiliki efektivitas pengecoh yang cukup baik.
- d. Apabila hanya terdapat satu jawaban pengecoh yang berfungsi maka soal dikatakan memiliki efektivitas pengecoh yang kurang baik.
- e. Apabila semua jawaban pengecoh tidak berfungsi maka soal dikatakan memiliki efektivitas pengecoh yang tidak baik.

Analisis soal Penilaian Akhir Tahun (PAT) mata kuliah Sejarah Indonesia kelas XII IPS SMAN 2 Pandeglang tahun pelajaran 2021–2022 diketahui bahwa tiga puluh soal (100%) memenuhi standar yang dipersyaratkan. Karena distraktor memenuhi kriteria baik yaitu 100 persen, maka nilai keefektifan distraktor berada pada kategori baik. Berikut ini dapat diikuti (Anas Sudijono, 2011: 417).

- a. Pengecoh yang berfungsi dengan baik dapat digunakan lagi pada tes hasil belajar yang akan datang.
- b. Pengecoh yang belum berfungsi dengan baik, diperbaiki atau diganti dengan pengecoh yang lain. Cara untuk membuat pengecoh yang baik yaitu sebagai berikut (Sumarna Surapranata. 2005: 136).
  1. Gunakan pilihan jawaban yang dapat dimengerti peserta didik.
  2. Gunakan kata-kata yang kedengarannya sama.
  3. Gunakan yang kira-kira ada kaitannya.
  4. Gunakan bahasa buku yang tidak diragukan lagi kebenarannya.

## **KESIMPULAN**

Analisis butir Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun ajaran 2021/2022 menghasilkan kesimpulan sebagai berikut: Validitas, reliabilitas, daya pembeda, tingkat kesukaran, dan efektivitas pengecoh.

1. Berdasarkan temuan penelitian, soal pilihan ganda Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun pelajaran 2021/2022 terdapat 13 soal yang valid (43,3%) dan 17 soal yang tidak valid (56,7%).
2. Berdasarkan temuan penelitian reliabilitas menunjukkan bahwa soal pilihan ganda Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun pelajaran 2021/2022 memiliki reliabilitas sebesar 0,49, hal ini menunjukkan bahwa baik pertanyaan maupun hasilnya tidak konsisten.
3. Berdasarkan temuan penelitian, Daya Diskriminasi mengungkapkan bahwa soal Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS di SMAN 2 Pandeglang tahun pelajaran 2021/2022 berisi soal pilihan ganda. Dua puluh pertanyaan (atau 66,7%) termasuk dalam kategori lemah, tujuh pertanyaan (atau 23,3%) termasuk dalam kategori cukup, satu item (atau 3,3%) termasuk dalam kategori baik, dua item (atau 6,7%) termasuk dalam kategori baik.
4. Berdasarkan temuan penelitian, soal Penilaian Akhir Tahun (PAT) mata pelajaran Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun pelajaran 2021/2022 terdapat empat soal pilihan ganda yang termasuk dalam kategori mudah, atau 13,3%, dan empat soal yang masuk dalam kategori sedang.
5. Berdasarkan penelitian, soal Penilaian Akhir Tahun (PAT) Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun pelajaran 2021/2022 memiliki 30 soal pilihan ganda (100%) dengan kriteria baik, hal ini menunjukkan keefektifan pembelajaran. Karena efektivitas pengecoh memenuhi kriteria baik yaitu 100%, maka nilai keefektifan pengecoh berada pada kategori baik.
6. Diketahui berdasarkan analisis keseluruhan soal, soal Penilaian Akhir Tahun (PAT) pilihan ganda Sejarah Indonesia kelas XI IPS SMAN 2 Pandeglang tahun pelajaran 2021/22 berjumlah 5 soal (16,7%). dengan kategori kualitas, 23 soal (76,7%) dengan kategori kualitas rendah, dan 2 soal (6,6%) dengan kategori kualitas rendah.
7. Terima kasih kepada Kepala Sekolah SMAN 2 Pandeglang Dra. Hj. Lilis Lismunah, M.Pd yang telah memberi izin sebagai tempat penelitian dalam tulisan ini. Serta terima kasih kepada Desi Tresnasari, S.Pd. sebagai guru mata pelajaran Sejarah Indonesia yang telah banyak membantu dalam proses penelitian.

#### **DAFTAR PUSTAKA**

- Aiken, L.R. (1995). *Psychological Testing and Assessment*. Boston: Alliy and Bacon.
- Anastasi, Urbina, S. (1997). *Psychological Testing*. New Jersey: Prentice Hall, Inc.
- Arifin, Zainal. (2012). *Evaluasi Pembelajaran (Edisi Revisi)*. Jakarta. Direktorat Jendral Pendidikan Islam Kementerian Agama Islam.
- Arifin, Zainal. (2014). *Evaluasi Pembelajaran*. Bandung: PT Remaja Rosdakarya.
- Arikunto, S. (2013). *Dasar-dasar Evaluasi Pendidikan*. Jakarta: Bumi Aksara.

- Arikunto, S. (2014). *Prosedur Penelitian: Suatu Pendekatan Praktik*. Jakarta.: PT Rineka Cipta.
- Azwar, S. (2013). *Sikap Manusia: Teori dan Pengukurannya*. Yogyakarta: Pustaka Pelajar.
- Brinkerhoff, R. O. et al. (1986). *Program Evaluation: A Practitioner's Guide for Trainers and Educators. Fourth Printing*. Massachusetts: Kluwer- Nijhoff Publishing.
- Creswell, J.W. (2016). *Research Design: Pendekatan Metode Kualitatif, Kuantitatif, dan Campuran (Terjemahan Edisi 4)*. Yogyakarta: Pustaka Pelajar.
- Darmawan, D. (2013). *Metode Penelitian Kuantitatif*. Bandung: Remaja Rosdakarya.
- Depdiknas. (2007). *Panduan Penulisan Soal Pilihan Ganda*. Jakarta: Pusat Penilaian Pendidikan Balitbang Depdiknas.
- Djamarah, S. (2005). *Guru dan Anak didik dalam interaksi edukatif suatu pendekatan teoritis psikologis*. Jakarta: PT. Rineka Cipta.
- Djiwandono, S. (2008). *Tes Bahasa dalam Pengajaran*. Bandung: ITB.
- Gronlund, E. Norman.(1982). *Constructing Achievement Test*. London: Prentice Hall.
- Kusaeri dan Suprananto. (2012). *Pengukuran dan Penilaian Pendidikan*. Yogyakarta: Graha Ilmu.
- Majid, A. (2014). *Penilaian Autentik Proses dan Hasil Belajar*. Bandung: PT. Remaja Rosdakarya Offset.
- Mardapi, D. (2008). *Teknik Penyusunan Instrumen Tes dan Nontes*. Yogyakarta: Mitra Cendekia.
- Muhson, A. (2015). *Panduan Penggunaan Anbuso Versi 6.1*. Yogyakarta: Universitas Negeri Yogyakarta.
- Mulyasa, E. (2005). *Kurikulum Tingkat Satuan Pendidikan*. Bandung: PT. Remaja Rosdakarya.
- Nurgiyantoro, B. (2011). *Penilaian Pembelajaran Bahasa*. Yogyakarta: BPFE.
- Popham, W. J. (1995). *Classroom Assesment: What Teacher Need to Know*. Boston: Allyn and Bacon.
- Pramana, Y.A. (2013). *Aplikasi Microsoft Excel 2010 untuk Menganalisis Butir Soal Pilihan Ganda*. Skripsi. Semarang: UNS
- Purwanto, Ngalm. (2013). *Prinsip-Prinsip dan Teknik Evaluasi Pengajaran*. Bandung: PT. Remaja Rosdakarya.
- Purwanto. (2011). *Evaluasi Hasil Belajar*. Yogyakarta: Pustaka Pelajar.
- Ratnawulan, E & Rusdiana. (2014). *Evaluasi Pembelajaran*. Bandung: Pustaka Pelajar.
- Slameto. (2003). *Belajar dan Faktor-faktor yang Mempengaruhi Edisi Revisi*. Jakarta: Rineka Cipta.
- Subali, B. (2014). *Evaluasi Pembelajaran (Proses dan Produk)*. Purwokerto. Universitas Muhammadiyah Purwokerto.
- Sudijono, A. (2009). *Pengantar Evaluasi Pendidikan*. Yogyakarta: PT Raja
- Tjalla, A.(2020) *Potret Mutu Pendidikan Indonesia Ditinjau dari Hasil-hasil Studi Internasional*. Jurnal. Jakarta: Universitas Negeri Jakarta.
- Toksosz, S & Ertunc, A. (2017). *Item Analysis of a Multiple Choice Exam*. Jurnal. Turkey. University, İstiklal Yerleşkesi Turkey.

- Umamah, N. (2014). *Bahan Ajar: Perencanaan Pembelajaran Bidang Studi*. Jember: Universitas Jember.
- Utami, I. (2016). *Analisis Butir Soal Pilihan Ganda Ulangan Akhir Semester Genap Tahun Pelajaran 2014/2015 Mata Pelajaran PKn Kelas IV SD di Kecamatan Depok, Sleman, Yogyakarta*. Skripsi. Yogyakarta: Universitas Sanata Yogyakarta.
- Wahyudi, D. (2011). *Analisis Kualitas Butir Soal Mata Pelajaran Pendidikan Agama Islam dalam Pencapaian Kompetensi Siswa SMA Negeri 2 Kebumen*. Skripsi. Yogyakarta: Universitas Islam Negeri Sunan Kalijaga Yogyakarta.
- Werdiningsih, G. (2015). *Analisis Kualitas Butir Soal Ulangan Akhir Semester Gasal Mata Pelajaran Ekonomi Kelas XII IPS SMAN 2 Banguntapan Tahun Ajaran 2014/2015*. Skripsi. Yogyakarta: Universitas Negeri Yogyakarta.
- Wijayanti, H. (2014). *Analisis Butir Soal Objektif UAS Semester Genap Kelas VII Mata Pelajaran Ilmu Pengetahuan Sosial (IPS) Tahun Pelajaran 2013/2014 Di SMP 3 Balung*. Skripsi. Jember. Universitas Jember Press.
- Yulista H., Zulfan, & M. Arifin (2018). *Analisis Tingkat Kesukaran Soal dan Daya Pembeda Soal Mata Pelajaran Sejarah Kelas Xi Semester Ganjil di SMA Negeri 5 Banda Aceh Tahun Pelajaran 2015-2016*. Jurnal. Banda Aceh: Universitas Syiah Kuala.
- Yuniarti, A.D. (2013). *Analisis Butir Soal Olimpiade Ekonomi VI pada Prodi Pendidikan Ekonomi FKIP UNEJ Tingkat SMA Sederajat*. Skripsi. Jember. Universitas Jember Press.