

BAB I

PENDAHULUAN

1.1 Latar Belakang

Pendidikan merupakan suatu proses yang sangat penting dalam kehidupan manusia. Melalui pendidikan, seseorang dapat memperoleh pengetahuan keterampilan serta berbagai nilai yang diperlukan agar dapat membentuk kesadaran dan kepribadian individu sehingga siap menyongsong masa depan kehidupan bangsa dan negara yang lebih cerah (Nurkholis, 2013). Upaya dalam melaksanakan pendidikan tidak terlepas dari komponen dan faktor-faktor yang dapat mendukung keberhasilan pelaksanaan pendidikan. Salah satu faktor penting dalam sistem pendidikan adalah penilaian (Kusainun, 2020). Informasi yang didapatkan dari kegiatan penilaian peserta didik dapat digunakan dalam menentukan keputusan ketercapaian tujuan pendidikan serta mempertanggung jawabkan penyelenggaraan pendidikan. Pendidik dapat mengetahui kemampuan yang dimiliki peserta didik, mengevaluasi metode mengajar yang digunakan serta keberhasilan peserta didik dalam meraih kompetensi yang telah ditetapkan dengan melakukan kegiatan penilaian. Melihat bagaimana pentingnya penilaian dalam pelaksanaan pendidikan, pemerintah dalam Permendikbudristek Nomor 21 Tahun 2022 mengatur standar penilaian pendidikan di Indonesia. Standar penilaian merupakan kriteria mengenai ruang lingkup, tujuan, manfaat, prinsip, mekanisme, prosedur, dan instrumen penilaian hasil belajar peserta didik pada pendidikan dasar dan menengah. Standar yang telah ditentukan ini perlu dipahami serta dilaksanakan berdasarkan ketentuan. Pendidik pada setiap satuan pendidikan dalam mengembangkan instrumen penilaian perlu berpedoman pada standar penilaian yang telah ditentukan. Penilaian hasil belajar oleh pendidik terdiri atas penilaian sikap, pengetahuan, dan keterampilan. Pada ranah pengetahuan, teknik penilaian yang dapat digunakan adalah tes tertulis, tes lisan dan penugasan (Setiawati, dkk., 2019). Penilaian dengan teknik tes tertulis di satuan pendidikan dikenal dengan istilah Penilaian Harian, Penilaian Tengah Semester, Penilaian Akhir Tahun serta Penilaian Sumatif Akhir Jenjang untuk peserta didik jenjang akhir di tingkat sekolah dasar dan menengah sebagai salah satu ketentuan kelulusan peserta didik.

Penilaian sumatif yang diadakan oleh pihak sekolah untuk peserta didik jenjang akhir resmi dilaksanakan setelah kemendikbud memutuskan Peniadaan Ujian Nasional dan Ujian Kesetaraan dalam Surat Edaran (SE) Mendikbud Nomor 1 Tahun 2021. Meskipun Ujian Nasional ditiadakan, peserta didik pada tingkat akhir tidak serta-merta dinyatakan lulus. Berdasarkan Surat Edaran Dinas Pendidikan Pemerintah Provinsi Jawa Barat tahun 2024, peserta didik pada tingkat akhir dapat dinyatakan lulus dari satuan pendidikan jika telah memenuhi beberapa ketentuan diantaranya, 1) Telah menyelesaikan program pembelajaran yang dibuktikan rapor pada tiap semester; 2) Memperoleh nilai sikap/perilaku minimal baik; dan 3) Mengikuti Penilaian Sumatif Akhir Jenjang (PSAJ) yang diselenggarakan oleh satuan pendidikan. Penilaian Sumatif Akhir Jenjang merupakan evaluasi pembelajaran yang diselenggarakan oleh satuan pendidikan dalam rangka mengukur ketercapaian dan keberhasilan standar kompetensi lulusan untuk semua mata pelajaran sebagai dasar dalam menentukan kelulusan peserta didik dari satuan pendidikan (Anggraena, dkk., 2022). PSAJ dapat dilaksanakan dalam bentuk: a) Tes tulis; b) Tugas unjuk performa (berupa praktik, menghasilkan produk, melakukan proyek); c) Portofolio (Kumpulan dokumen hasil penilaian, penghargaan, dan karya peserta didik dalam bidang tertentu yang mencerminkan perkembangan (reflektif-integratif) dalam kurun waktu tertentu), atau kombinasi. Teknis pelaksanaan penilaian diserahkan sepenuhnya kepada pihak sekolah sesuai dengan kondisi sekolah maupun peserta didik.

Penilaian Sumatif Akhir Jenjang (PSAJ) merupakan tahap akhir dalam proses pembelajaran yang perlu ditempuh oleh peserta didik di satuan pendidikan serta menjadi salah satu syarat penentu kelulusan. Berdasarkan wawancara yang dilakukan oleh peneliti pada salah satu guru penanggung jawab kegiatan PSAJ mata pelajaran fisika di salah satu Sekolah Menengah Atas Negeri kota Tasikmalaya, pelaksanaan Penilaian Sumatif Akhir Jenjang dilaksanakan dalam 2 bentuk tes, yaitu tes keterampilan dan tes tertulis secara daring di sekolah. Bentuk soal tertulis PSAJ di sekolah tersebut terdiri dari 4 bentuk soal, diantaranya soal pilihan ganda, soal isian singkat, soal benar salah, dan soal menjodohkan. Hasil dari wawancara tersebut juga didapatkan informasi bahwa nilai akhir ujian sekolah yang diperoleh peserta didik di sekolah tersebut merupakan gabungan dari 30%

nilai rapor semester 1-6 dan 70% nilai PSAJ. Hal ini menunjukkan bahwa PSAJ merupakan faktor penting dalam menentukan kelulusan peserta didik dari satuan pendidikan.

Dalam kegiatan PSAJ, instrumen penilaian memiliki peran yang sangat penting karena memiliki fungsi sebagai alat yang digunakan oleh guru untuk menilai tingkat ketercapaian dan keberhasilan standar kompetensi lulusan untuk semua mata pelajaran sebagai dasar dalam menentukan kelulusan peserta didik dari satuan pendidikan. Suatu instrumen penilaian dapat dikatakan baik jika memenuhi kaidah-kaidah penulisan soal, baik yang berkaitan dengan materi, konstruksi, maupun bahasa serta memiliki kualitas parameter butir soal yang teruji seperti tingkat kesulitan, daya pembeda dan faktor tebakan. Kualitas instrumen suatu penilaian dapat diketahui dengan dilakukan analisis instrumen penilaian. Analisis ini merupakan tahapan yang dimaksudkan untuk mengetahui derajat kualitas suatu instrumen penilaian baik untuk setiap butir instrumen maupun secara keseluruhan (Setiayawan dan Wijayanti, 2020). Meskipun demikian, berdasarkan hasil wawancara sebelumnya, fakta di lapangan menunjukkan bahwa selama ini guru belum pernah melakukan analisis terhadap soal yang digunakan dalam PSAJ mata pelajaran fisika. Ketika pembuatan instrumen penilaian, guru hanya diminta untuk memperhatikan bagaimana konstruksi, materi dan bahasa yang digunakan dalam soal. Hal tersebut dikarenakan guru dianggap telah memiliki kemampuan dalam membuat soal yang baik sehingga analisis kualitas soal PSAJ tidak dilakukan. Hal ini mengakibatkan guru tidak dapat mengetahui bagaimana kualitas soal PSAJ mata pelajaran fisika yang selama ini digunakan di sekolah tersebut.

Pada praktiknya, analisis kualitas instrumen penilaian dapat dilakukan melalui dua pendekatan yaitu, analisis secara kualitatif dan analisis secara kuantitatif (Supriyadi, 2011). Analisis kualitas instrumen penilaian secara kualitatif menggunakan format penelaahan oleh pakar untuk mengkaji butir soal dari aspek materi, konstruksi dan bahasa sehingga akan diketahui validitas instrumen penilaian berdasarkan pandangan para pakar dengan memperhatikan ketiga aspek tersebut. Sedangkan analisis kuantitatif merupakan analisis yang didasarkan pada data empirik yang diperoleh dari soal yang telah diujikan. Kedua teknik tersebut masing-masing memiliki keunggulan dan kelemahan, oleh sebab

itu teknik dalam menganalisis instrumen penilaian terbaik adalah dengan menggunakan keduanya (Fitrianawati, 2017). Analisis soal secara kualitatif dapat dilakukan dengan teknik moderator dan teknik panel. Teknik moderator dilakukan dengan metode berdiskusi oleh para ahli terkait bidang mata pelajaran serta moderator untuk menentukan kualitas soal tersebut. Sedangkan teknik panel merupakan kegiatan menelaah butir soal berdasarkan kaidah butir soal yang terdiri dari materi, konstruksi dan bahasa. Analisis kualitatif pada penelitian ini menggunakan format penelaahan oleh pakar yang menguasai materi/isi yang diujikan agar dapat diketahui validitas isinya (Subali, 2014). Selain dilakukan secara teoritis, seperti telaah butir berdasarkan aspek isi, konstruksi dan bahasa, analisis juga perlu dilakukan secara empirik. Analisis kualitas soal secara empirik dapat dibedakan menjadi dua, yaitu pendekatan teori tes klasik dan teori respon butir (Item Response Theory, IRT) (Retnawati, 2016). Salah satu dari pendekatan tersebut dapat digunakan untuk menganalisis kualitas soal penilaian secara kuantitatif. Akan tetapi, analisis soal dengan pendekatan teori tes klasik memiliki beberapa keterbatasan diantaranya, 1) tes klasik bergantung pada sampel dimana estimasi parameter soal bergantung pada karakteristik peserta tes, hal ini dapat berpengaruh pada tingkat kesulitan soal. Tingkat kesulitan soal akan menjadi rendah jika tes diberikan pada peserta didik yang berkemampuan tinggi begitupun sebaliknya, 2) Kesalahan pengukuran (standard error) dianggap sama untuk semua peserta tes, 3) Estimasi kemampuan peserta didik bergantung pada karakteristik butir soal sehingga estimasi kemampuan peserta didik akan menjadi rendah apabila soal yang diberikan di atas kemampuannya dan sebaliknya (Azevedo, dkk., 2019). Keterbatasan tersebut mengakibatkan hasil analisis dengan pendekatan teori tes klasik kurang merefleksikan kemampuan peserta didik. Untuk melengkapi keterbatasan pendekatan teori klasik, dikembangkan pendekatan instrumen modern berupa teori respon butir atau yang disebut *Item Response Theory* (IRT). Pada pendekatan ini, teori respon butir didasarkan pada tiga asumsi, yaitu 1) undimensi, setiap butir tes hanya mengukur satu kemampuan saja, 2) independensi lokal, jawaban peserta tes terhadap suatu butir tes tidak mempengaruhi jawaban peserta tes pada butir soal lainnya, dan 3) invariansi parameter, karakteristik butir soal tidak bergantung pada distribusi parameter

kemampuan peserta tes dan parameter yang menjadi ciri dari peserta tes tidak bergantung pada ciri butir soal (Eaton, dkk., 2019). Berdasarkan pertimbangan tersebut, kegiatan analisis kualitas soal pada Penilaian Sumatif Akhir Jenjang akan dilakukan secara kualitatif dengan melihat validitas isi menggunakan teknik gregory dan secara kuantitatif menggunakan pendekatan teori respon butir dengan bantuan program e-IRT yang terpasang pada add-in di aplikasi Microsoft Excel. Program ini dibuat oleh profesional untuk dapat melakukan estimasi Teori Respon Butir baik menggunakan pengukuran 1 parameter, pengukuran 2 parameter ataupun pengukuran 3 parameter. Program e-IRT dapat melakukan analisis butir soal berbentuk pilihan ganda, soal dikotomi serta soal berbentuk esai atau angket.

Berdasarkan uraian masalah yang telah dipaparkan diatas, peneliti tertarik untuk melakukan penelitian berupa analisis kualitas soal Penilaian Sumatif Akhir Jenjang (PSAJ) untuk mengetahui bagaimana kualitas instrumen soal tersebut. Analisis soal dilakukan untuk mengetahui derajat kualitas soal berdasarkan validitas, reliabilitas dan parameter butir tes. Dengan demikian, judul penelitian ini adalah **“Analisis Kualitas Soal Penilaian Sumatif Akhir Jenjang (PSAJ) Mata Pelajaran Fisika Kelas XII berbantuan program e-IRT”**.

1.2 Rumusan Masalah

Pada penelitian ini, rumusan masalah yang menjadi fokus peneliti adalah “Bagaimana kualitas soal pada Penilaian Sumatif Akhir Jenjang Mata Pelajaran Fisika kelas XII tahun ajaran 2023/2024 yang dianalisis secara kualitatif berupa validitas isi teknik gregory dan secara kuantitatif dengan pendekatan teori respon butir?”. Rumusan masalah tersebut dapat dikembangkan dalam pertanyaan penelitian sebagai berikut.

1. Bagaimana hasil analisis kualitatif instrumen Penilaian Sumatif Akhir Jenjang (PSAJ) Mata Pelajaran Fisika Kelas XII tahun ajaran 2023/2024 di salah satu SMA kota Tasikmalaya menggunakan teknik gregory?
2. Model parameter logistik mana yang paling sesuai untuk menganalisis kualitas soal PSAJ Fisika Kelas XII tahun ajaran 2023/2024 di salah satu SMA kota Tasikmalaya dengan pendekatan teori respon butir?
3. Bagaimana reliabilitas dan validitas empirik instrumen PSAJ Fisika Kelas XII tahun ajaran 2023/2024 di salah satu SMA Kota Tasikmalaya melalui analisis

parameter logistik?

4. Bagaimana parameter butir tes (daya pembeda, tingkat kesulitan, faktor tebakan) instrumen PSAJ Fisika Kelas XII tahun ajaran 2023/2024 di salah satu SMA Kota Tasikmalaya melalui analisis parameter logistik?

1.3 Tujuan Penelitian

Tujuan penelitian ini secara umum adalah untuk mengetahui bagaimana kualitas instrumen penilaian yang digunakan pada Penilaian Sumatif Akhir Jenjang mata pelajaran Fisika kelas XII tahun ajaran 2023/2024 di salah satu SMA Kota Tasikmalaya yang dianalisis secara kualitatif berupa validitas isi dengan teknik gregory dan secara kuantitatif dengan pendekatan teori respon butir berbantuan program *e-IRT*.

1.4 Manfaat Penelitian

Setelah penelitian dilaksanakan, hasil dari penelitian ini diharapkan dapat memberikan manfaat sebagai berikut.

1. Secara teoritis, penelitian ini diharapkan dapat menambah pengetahuan khususnya dalam pembahasan mengenai analisis instrumen penilaian hasil belajar peserta didik.
2. Secara praktis, penelitian ini diharapkan dapat memberikan gambaran bagaimana menganalisis soal penilaian secara kualitatif berupa validasi isi dengan teknik gregory dan secara kuantitatif dengan pendekatan teori respon butir berbantuan program *e-IRT*.

1.5 Definisi Operasional

1.5.1 Reliabilitas

Reliabilitas tes merupakan suatu nilai yang menunjukkan sejauh mana suatu tes mampu menunjukkan ke-konsistenan hasil pengukurannya. Pada pendekatan teori respon butir, reliabilitas suatu soal tes dapat diketahui melalui fungsi informasi yang kebenarannya bersifat probabilitas dan tidak luput dari kesalahan pengukuran atau *Standard Error of Measurement* (SEM). Reliabilitas soal didapatkan dengan menghitung nilai SEM lalu membandingkannya dengan nilai fungsi informasi soal. Semakin besar nilai fungsi informasi maka semakin kecil

nilai SEM dan semakin kecil nilai SEM maka soal tes tersebut semakin tepat dan reliabel. Selain membandingkan nilai fungsi informasi dengan nilai SEM, reliabilitas suatu tes pada taraf kemampuan tertentu dapat diketahui dengan memanfaatkan kurva fungsi informasi dengan kurva SEM dengan melihat titik perpotongan diantara kedua kurva tersebut.

1.5.2 Validitas

Validitas merupakan ciri yang menandai tes hasil belajar yang baik serta ketepatan suatu alat ukur dalam mengukur apa yang ingin diukur berdasarkan tujuan yang ingin dicapai. Pada penelitian ini, validasi dilakukan secara teoritis berupa validitas isi dengan teknik gregory dan validitas empirik dengan menentukan butir soal mana yang sesuai dengan model parameter logistik yang digunakan. Butir soal yang cocok dengan model parameter logistik yang digunakan dapat diketahui dengan membandingkan nilai *chi-square* hitung dengan nilai *chi-square* tabel atau dengan melihat nilai *p-values* yang didapatkan dari hasil perhitungan menggunakan program *e-IRT*.

Validasi isi dilakukan menggunakan teknik gregory dengan tahapan yang dimulai dengan melakukan penilaian oleh 2 pakar sebagai validator. Dari hasil penilaian, skor 1 hingga 4 yang diberikan validator dikelompokkan menjadi dua kategori yaitu relevansi kuat (skor 3 dan 4) dan relevansi lemah (skor 1 dan 2) lalu dimasukkan ke tabulasi silang 2x2. Kemudian informasi dari tabulasi silang 2x2 dimasukkan kedalam persamaan validitas isi gregory agar diketahui nilai validitas isi soal penilaian tersebut.

1.5.3 Tingkat Kesulitan

Tingkat kesulitan merupakan bilangan yang menunjukkan besarnya proporsi peserta tes yang menjawab benar pada butir soal. Bilangan ini biasanya dinyatakan dalam bentuk indeks kesukaran dengan nilai antara 0 dan 1. Tingkat kesulitan pada penelitian ini diukur menggunakan bantuan program *e-IRT*. Pada eori respon butir, indeks tingkat kesulitan yang baik berada pada interval -2 sampai dengan 2. Apabila tingkat kesulitan suatu butir soal dibawah -2,0 atau di atas 2,0 maka butir soal tersebut tidak akan digunakan atau dibuang.

1.5.4 Daya Pembeda

Daya pembeda merupakan kemampuan soal dalam membedakan peserta tes berkemampuan tinggi dengan berkemampuan rendah setelah menjawab soal. Pada penelitian ini, daya pembeda diukur menggunakan bantuan program *e-IRT*. Daya pembeda yang baik berada pada interval mulai dari 0 sampai dengan 2. Apabila indeks daya pembeda suatu butir soal bernilai negatif atau diatas 2 maka daya pembeda dari butir tersebut kurang baik sehingga butir sola tersebut sebaiknya tidak digunakan

1.5.5 Faktor Tebakan

Faktor tebakan merupakan besar peluang peserta tes untuk menebak soal dengan benar. Pada penelitian ini faktor tebakan dapat diketahui dengan bantuan program *e-IRT*. Faktor tebakan dikatakan baik jika nilai faktor tebakan mendekati nol atau tidak lebih dari $1/k$ dimana k adalah banyaknya pilihan jawaban pada butir soal.