

Analisis kualitas soal Penilaian Sumatif Akhir Jenjang Pendidikan Pancasila di SMA Negeri 2 Sukoharjo

Desinta Ayu Liana ^{a,1}, Suyato ^{b,2}

¹ (desintaayu.2021@student.uny.ac.id) Mahasiswa Departemen PKn, Fisip, UNY

² (suyato@uny.ac.id) Dosen Fakultas Ilmu Sosial dan Ilmu Politik, UNY

^a Mahasiswa (Departemen Pendidikan Kewarganegaraan FISIP UNY), Yogyakarta Indonesia

^b Dosen (Departemen Pendidikan Kewarganegaraan FISIP UNY), Yogyakarta Indonesia

ABSTRAK

Penelitian ini bertujuan untuk mendeskripsikan kualitas butir soal yang ditelaah secara kualitatif, kuantitatif dan jumlah butir soal yang berkualitas pada Penilaian Sumatif Akhir Jenjang mata pelajaran Pendidikan Pancasila di SMA Negeri 2 Sukoharjo. Penelitian ini merupakan suatu penelitian eksploratif dengan pendekatan kuantitatif deskriptif. Populasi yang menjadi fokus dalam penelitian ini adalah peserta didik kelas XII di SMA Negeri 2 Sukoharjo. Data penelitian diperoleh melalui lembar telaah dan hasil analisis. Pada teknik analisis data peneliti menggunakan uji instrumen (uji validitas dan reliabilitas) serta uji kualitas soal, yang terdiri dari uji analisis kualitatif dan uji analisis kuantitatif (tingkat kesulitan, daya beda dan *distractor*). Hasil penelitian ini menunjukkan bahwa Kualitas butir soal PSAJ yang ditelaah secara teoritis terdapat 22 butir soal yang dapat diterima dan dinyatakan baik oleh ketiga ahli dari 30 butir soal pilihan ganda, Kualitas butir soal PSAJ yang ditelaah secara empiris terdapat 9 butir soal yang memiliki tingkat kesulitan, daya beda, dan pengecoh yang baik, Jumlah butir soal PSAJ yang berkualitas sebanyak 7 butir menurut hasil analisis validitas *Aiken's*, telaah secara teoritis, dan telaah secara empiris.

ABSTRACT

This study aims to Describe the quality of the items studied qualitatively, quantitatively and the number of quality items in the Final Sumative Assessment (PSAJ) of Pancasila Education at SMA Negeri 2 Sukoharjo. This study is explorative research with descriptive qualitative approaches. The population that became the focus of this study were students of class XII at SMA Negeri 2 Sukoharjo. The research data were obtained through a review sheet and the results of the analysis. In the data analysis technique, the researcher used instrument tests (validity and reliability tests) and question quality tests, which consisted of qualitative analysis tests and quantitative analysis tests (difficulty, difference and distractor levels). The results of this study indicate that The quality of the PSAJ question items studied theoretically, there are 22 questions that can be accepted and stated well by the three experts out of 30 multiple choice questions, The quality of the PSAJ question items studied empirically, there are 9 questions that have a good level of difficulty, different power, and deceiver, The number of quality PSAJ question items is 7 according to the results of Aiken's validity analysis, theoretical study, and empirical study

Sejarah Artikel

Diterima : 15-07-2025

Disetujui : 19-07-2025

Kata kunci:

Kualitas Soal, Pendidikan Pancasila

Keywords:

Question Quality, Pancasila education

Pendahuluan

Pendidikan memiliki kontribusi yang cukup penting guna memperbaiki kualitas atau mutu pada individu maupun kelompok. Tujuan dari Negara Kesatuan Republik Indonesia salah satunya telah disebutkan dalam pembukaan Undang-undang Dasar 1945 yaitu mencerdaskan kehidupan bangsa. Definisi dari Pendidikan dalam Undang-undang Nomor 20 Tahun 2003 adalah usaha secara sadar serta sudah tersusun guna merealisasikan kondisi belajar serta proses pembelajaran supaya peserta didik dengan cakap dan bersungguh-sungguh memperdalam kemampuan pada diri mereka guna memiliki kecakapan spiritual dalam agama, pengelolaan dalam diri, karakter, intelektual, akhlak yang

baik, serta keterampilan yang dibutuhkan pada dirinya, lingkungan di masyarakat, bangsa serta negara.

Proses pembelajaran memiliki rangkaian program dengan beberapa komponen yang diantaranya yaitu tujuan pembelajaran, adanya kegiatan pembelajaran atau proses belajar mengajar, dan evaluasi terhadap pembelajaran. Proses atau kegiatan evaluasi pembelajaran juga termasuk tahap yang penting dalam kegiatan pembelajaran hal tersebut dikarenakan dari adanya kegiatan evaluasi maka dapat memberikan informasi serta data terkait tingkatan capaian dari adanya tujuan pembelajaran oleh peserta didik, maka diharapkan dari adanya informasi tersebut guru bisa memberikan dan melakukan upaya-upaya sebagai kegiatan tindak lanjutnya.

Proses evaluasi setidaknya juga mencakup tiga hal yang dijadikan sebagai sasaran dalam melakukan pengukuran kompetensi pada hasil belajar, tiga hal tersebut antara lain yaitu terkait aspek pengetahuan (kognitif), aspek sikap (afektif), serta aspek keterampilan (psikomotorik) (Zainudin & Ubabuddin, 2023). Kegiatan evaluasi memiliki fungsi sebagai solusi guna meningkatkan kualitas dalam proses pembelajaran, serta adanya evaluasi juga dapat menjadi solusi motivasi pendidik guna meningkatkan sistem kerjanya terkait elemen program-program pembelajaran yang belum kuat serta perlu adanya perbaikan, dan evaluasi dapat bermanfaat pula untuk meningkatkan semangat belajar pada peserta didik (Magdalena et al., 2020).

Dalam kaitannya sebagai alat guna melakukan evaluasi hasil belajar, setidaknya tes memiliki dua fungsi, yaitu: pertama, untuk mengukur sejauh mana penguasaan terhadap materi tertentu atau tingkat pencapaian terhadap tujuan-tujuan tertentu, fungsi kedua, guna menetapkan posisi atau tingkatan peserta didik dalam kelompoknya, terkait kemampuan dalam menguasai materi atau capaian dalam tujuan pembelajaran. Fungsi evaluasi yang pertama lebih difokuskan pada pengukuran berhasil atau tidaknya rancangan dalam pembelajaran, sementara itu, fungsi evaluasi yang kedua lebih menitikberatkan pada penilaian berhasil atau tidaknya proses belajar pada perorangan peserta tes (Sunaryati et al., 2024). Selain itu menurut Sudjiono (Zainuri et al., 2021) terdapat tiga jenis fungsi tes, diantaranya adalah: 1) Tes dipergunakan untuk menilai atau mengukur peserta didik, 2) Tes digunakan sebagai alat untuk mengukur dan menilai berhasil atau tidaknya suatu rancangan pembelajaran atau kurikulum, 3) Tes digunakan untuk tinjauan guna menyempurnakan kegiatan pembelajaran yang belum mencapai tujuan.

Pengukuran, penilaian dan evaluasi adalah tiga hal yang berbeda. Evaluasi memiliki lingkup yang lebih luas dari penilaian, karena penilaian fokusnya lebih kepada aspek tertentu saja (Rahman & Narsyah, 2019). Menurut Arifin (2014) Penilaian merupakan suatu kegiatan secara tersusun dan berkelanjutan guna mendapatkan data atau informasi terkait prosedur dan hasil belajar dari peserta didik sebagai cara dalam menentukan ketetapan-ketetapan yang telah didasarkan pada kompetensi serta pertimbangan yang telah ditentukan. Menurut Sudijono (2011) Penilaian memiliki arti yaitu memberikan nilai pada sesuatu, dan memberi nilai memiliki arti penarikan keputusan pada objek dengan adanya dasar atau patokan untuk sehat atau sakit, baik atau buruk dan sebagainya. Mardapi (2017) berpendapat bahwa Penilaian meliputi seluruh cara yang dipakai untuk memperoleh data individu. Proses penilaian dalam pendidikan memiliki ciri-ciri yang diantaranya yaitu dilaksanakan dengan tidak langsung, adanya suatu ukuran secara kuantitatif, dalam penilaian pendidikan memakai satuan yang konstan atau tetap termasuk dalam pengukuran terkait IQ, penilaian pendidikan juga memiliki sifat yang relatif, dalam penilaian pendidikan juga tidak jauh dari adanya kesalahan-kesalahan (Arikunto, 2018).

Sedangkan pada kegiatan pengukuran didapatkan dengan alat ukur atau instrumen. Menurut Azwar (2015) Pengukuran adalah suatu cara memberikan angka pada variabel dengan suatu urutan. Pada dasarnya kegiatan pengukuran yaitu membandingkan sesuatu dengan atau atas dasar ukuran

yang ditentukan. Secara kesimpulan, pengukuran yaitu suatu kegiatan dalam memberikan angka atau deskripsi angka pada suatu individu. Alat ukur tersebut bisa berbentuk tes atau non tes, dalam bentuk tes maka dapat berupa soal dengan pilihan ganda ataupun uraian (kognitif), sedangkan non tes dapat berupa pedoman untuk observasi, wawancara, penilaian untuk bakat, ataupun angket (afektif) (Rahman & Narsyah, 2019).

Secara etimologis, istilah 'tes' berasal dari kata dalam bahasa Prancis Kuno yakni "*testum*" dengan arti piring untuk menyisihkan logam-logam mulia (Riinawati, 2021). Tes merupakan sekumpulan pertanyaan atau latihan ataupun alat lain yang dapat digunakan untuk mengukur seseorang atau kelompok orang pada aspek pengetahuan, keterampilan, kemampuan, kecerdasan ataupun bakat (Arikunto, 2013). Tes merupakan alat ukur dengan bersifat baku dan objektif, maka mampu dipergunakan secara luas guna menilai serta membandingkan kondisi psikologis atau perilaku individu (Setyawarno, 2017). Oleh karena itu, tes dapat memberikan keterangan yang akurat dan objektif mengenai objek yang diukur, baik dalam aspek psikologis maupun perilaku, serta memungkinkan perbandingan antarindividu. Maka dari itu, dapat ditarik kesimpulan bahwa tes merupakan metode atau alat penilaian berupa satu pekerjaan atau serangkaian tanggung jawab yang wajib untuk diselesaikan peserta didik, yang bertujuan guna memperoleh nilai prestasi belajar dan tingkah laku mereka (Bhakti et al., 2022).

Tes dapat dibedakan menjadi berbagai macam kelompok, jika dibedakan dari segi konstruksinya maka dapat dikelompokkan seperti berikut: 1) Berdasarkan bentuk : pada umumnya bentuk tes ada dua macam yaitu bentuk soal uraian (*essay test*) dan bentuk soal objektif (*objective test*). 2) Menurut jenisnya: butir soal *essay* dikelompokkan dalam dua jenis, yaitu soal uraian terbatas (*restricted essay*), dan soal uraian bebas (*extended essay*). Sedangkan untuk butir soal objektif berdasarkan jenisnya dikelompokkan dalam tiga jenis, diantaranya adalah soal benar dan salah (*true-false*), butir soal menjodohkan (*matching*), dan butir soal pilihan ganda (*multiple choice*) (Inanna et al., 2021)

Menurut Arikunto (2018), sebagai alat pengukur, seperangkat tes yang baik harus mempunyai persyaratan tes. Menurut Reynolds, Livingston, & Willson, syarat yang setidaknya perlu dipenuhi sebagai karakteristik instrumen yang baik yakni valid dan reliabel (Faridah, 2021). Sebuah instrumen dapat dikatakan valid apabila instrumen yang dipakai benar-benar mampu mengukur sesuatu yang ingin diukur secara akurat, serta dikatakan reliabel jika memiliki hasil tes yang konsisten atau tetap. Sudijono (2011) mengatakan bahwa terdapat dua cara yang dapat dipakai guna menguji validasi, yaitu penganalisisan yang dilakukan dengan memakai logika (*logical analysis*) dan penganalisisan dengan berdasarkan kenyataan empiris (*empirical analysis*). Dalam berbagai buku terkait penelitian kuantitatif dalam Budiastuti & Bandur (2018), terdapat beberapa jenis validitas yang diantaranya yaitu: Validitas Kriteria (*Criterion Validity*) , Validitas Konten (*Content Validity*), Validitas Konstruk (*Construct Validity*). Reliabilitas dalam instrumen bisa diujikan dengan berbagai macam uji reliabilitas. Uji tersebut diantaranya yaitu *test retest*, *equivalen*, dan *internal consistency*. Dalam *internal consistency* masih terdapat uji reliabilitas yang berbeda yang diantaranya terdiri dari uji *splitt half*, KR20, KR21, serta *Alpha Cronbach* (Yusup, 2018).

Kualitas pada soal dapat menentukan kualitas tes secara menyeluruh, maka diperlukannya uji coba secara langsung guna mengetahui butir soal yang diubah, direvisi atau diperbaiki, bahkan soal yang tidak dipakai serta item soal yang masih bisa digunakan berikutnya (Khaerudin, 2015). Melakukan analisis butir soal adalah suatu hal atau aktivitas yang perlu dilaksanakan oleh tenaga pendidik untuk meningkatkan kualitas soal yang telah dibuat. Seorang pendidik juga mempunyai tugas untuk melaksanakan kegiatan penilaian dan evaluasi terhadap peserta didik, tidak hanya mempunyai peran dalam mengajar, mendidik serta membimbing peserta didik. Ketika tahap pelaksanaan evaluasi, seorang guru wajib mempunyai kemahiran dalam menyusun instrumen tes, tujuannya agar guru bisa membuat instrumen tes dengan memiliki kualitas yang baik, mutu yang baik,

serta memiliki kelayakan guna diujikan pada siswa atau peserta didik sehingga tes yang telah dibuat mampu menguji kecakapan peserta didik. Apabila instrumen tes memiliki kualitas dan mutu yang baik maka guru mampu memperoleh deskripsi yang sah serta jelas yang berkaitan dengan hasil belajar siswanya (Juliani & Erita, 2023).

Analisis pada soal dilaksanakan guna mengetahui apakah soal tersebut berfungsi atau tidak. Dari adanya analisis kualitas soal tersebut maka mampu untuk mengetahui bagaimana keadaan atau kondisi soal yang digunakan tersebut (Rudini, 2020). Pendapat Surharsimi Arikunto analisis soal merupakan suatu cara yang tersusun dalam memberi gambaran informasi kekhususan pada butir-butir tes yang sudah kita buat, sedangkan Nana Sudjana berpendapat bahwa Analisis butir soal atau biasa disebut analisis item merupakan pendalaman apa yang ditanyakan dalam tes agar memperoleh instrumen pertanyaan yang mempunyai mutu baik dan memadai (Nurjanah & Marlianingsih, 2015). Tujuan dari adanya analisis butir soal yakni agar dapat mengetahui seberapa tinggi tingkat kesukaran pada item soal, daya beda serta keberfungsian distraktor atau pengecoh pada soal. Selain itu analisis item juga bisa menjadi petunjuk sehingga mutu pada butir soal yang telah disusun oleh guru dapat meningkat menjadi lebih baik (Sutrisno, 2014).

Khaerudin (2017) berpendapat bahwa, Analisis kualitatif atau sering dikatakan validitas logis (*logical validity*) merupakan pengkajian yang memiliki tujuan guna menganalisis soal dengan tinjauan segi teknis, isi, dan editorial. Selain itu analisis kualitatif juga bisa dikelompokkan berdasarkan segi materi, konstruksi, dan bahasa. Pada analisis materi memiliki tujuan guna menelaah hal yang memiliki keterkaitan dengan pokok ilmu yang menjadi pertanyaan butir soal dan tingkat keahlian yang setara dengan butir soal, pada analisis konstruksi menelaah hal yang pada umumnya memiliki keterkaitan dengan cara penulisan soal, dan analisis bahasa memiliki tujuan guna menelaah butir soal yang memiliki keterkaitan dengan pemakaian bahasa Indonesia yang baik dan benar menurut Ejaan Yang Disempurnakan (EYD), sehingga dari penganalisisan ini sebuah soal dapat diketahui keberfungsian (Kurniawan & Mutaqimah, 2009).

Validitas logis atau analisis kualitatif dilaksanakan ketika soal belum digunakan, hal tersebut guna mengetahui keberfungsian soal yang telah dibuat. Sedangkan, analisis soal empiris atau kuantitatif dilaksanakan ketika soal sudah diujicobakan terhadap sampel yang representatif guna mengetahui lebih berfungsi atau tidaknya soal tersebut dalam mencari informasi tentang tingkat kesulitan, daya pembeda serta pengecoh (Khaerudin, 2017), sehingga soal yang telah digunakan bisa dilakukan analisis pada tiap butir soalnya.

Untuk melakukan analisis tes terdapat dua cara atau dua teori yang dapat digunakan diantaranya yaitu teori tes klasik dan teori tes modern (Suryabrata dalam Iskandar & Rizal, 2018) Analisis dapat dilakukan dengan menggunakan salah satu dari teori tersebut. Menurut Mardapi (2017) penerapan model teori tes klasik yaitu dengan bentuk matematika sederhana guna menentukan korelasi terkait skor pengamatan, skor asli, serta skor kesalahan. Hal ini terdapat asumsi utama yaitu tidak ada hubungan antara nilai asli atau nilai sebenarnya dengan nilai kesalahan, sehingga model skor pengamatan adalah jumlah model skor asli dan model skor kesalahan. Tolak ukur pada teori ini adalah kemampuan serta parameter butir yang terdiri dari tingkat daya beda pada soal, tingkat kesulitan pada soal serta dugaan.

Menurut pendapat Allen & Yen (1979, p. 56) model tes atau teori tes merupakan penggambaran simbol dari faktor yang memiliki pengaruh terhadap nilai tes yang dilihat dan dijabarkan dengan dugaan. Sehingga dapat dikatakan bahwa teori tes klasik adalah sebuah pendekatan guna menganalisis perbedaan antar individu yang lebih sederhana dari anggapan teori respon butir serta lebih kecil untuk takaran sampel yang berlaku. Menurut Stage dalam Sumaryanta (2021) Pada teori tes klasik nilai tes atau skor tes dalam pandangannya terdapat dua komponen yaitu

nilai sebenarnya dan kesalahan dalam pengukuran. Parameter utama yang dianalisis dalam teori ini meliputi tingkat kesulitan butir soal, daya pembeda butir, serta reliabilitas instrumen (Summaryanta, 2021). Selain ketiga parameter tersebut, analisis menggunakan teori tes klasik juga dimanfaatkan untuk mengevaluasi efektivitas pengecoh pada soal bentuk pilihan ganda. Informasi dari parameter-parameter tersebut memiliki peranan penting dalam memahami mutu tes yang sedang dikembangkan.

Pentingnya cara penilaian hasil belajar salah satunya dengan tes guna menguji tingkatan peserta didik pada prestasi hasil belajarnya. Tes bisa dilaksanakan secara berkala pada saat ulangan tengah semester, ulangan akhir semester ataupun ulangan harian. Evaluasi dapat dilaksanakan melalui hasil ulangan akhir semester, nilai tersebut dapat memberikan gambaran terkait penguasaan capaian pembelajaran selama satu semester oleh peserta didik sehingga perlu adanya soal yang cakupannya lebih luas untuk mengukur hal tersebut. Setelah keseluruhan soal telah disusun, maka lebih baiknya soal tersebut didiskusikan dan dilakukan telaah soal terkait dengan aturan penulisan soal. Hal-hal yang harus diperhatikan ketika melakukan telaah soal yakni mulai dari segi bahasa, isi, konstruksi serta materi. Kegiatan ini dapat dikatakan sebagai analisis validitas soal. Langkah ini perlu dilakukan untuk mengujikan soal yang telah dibuat guna memperbaiki soal tersebut apabila terdapat kekurangan.

Namun, kurangnya rasa sadar tenaga pendidik terhadap pentingnya melaksanakan analisis pada kualitas butir soal yang mengakibatkan guru hanya melakukan analisis sederhana dan subjektif, yakni hanya menganalisis pada validitas logis, yang dimana penilaian didasarkan pada kemampuan menjawab peserta didik. Guna memperoleh informasi terkait suatu soal yang mempunyai kualitas atau tidak mempunyai kualitas, kurang cukup apabila hanya menganalisis validitas logis saja. Pengamatan lapangan menunjukkan bahwa sebagian besar guru saat menyusun atau membuat soal sering memiliki tingkat kesulitan yang tinggi dan sangat rendah (Sutrisno, 2014).

Dua permasalahan itu memiliki dampak yang buruk bagi siswa karena soal yang dinilai bagus yaitu soal yang memiliki taraf kesulitan yang tidak terlalu ataupun rendah. Butir soal yang dinilai cukup mudah tidak memotivasi peserta didik dalam berusaha untuk menyelesaikan tes. Agar hasil tes dapat memberikan gambaran keadaan sebenarnya, maka rangkaian butir pada soal yang digunakan wajib memenuhi semua persyaratan yang ada dalam tes dan alat ukur yang baik serta dapat diuji dari berbagai aspek. Soal-soal yang dipakai oleh pendidik harusnya memiliki mutu yang bagus dari sudut pandang manapun, antara lain validitas, reliabilitas, tingkat kesulitan, daya beda, dan efektivitas pengecoh (*distractor*) (Mardapi dalam Charismana, 2016).

Pada penelitian terdahulu dengan judul *"Analisis Kualitas Soal Buatan Guru Pendidikan Pancasila dan Kewarganegaraan (PPKn)"*. Peneliti melakukan wawancara terhadap guru PPKn di SMAN 6 Banjarmasin dengan hasil apabila guru selama ini kurang dalam melakukan analisis terhadap butir-butir soal yang diujikan pada peserta didik dalam mata pelajaran PPKn baik terhadap soal penilaian harian maupun dalam penilaian akhir semester. Hasil dari analisis terhadap butir soal mulai dari tingkat kesulitan, daya beda serta keberfungsian pengecoh atau *distractor* mata pelajaran PPKn kelas XI tahun ajaran 2018 / 2019 di SMAN 6 Banjarmasin dengan bantuan program AnBuso versi 8.0 yaitu dengan hasil terdapat 10% atau 5 soal dengan kualitas yang baik, 14% atau 7 soal harus dilakukan perbaikan pada pilihan jawaban, 20% atau 10 soal dengan kategori cukup baik, dan 56% atau 28 soal dengan kategori tidak baik (Adawiah & Ruchliyadi, 2019).

Penelitian terdahulu yang berjudul *"Analisis Butir Soal Ulangan Akhir Semester PPKn Kelas X SMA N 2 Mempawah Tahun Ajaran 2018/2019"*. Peneliti menjelaskan bahwa data yang terdapat di lapangan telah memberikan gambaran bahwa selama ini guru tidak melakukan analisis terhadap tiap-tiap butir soal, baik analisis kualitatif atau dari segi bahasa, isi, serta konstruksi maupun analisis kuantitatif yang dapat dilakukan pada tingkat kesulitan soal, daya beda soal serta keberfungsian

distraktor dan pada validitas maupun reliabilitasnya, serta guru juga kurang melakukan persiapan saat menyusun soal yang digunakan untuk tes pada ulangan akhir semester. Hal tersebut dijadikan alasan karena guru menganggap bahwa tugasnya sangat banyak sehingga kurang waktu dalam melakukan penyusunan instrumen tes agar sesuai dengan aturan dalam penulisan tes. (Wahyuni et al., 2019).

Berdasarkan uraian diatas guru selama ini kurang dalam melakukan analisis terhadap butir-butir soal, sehingga belum diketahuinya kualitas soal yang diujikan pada peserta didik. Oleh karena itu peneliti akan menganalisis butir soal guna mengetahui kualitas soal Penilaian Sumatif Akhir Jenjang (PSAJ) mata pelajaran Pendidikan Pancasila Kelas XII Tahun Ajaran 2024/2025 di SMA Negeri 2 Sukoharjo dengan pengimplementasian Kurikulum Merdeka. Analisis ini mencakup aspek kualitatif dan kuantitatif, meliputi validitas, reliabilitas, tingkat kesulitan, daya beda dan efektivitas pengecoh (*distractor*).

Metode

Penelitian ini merupakan suatu penelitian eksploratif dengan pendekatan kuantitatif deskriptif. Menurut Sugiyono (2019) menyatakan bahwa Penelitian deskriptif eksploratif merupakan penelitian yang menggunakan metode guna memberikan suatu gambaran terkait hasil penelitian, namun temuan tersebut tidak dijadikan sebagai suatu kesimpulan yang lebih umum. Prosedur penelitian ini diantaranya yaitu mengumpulkan perangkat tes dan lembar jawaban terhadap perangkat tes, melakukan analisis validasi terhadap butir tes, melakukan analisis butir soal secara kualitatif dan kuantitatif, dan menentukan butir soal yang berkualitas. Teknik analisis data yang digunakan dalam penelitian ini yaitu teori tes klasik, analisis ini melalui dua tahapan yaitu telaah butir soal oleh tim ahli dan analisis dengan bantuan program komputer *Iteman*. Pengumpulan data dalam penelitian ini menggunakan teknik dokumentasi.

Hasil dan Pembahasan

Validitas yang digunakan dalam mengukur perangkat tes PSAJ Pendidikan Pancasila Kelas XII di SMA Negeri 2 Sukoharjo yaitu validitas isi. Jenis validitas isi yang lebih sering digunakan adalah validitas isi formula *Aiken's*. Dasar dari validitas *Aiken's* adalah hasil penilaian dari panel ahli terhadap suatu item sejauh mana item mewakili konstruk yang diukur. Penelitian ini mengukur validitas isi item dengan menggunakan tiga panel penilai, diantaranya yaitu satu orang dosen pengampu mata kuliah Pendidikan Kewarganegaraan di Fakultas Ilmu Sosial, Hukum, dan Ilmu Politik Universitas Negeri Yogyakarta, dan dua orang guru mata pelajaran Pendidikan Pancasila di SMK Negeri 4 Surakarta. Analisis validitas isi *Aiken's* memiliki tujuan guna mengukur seberapa ketepatan butir-butir pada perangkat tes PSAJ Pendidikan Pancasila dalam melakukan fungsi ukurnya. Berikut ini hasil analisis validitas isi *Aiken's* item pada perangkat tes PSAJ Pendidikan Pancasila Kelas XII di SMA Negeri 2 Sukoharjo.

Tabel 1

Hasil Analisis Validitas Isi Aiken's Item Tes PSAJ Pendidikan Pancasila.

No. Butir	s1	s2	s3	Σs	V
Soal 1	3	3	2	8	0,88
Soal 2	2	3	2	7	0,77
Soal 3	2	2	2	6	0,66
Soal 4	3	2	3	8	0,88
Soal 5	2	3	3	8	0,88
Soal 6	1	2	1	4	0,44
Soal 7	3	3	2	8	0,88
Soal 8	3	3	2	8	0,88
Soal 9	2	2	2	6	0,66
Soal 10	2	2	1	5	0,55
Soal 11	3	3	2	8	0,88
Soal 12	1	3	2	6	0,66
Soal 13	3	3	2	8	0,88
Soal 14	2	3	3	8	0,88
Soal 15	3	3	2	8	0,88
Soal 16	2	3	2	7	0,77
Soal 17	2	1	2	5	0,55
Soal 18	3	3	2	8	0,88
Soal 19	3	3	2	8	0,88
Soal 20	3	3	2	8	0,88
Soal 21	2	3	3	8	0,88
Soal 22	2	3	3	8	0,88
Soal 23	2	2	2	6	0,66
Soal 24	2	3	3	8	0,88
Soal 25	3	3	2	8	0,88
Soal 26	1	1	2	4	0,44
Soal 27	3	3	1	7	0,77
Soal 28	3	2	1	6	0,66
Soal 29	1	1	1	3	0,33
Soal 30	2	1	3	6	0,66

Sumber: Data yang sudah diolah, 2025

Berdasarkan data pada tabel diatas maka keputusan pada butir soal yaitu:

- Nilai validitasnya $> 0,7$ maka soal dinyatakan valid
- Nilai validitasnya $0,5 - 0,7$ maka soal dinyatakan cukup valid
- Nilai validitasnya $< 0,5$ maka soal dinyatakan tidak valid (Charismana, 2016).

Hasil analisis pada item tes PSAJ Pendidikan Pancasila memberikan hasil bahwa terdapat 19 butir soal yang valid (tinggi) yaitu pada nomor soal 1, 2, 4, 5, 7, 8, 11, 13, 14, 15, 16, 18, 19, 20, 21, 22, 24, 25, 27, selanjutnya terdapat 8 butir soal yang cukup valid (sedang) yaitu pada nomor soal 3, 9, 10, 12, 17, 23, 28, 29, dan 3 butir soal yang dinyatakan tidak valid (rendah) yaitu nomor 6, 26, 29.

Setelah seluruh data respon peserta didik telah terkumpul dalam satu file, maka data telah siap diolah pada analisis empiris, akan tetapi sebelum dilakukan analisis empiris terlebih dahulu

dilakukan analisis secara teoritis pada perangkat soal PSAJ Pendidikan Pancasila di SMA Negeri 2 Sukoharjo. Analisis teoritis digunakan untuk melihat kualitas soal berdasarkan tiga aspek yaitu kesesuaian materi, konstruksi dan bahasa oleh tim penelaah (ahli). Tahap analisis ini melibatkan tiga ahli untuk melakukan penelaahan terhadap perangkat soal, diantaranya yaitu satu dosen pengampu mata kuliah Pendidikan Kewarganegaraan di Fakultas Ilmu Sosial, dan Ilmu Politik Universitas Negeri Yogyakarta dan dua guru Pendidikan Pancasila di SMK Negeri 4 Surakarta.

Pemilihan ketiga ahli untuk menelaah perangkat soal tersebut dengan dilakukan berbagai pertimbangan, salah satu pertimbangan tersebut bahwa ketiga penelaah memiliki latar belakang yang kompeten pada bidangnya. Dengan adanya pertimbangan tersebut maka secara logis penelaah dapat melakukan telaah pada tiap butir soal dengan baik. Hasil penelaahan pada perangkat soal dengan penilaian ahli (*Expert judgement*) dalam penelitian ini dapat dilihat dalam tabel 7 tentang rangkuman hasil telaah perangkat soal PSAJ Pendidikan Pancasila Kelas XII di SMA Negeri 2 Sukoharjo, untuk lebih jelas data yang lengkap terkait hasil penelaahan dengan *expert judgement* dapat dilihat pada halaman lampiran penelitian ini.

Tabel 2

Rangkuman Hasil Telaah Perangkat Soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo.

Jenis Soal	Butir Soal Diterima		Butir Soal Ditolak	
	Nomor Butir	Jumlah	Nomor Butir	Jumlah
Pilihan Ganda	1, 7, 8, 11, 13, 14, 15, 17, 18, 19, 20, 21, 22, 23, 24, 25 (Baik) 2, 5, 9, 26, 29, 30 (Cukup)	22	3, 4, 6, 10, 12, 16, 27, 28	8

Sumber: Data yang sudah diolah, 2025

Tabel 2 memberikan gambaran hasil telaah oleh ahli secara teoritis yang menunjukkan bahwa dari 30 soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo dengan jenis soal pilihan ganda yaitu terdapat 22 butir soal yang dapat diterima. Soal yang dapat diterima adalah soal yang dinyatakan baik dan cukup baik oleh ketiga ahli yang terdiri dari dosen dan guru dengan dasar pada aspek materi, konstruksi dan bahasa, sementara itu 8 butir soal yang lain dinyatakan ditolak. Soal yang ditolak adalah soal yang dinyatakan tidak baik karena tidak memenuhi aspek yang ditelaah. Soal tersebut tidak mencukupi standar pada aspek materi dengan tidak mengukur kemampuan berpikir tingkat tinggi dan tidak menggunakan stimulus yang menarik peserta didik, serta tidak memenuhi konstruksi pembuatan soal yang baik. Oleh karena itu, 22 butir soal pilihan ganda telah sesuai dengan kaidah penyusunan butir soal yang baik dari aspek materi, konstruksi dan bahasa.

Data yang diperoleh dalam penelitian ini akan dianalisis menggunakan bantuan program komputer *Iteman 4.3* guna mengetahui informasi terkait ciri-ciri perangkat tes yang mencakup tingkat kesulitan soal (diperlihatkan dengan perbandingan peserta didik yang menjawab benar), daya pembeda (diperlihatkan dengan korelasi poin biserial) dan efektivitas pengecoh/distraktor.

Hasil analisis menggunakan bantuan program *Iteman*, pada tingkat kesulitan soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo digambarkan pada tabel 8 sebagai berikut

Tabel 3
Tingkat Kesulitan Butir Soal Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Presentase
$p > 0,70$ (Mudah)	1, 2, 4, 7, 15, 16, 17, 21, 25, 26, 28, 30	12	40
$0,31 \leq p \leq 0,70$ (Sedang)	5, 6, 8, 9, 10, 11, 12, 13, 20, 22, 23, 27, 29	13	43
$p < 0,30$ (Sulit)	3, 14, 18, 19, 24	5	17
Jumlah		30	100

Sumber: Data yang sudah diolah, 2025

Hasil dari tabel diatas dapat diketahui bahwa terdapat 40% atau 12 butir soal yang termasuk kategori mudah, lalu 43% atau 13 butir soal yang termasuk kategori sedang, serta 17% atau 5 butir soal yang termasuk kategori sulit. Dari data tersebut maka dapat dikatakan bahwa tingkat kesulitan pada butir soal termasuk dalam kategori sedang.

Hasil analisis menggunakan bantuan program Iteman, pada daya beda butir soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo digambarkan pada tabel 4 sebagai berikut

Tabel 4.
Daya Beda Butir Soal Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Presentase
$D < 0,20$ (Jelek)	3, 7, 9, 10, 17, 19, 27, 29, 30	9	30
$0,21 \leq D \leq 0,40$ (Cukup)	1, 2, 4, 6, 8, 11, 12, 13, 15, 18, 20, 22, 23, 24, 25, 28	16	54
$0,41 \leq D \leq 0,70$ (Baik)	5, 16, 21, 26	4	13
$D > 0,70$ (Sangat Baik)	-	-	-
$D = \text{Negatif}$ (Soal dibuang)	14	1	3
Jumlah		30	100

Sumber: Data yang sudah diolah, 2025

Hasil dari tabel tersebut dapat diketahui bahwa terdapat 30% atau 9 butir soal yang mempunyai daya beda jelek, 54% atau 16 butir soal yang mempunyai daya beda cukup baik, 13% atau 4 butir soal dengan daya beda baik, tidak ada butir soal yang mempunyai daya beda sangat baik, dan 3% atau 1 soal yang mempunyai nilai daya beda negatif.

Hasil analisis menggunakan bantuan program Iteman, pada efektivitas pengecoh butir soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo digambarkan pada tabel 5 sebagai berikut.

Tabel 5.
Efektivitas Pengecoh Butir Soal Pilihan Ganda

Kategori	Nomor Butir Soal	Jumlah	Presentase
> 2% dan r_{pbis} negatif (Baik)	2, 3, 4, 5, 6, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 20, 21, 22, 23, 24, 25, 26, 27, 29, 30	26	87
<2% dan r_{pbis} positif selain kunci jawaban (Tidak Baik)	1, 7, 19, 28	4	13
Jumlah		30	100

Sumber: Data yang sudah diolah, 2025

Berdasarkan hasil dari analisis butir soal PSAJ Pendidikan Pancasila yang digambarkan pada tabel diatas dapat dilihat bahwa terdapat 87% atau 26 butir soal yang memiliki pengecoh/distraktor baik, hal tersebut dikarenakan option pengecoh dipilih lebih dari 2% oleh peserta didik dan r_{pbis} negatif. Selain itu terdapat 13% atau 4 butir soal yang memiliki pengecoh/distraktor tidak baik, karena peserta tes yang milih opsi pengecoh tersebut kurang dari 2% atau r_{pbis} bernilai positif selain kunci jawaban. Tiap butir soal pada perangkat tes tidak ada yang mendapatkan peringatan “*chek the key*”, artinya tidak perlu melakukan cek kembali pada kunci jawaban tiap butir soal.

Hasil analisis validitas menggunakan pendekatan *Aiken's* terhadap 30 butir soal pilihan ganda menunjukkan bahwa 27 butir soal mencukupi kriteria validitas dalam kategori cukup baik hingga baik. Selain itu untuk hasil analisis instrumen tes PSAJ mata pelajaran Pendidikan Pancasila di SMA Negeri 2 Sukoharjo melalui pendekatan teoritis (kualitatif) menunjukkan bahwa dari 30 soal yang ditelaah, sebanyak 22 soal tergolong baik berdasarkan aspek materi, konstruksi dan bahasa. Kemudian, butir soal yang telah melalui analisis teoritis dianalisis kembali menggunakan pendekatan empiris (kuantitatif).

Dalam analisis kualitatif seluruh butir soal dengan jenis pilihan ganda dilakukan penganalisisan pada tiap butir soalnya, Meskipun 22 butir soal telah dinyatakan baik berdasarkan analisis teoritis, data tersebut digunakan untuk mengetahui gambaran menyeluruh mengenai butir soal agar dapat menjadi bahan pertimbangan bagi pembuat soal di masa yang akan datang. Dengan menggunakan format yang serupa untuk peserta didik yang sama, sudah diperoleh informasi mengenai tingkat kesulitan dan daya pembeda setiap butir soal.

Berdasarkan analisis menggunakan teori tes klasik dan bantuan program *Iteman*, ditemukan 9 dari 30 butir soal pilihan ganda yang memenuhi kriteria baik. hasil rangkuman akhir proses pemilihan butir soal tes PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo menghasilkan 7 butir soal pilihan ganda yang berkualitas, yaitu butir soal nomor 5, 8, 11, 13, 20, 22 dan 23.

Simpulan

Kualitas butir soal PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo yang ditelaah secara kualitatif (teoritis), menunjukkan hasil bahwa terdapat 22 butir soal yang diterima serta dinyatakan baik oleh ketiga ahli dari 30 butir soal pilihan ganda. Butir soal yang dinyatakan tidak baik serta tidak diterima yaitu sejumlah 8 soal, alasan butir ditolak karena tidak mencukupi standar pada aspek materi dan konstruksi dalam pembuatan soal yang baik. Lalu Kualitas butir soal PSAJ Pendidikan Pancasila

SMA Negeri 2 Sukoharjo yang ditelaah secara kuantitatif (empiris), sebagai berikut: Hasil analisis validitas isi menggunakan *Aiken's*, diperoleh 27 soal yang memiliki validitas tergolong baik dari 30 butir soal keseluruhan. Berdasarkan pendekatan teori tes klasik dengan menggunakan program *Itean* diperoleh hasil yaitu dari 30 butir soal pilihan ganda, terdapat 9 soal (30%) yang mempunyai tingkat kesulitan, daya beda, dan pengecoh yang baik. Terdapat 21 butir soal (70%) tidak mencukupi standar karena termasuk dalam butir soal yang memiliki tingkat kesulitan terlalu mudah atau terlalu sulit, daya beda yang tidak baik dan pengecoh yang tidak berfungsi, sehingga tidak dapat membedakan kemampuan peserta didik yang sebenarnya. Untuk hasil dari analisis secara keseluruhan, baik analisis teoritis maupun empiris pada perangkat tes PSAJ Pendidikan Pancasila SMA Negeri 2 Sukoharjo diperoleh hasil yaitu terdapat 7 butir soal yang berkualitas menurut hasil analisis menggunakan validitas *Aiken's*, telaah butir soal secara teoritis, dan teori tes klasik. Butir soal yang berkualitas diantaranya yaitu butir soal nomor 5, 8, 11, 13, 20, 22 dan 23.

Referensi

- Adawiah, R., & Ruchliyadi, D. A. (2019). Analisis Kualitas Soal Buatan Guru Pendidikan Pancasila dan Kewarganegaraan (PPKn). *Jurnal Pendidikan Kewarganegaraan*, 9, 58–65.
- Allen, M. J., & Yen, W. M. (1979). *Introduction to Measurement Theory*. Brooks/Cole Publishing Company.
- Arifin, Z. (2014). *Evaluasi Pembelajaran*. Remaja Rosdakarya.
- Arikunto, S. (2013). *Prosedur Penelitian Suatu Pendekatan dan Praktek* (15th ed.). PT Rineka Cipta.
- Arikunto, S. (2018). *Dasar-dasar Evaluasi Pendidikan* (R. Damayanti (ed.); 3rd ed.). Bumi Aksara.
- Azwar, S. (2015). *Reliabilitas dan Validitas* (4th ed.). Pustaka Pelajar.
- Bhakti, Y. B., Awaludin, A. A. R., Chusni, M. Mi., Zakwandi, R., Astuti, I. A. D., Mariasih, Dewi, S., Fitriani, A., Okyranida, I. Y., Haryadi, D., & Gumilar, E. B. (2022). *Evaluasi Pembelajaran Dalam Bidang Pendidikan* (1st ed.). CV. Bintang Semesta Media.
- Budiastuti, D., & Bandur, A. (2018). *Validitas dan Reliabilitas Penelitian Dilengkapi Analisis dengan NVIVO, SPSS, dan AMOS*. Penerbit Mitra Wacana Media.
- Charismana, D. S. (2016). *ANALISIS KUALITAS TES UJIAN AKHIR SEMESTER BUATAN MGMP PADA KURIKULUM 2013 MATA PELAJARAN PPKN SMP KABUPATEN KUDUS*. Tesis. Universitas Negeri Yogyakarta
- Faridah, A. (2021). Karakteristik butir soal ujian akhir semester mata pelajaran sejarah. *Jurnal Penelitian Hukum Dan Pendidikan*, 20(2), 1281–1288.
- Inanna, Rahmatullah, & Hasan, M. (2021). *Evaluasi Pembelajaran: Teori dan Praktek* (1st ed.). Tahta Media Group.
- Iskandar, A., & Rizal, M. (2018). Analisis kualitas soal di perguruan tinggi berbasis aplikasi TAP. *Jurnal Penelitian Dan Evaluasi Pendidikan*, 22(1), 12–23. <https://doi.org/10.21831/pep.v22i1.15609>
- Juliani, R. P., & Erita, S. (2023). Analisis Validitas dan Reliabilitas Instrumen Penilaian Kemampuan Berpikir Kritis dalam Konteks Sekolah Menengah. *Journal of Educational Integration and Development*, 3(3), 169–179.
- Khaerudin. (2015). Kualitas Instrumen Tes Hasil Belajar. *Jurnal Madaniyah*, 2(9), 212–235.
- Khaerudin. (2017). Administarsi, Analisis Butir, dan Kaidah Penulisan Tes. *Madaniyah*, 1, 97–128.
- Kurniawan, E., & Mutaqimah, E. (2009). *Penilaian*. Departemen Pendidikan Nasional.
- Magdalena, I., Fauzi, H. N., Putri, R., & Tangerang, U. M. (2020). *DAN AKIBAT MEMANIPULASINYA*. 2,

244–257.

- Mardapi, D. (2017). *pengukuran, penilaian dan evaluasi pendidikan* (2nd ed.). parama publishing.
- Nurjanah, & Marlianingsih, N. (2015). Analisis Butir Soal Pilihan Ganda Dari Aspek Kebahasaan. *Faktor: Jurnal Ilmu Kependidikan*, 11(1), 69–78.
- Rahman, A. A., & Narsyah, C. E. (2019). *Evaluasi Pembelajaran*. Uwais Inspirasi Indonesia.
- Riinawati. (2021). *Pengantar Evaluasi Pendidikan* (1st ed.). Thema Publishing.
- Rudini, M. (2020). Efektivitas Analisis Butir Soal Mata Pelajaran Matematika Pada Siswa Kelas IV Dalam Meningkatkan Kualitas Guru di SDN Sabang. *Tolis Ilmiah: Jurnal Penelitian*, 2(1), 17–27.
- Setyawarno, D. (2017). Penggunaan Aplikasi Software Itean (Item and Test Analysis) untuk Analisis Butir Soal Pilihan Ganda Berdasarkan Teori Tes Klasik. *Jurnal Ilmu Fisika Dan Pembelajarannya*, 1(1), 11–21. <https://doi.org/doi.org/10.19109/jifp.v1i1.866>
- Sudijono, A. (2011). *Pengantar Evaluasi Pendidikan* (1st ed.). Rajawali Pers.
- Sugiyono. (2019). *Metode Penelitian Kuantitatif Kualitatif Dan R&D* (2nd ed.). Alfabeta.
- Sumaryanta. (2021). *Teori Tes Klasik & Teori Respon Butir: Konsep & Contoh Penerapannya* (1st ed.). CV. Confident.
- Sunaryati, T., Meilania, D. K., Lestari, F., Aliifah, S. N., & Saphira, V. N. (2024). Analisis Instrumen Tes dan Non Tes dalam Evaluasi Pembelajaran. *Jurnal Pendidikan Tambusai*, 8(2), 30461–30472.
- Sutrisno, D. A. (2014). *ANALISIS BUTIR SOAL ULANGAN AKHIR SEMESTER GASAL TAHUN AJARAN 2013/2014 MATA PELAJARAN PKN KELAS VII SEKOLAH MENENGAH PERTAMA NEGERI 5 DEPOK KABUPATEN SLEMAN*. Skripsi. Universitas Negeri Yogyakarta
- Wahyuni, F., Rosyid, R., & Chalimi, I. R. (2019). Analisis Butir Sola Ulangan Akhir Semester PPkn Kelas X SMA N 2 Mempawah tahun Ajaran 2018/2019. *Jurnal Ilmiah Universitas Tanjungpura*, 1(1), 1–7.
- Yusup, F. (2018). uji validitas dan reliabilitas instrumen penelitian kuantitatif. *Jurnal Ilmiah Kependidikan*, 7, 17–23.
- Zainudin, & Ubabuddin. (2023). Ranah Kognitif, Afektif dan Psikomotorik Sebagai Objek Evaluasi Hasil Belajar Peserta Didik. *Islamic Learning Journal*, 1(3), 915–931.
- Zainuri, A., Aquami, & AnNur, S. (2021). *Evaluasi Pendidikan (Kajian Teoritik)* (1st ed.). Penerbit Qiara Media.