

BAB II

KAJIAN PUSTAKA DAN KERANGKA BERPIKIR

A. Kajian Pustaka

1. Asesmen Pembelajaran

Berdasarkan Peraturan Menteri Pendidikan Nasional Nomor 20 Tahun 2007 tentang Standar Penilaian Pendidikan dan Peraturan Menteri Pendidikan dan Kebudayaan Nomor 66 Tahun 2013 tentang Standar Penilaian Pendidikan, yang dimaksud dengan penilaian ialah proses mengumpulkan dan mengolah informasi untuk memutuskan capaian hasil belajar siswa. Penilaian (*asesment*) adalah serangkaian kegiatan yang dilakukan untuk mendapatkan, menelaah, dan merumuskan data mengenai tahap dan hasil belajar siswa dengan runtut dan berkelanjutan agar dapat menjadi penjelasan penting dalam mengambil keputusan pada akhir pembelajaran.

Hill dan Ruptic (dalam Basuki dan Hariyanto, 2017: 8) mengemukakan bahwa asesmen ialah suatu prosedur mengumpulkan bukti dan pendokumentasian pembelajaran serta perkembangan anak-anak. Dari pernyataan tersebut dapat disimpulkan bahwa asesmen merupakan tahap yang rinci dan berkelanjutan dalam mengumpulkan data mengenai hasil belajar siswa, yang bertujuan untuk meningkatkan keefektifan proses belajar dan mengajar.

Basuki dan Hariyanto (2017: 154) mengemukakan bahwa tujuan penilaian (*asesment*) meliputi beberapa aspek sebagai berikut:

- a. Pemberian tugas tertentu untuk mengevaluasi kemampuan individu.
- b. Mengidentifikasi keperluan pembelajaran.
- c. Menolong dan memotivasi siswa dalam proses belajar.
- d. Mendukung proses guru dalam mengajar dengan baik.
- e. Menetapkan strategi pembelajaran yang efektif.
- f. Menunjukkan tanggung jawab lembaga pendidikan.
- g. Meningkatkan mutu pendidikan.

Santoso (dalam Basuki dan Hariyanto, 2017: 156), menjabarkan prinsip-prinsip penilaian (*asesment*) secara umum sebagai berikut:

- a. *Keeping track*, adalah kemampuan untuk memantau antara perkembangan kemampuan siswa terhadap rencana pembelajaran yang telah ditentukan.
- b. *Checking up*, adalah kemampuan untuk memeriksa sejauh mana kemampuan siswa tercapai selama tahap pembelajaran.
- c. *Finding out*, adalah kemampuan guna menemukan kekeliruan sehingga menyebabkan adanya kelemahan dalam tahap pembelajaran.
- d. *Summing up*, adalah kemampuan guna menarik kesimpulan apakah kompetensi yang ditetapkan telah dicapai oleh siswa atau belum.

Basuki dan Hariyanto (2017: 158-159) mengidentifikasi fungsi asesmen dengan tiga tujuan utama sebagai berikut.

- a. Asesmen untuk pembelajaran (*assessment for learning*)

Ciri-cirinya meliputi :

- 1) Terdapat dua fase yakni asesmen diagnostik dan formatif.

- 2) Penggunaan portofolio, kemajuan kerja, pengamatan guru, dan percakapan sebagai sumber informasi.
- 3) Umpan balik kepada siswa, baik lisan maupun tertulis, bersifat deskriptif dan penekanan pada kekuatan, identifikasi tantangan, serta langkah selanjutnya.
- 4) Guru memantau pengetahuan siswa dan mengelola pembelajaran agar tujuan yang dicapai sesuai.
- 5) Tidak melibatkan skor atau nilai numerik.
- 6) Terjadi selama proses pembelajaran, dimulai dari perencanaan awal hingga penilaian sumatif.

Dapat disimpulkan, asesmen untuk pembelajaran merupakan asesmen yang terdiri atas asesmen diagnostik dan asesmen formatif, tidak ada skor atau angka mutu, dan berlangsung selama proses pembelajaran hingga asesmen sumatif dilaksanakan.

b. Asesmen sebagai pembelajaran (*assessment as learning*)

Ciri-cirinya meliputi :

- 1) Dilaksanakan ketika siswa memahami capaian pembelajaran dan kriteria pengerjaan yang harus dicapai.
- 2) Meliputi rumusan capaian pembelajaran, pemantauan, dan refleksi atas hasil pembelajaran.
- 3) Penekanan pada kepemilikan hasil belajar oleh siswa dan tanggung jawab mereka dalam mengarahkan pemikiran ke depan (metakognitif).

4) Terjadi sepanjang tahap kegiatan pembelajaran.

Dapat disimpulkan bahwa asesmen sebagai pembelajaran adalah asesmen dilakukan di setiap proses pembelajaran yang berimplikasi pada hasil belajar siswa.

c. Asesmen pembelajaran (*assessment of learning*)

Ciri-cirinya meliputi :

- 1) Melibatkan nilai atau huruf mutu (sumatif).
- 2) Adanya perbandingan antara prestasi siswa dengan standar yang ditetapkan.
- 3) Hasil penilaian dapat dibicarakan kepada siswa serta orang tua.
- 4) Dilakukan pada akhir setiap satuan pembelajaran.

Dapat disimpulkan bahwa asesmen pembelajaran adalah asesmen yang dilakukan pada akhir satuan pembelajaran yang dilengkapi dengan skor atau angka mutu dan hasil belajar siswa disampaikan kepada yang bersangkutan.

Jenis asesmen pembelajaran yang dapat digunakan pada kurikulum merdeka terbagi menjadi 3 jenis, yaitu:

a. Asesmen Diagnostik

Menurut Nasution (2022) asesmen diagnostik merupakan asesmen yang dilakukan untuk mengetahui kondisi awal siswa atau kemampuan dasar siswa. Hasil informasi dari asesmen diagnostik yang berkaitan dengan latar belakang keluarga, kesiapan belajar, motivasi, serta minat belajar peserta didik bisa digunakan untuk

mempertimbangkan perencanaan pembelajaran. Dalam pelaksanaannya, asesmen diagnostik dapat dilaksanakan sesuai dengan kebutuhan, misalnya pada awal tahun pelajaran, awal materi baru, dan sebelum menyusun modul ajar mandiri (Sufyadi et al., 2021). Terdapat dua jenis asesmen diagnostik yaitu: asesmen diagnostik kognitif dan asesmen diagnostik non-kognitif. Berikut penjelasannya menurut Nasution (2022: 138) :

- 1) Asesmen Kognitif, asesmen yang bertujuan untuk mengetahui capaian kompetensi siswa, penyesuaian pembelajaran berdasarkan kompetensi rata-rata siswa, dan memberikan remedial atau pelajaran tambahan untuk siswa yang berkompotensi di bawah rata-rata.
- 2) Asesmen Non-Kognitif, asesmen yang dilakukan pada awal pembelajaran untuk mengetahui psikologi dan sosial emosi siswa, aktivitas siswa ketika belajar di rumah, kondisi keluarga siswa, latar belakang pertemanan dan pergaulan siswa, serta untuk mengetahui bagaimana gaya belajar, karakter, dan minat siswa.

b. Asesmen Formatif

Nasution (2021), menyatakan bahwa asesmen formatif ialah asesmen yang dilakukan untuk mengetahui pemahaman murid, keperluan pembelajaran, dan kemajuan akademik selama proses pembelajaran. Asesmen formatif juga digunakan untuk mengamati

pembelajaran dan memberikan umpan balik (*feedback*) bagi guru secara berkelanjutan. Menurut Muktamar et al., 2024) manfaat pemberian asesmen formatif bagi guru sebagai berikut.

- 1) Mendapatkan informasi mengenai keperluan belajar siswa.
- 2) Mengetahui penguasaan serta kelemahan siswa dalam memahami materi.
- 3) Memudahkan guru dalam meramalkan tingkat keberhasilan siswa saat asesmen sumatif.
- 4) Memperkirakan tingkat keberhasilan suatu proses pembelajaran yang diberikan.
- 5) Memudahkan guru dalam merencanakan topik pembelajaran.
- 6) Sebagai bahan untuk menciptakan proses pembelajaran yang lebih baik kedepannya.

Sedangkan manfaat asesmen formatif bagi siswa sebagai berikut.

- 1) Siswa dapat mengetahui tingkat pemahamannya terhadap materi pelajaran.
- 2) Memudahkan siswa dalam mengenali tantangan apa saja yang membuatnya kesulitan dalam memahami materi.
- 3) Memudahkan siswa dalam menentukan langkahnya agar dapat mencapai hasil belajar yang maksimal.
- 4) Membuat siswa untuk lebih menghargai proses pembelajaran daripada hasil akhir.

Dengan demikian dapat disimpulkan bahwa asesmen formatif dilakukan untuk mengetahui perkembangan siswa serta umpan balik kepada guru selama proses pembelajaran.

c. Asesmen Sumatif

Asesmen sumatif menurut Nasution (2022), ialah asesmen pembelajaran yang dilakukan pada akhir pembelajaran. Asesmen sumatif dianggap memiliki pengaruh yang besar terhadap nilai akhir siswa, oleh karena itu asesmen sumatif sering diutamakan daripada asesmen formatif. Selaras dengan pendapat (Muktamar et al., 2024) asesmen sumatif dilakukan untuk memastikan tujuan pembelajaran secara keseluruhan telah tercapai. Asesmen ini sering dilakukan pada akhir pembelajaran, seperti akhir semester dan akhir tahun ajaran pendidikan. Hasil dari asesmen sumatif ini berfungsi untuk menentukan apakah siswa layak untuk naik kelas atau lulus (Nurgiantoro, 2001: 69).

Pemberian asesmen sumatif juga dapat bermanfaat bagi guru maupun siswa. Berikut manfaat asesmen sumatif bagi guru menurut (Muktamar et al., 2024) :

- 1) Memudahkan guru dalam menetapkan nilai setiap siswa agar dibandingkan dengan siswa yang lain.
- 2) Memberikan umpan balik bagi guru.
- 3) Sebagai dasar dalam menentukan apakah siswa dapat naik kelas atau tidak.

- 4) Untuk informasi terkait kemajuan belajar siswa serta bahan pemberitahuan hasil belajar siswa kepada orang tuanya.

Manfaat asesmen sumatif bagi siswa sebagai berikut.

- 1) Sebagai umpan balik untuk siswa agar mereka dapat meningkatkan atau mempertahankan hasil belajar mereka untuk kedepannya.
- 2) Sebagai informasi bagi siswa apakah mereka dapat naik kelas atau tidak.
- 3) Sebagai bukti terkait apa saja yang sudah berhasil mereka capai selama proses pembelajaran.

Berdasarkan pernyataan yang telah dipaparkan di atas, dapat disimpulkan bahwa asesmen merupakan bagian penting dalam proses pembelajaran. Asesmen diagnostik untuk mengetahui kekuatan dan kelemahan siswa. Asesmen formatif untuk memberikan umpan balik bagi guru maupun siswa secara berkelanjutan. Asesmen sumatif untuk mengetahui pencapaian siswa secara keseluruhan serta pengambilan keputusan terkait kelanjutan proses belajar siswa.

Pada penelitian ini yang menjadi fokus adalah kegiatan asesmen sumatif akhir semester, di mana asesmen ini dilakukan pada akhir semester lebih tepatnya semester ganjil tahun pelajaran 2023/2024 di SMAN 1 Nglames. Pelaksanaan asesmen ini bertujuan untuk mengetahui tingkat pemahaman siswa serta tujuan pembelajaran selama satu semester telah tercapai.

2. Jenis-Jenis Tes

a. Teknik Tes

Menurut Nurkencana dan Sumartana (dalam Nurgiantoro, 2001: 58), tes ialah suatu metode evaluasi yang menyertakan penugasan yang harus diselesaikan oleh siswa guna mengumpulkan data mengenai prestasi mereka, yang nantinya dapat digunakan sebagai dasar perbandingan terhadap pencapaian teman sejawat mereka atau standar yang telah ditetapkan. Anderson (dalam Nurgiantoro, 2001: 59) menyatakan pandangan serupa, bahwa tes ialah serangkaian pertanyaan yang digunakan untuk mengukur keterampilan, pengetahuan, atau bakat individu dan kelompok. Berdasarkan penyusunannya jenis tes dibedakan menjadi dua, yakni tes buatan guru dan tes standar.

1) Tes Buatan Guru

Tes buatan guru disusun oleh guru kelas masing-masing yang bersangkutan. Fungsinya adalah untuk mengevaluasi seberapa jauh siswa mencapai capaian pembelajaran yang sudah diatur oleh guru tersebut setelah proses pengajaran di kelas berlangsung (Nurgiantoro, 2001: 60).

Menurut Arikunto (2018: 104), tes buatan guru dibuat berdasarkan materi dan capaian yang ditetapkan oleh guru demi kepentingan kelas, meliputi materi pengetahuan dan keterampilan yang terbatas. Guru menyusun tes ini sendiri tanpa dibantu orang lain atau ahli, guru juga jarang menggunakan item tes yang sudah pernah

diuji coba, memiliki reliabilitas menengah atau rendah, dan aturan kelompok yang terbatas pada kelas tertentu.

Umumnya, tes yang dibuat oleh guru tidak diuji coba terlebih dahulu dalam praktik pendidikan, keterbatasan waktu, kesempatan, tenaga, biaya, dan kemampuan guru menjadi faktor kendala ketika akan menganalisis butir tes. Akibatnya, tes buatan guru cenderung memiliki tingkat keandalan atau validitas yang rendah (Nurgiantoro, 2001: 61).

Arikunto (2018: 106) menjelaskan beberapa manfaat dari penggunaan tes buatan guru sebagai berikut:

1. Dapat mengetahui tingkatan siswa dalam menguasai materi yang telah diberikan pada periode waktu tertentu.
2. Dapat mengevaluasi pencapaian tujuan pembelajaran.
3. Dapat memberikan penilaian kepada siswa.

2) Tes Standar

Menurut Arikunto (2018: 102), istilah "standar" dalam konteks tes mengacu pada situasi di mana semua siswa menjalankan tes dengan mengikuti instruksi yang seragam. Dengan demikian, tercipta kesan bahwa ada norma atau ukuran yang akan menghasilkan hasil yang serupa, memungkinkan perbandingan antara kinerja kelompok dengan standar tertentu.

Tes standar dibuat berdasarkan pada bahan dan tujuan umum dari sekolah di semua negara, mencakup aspek yang luas dan

pengetahuan dengan butir tes yang sedikit di setiap keterampilannya, disusun dengan beberapa kelengkapan seperti: staf profesor, pembahas, editor, butir tes, penggunaan butir tes yang telah diujicobakan, memiliki reliabilitas tinggi, dan penggunaan norma yang berlaku di seluruh negara (Arikunto, 2018: 104).

Sebuah tim sengaja dibentuk dalam kegiatan penulisan tes standar. Kurikulum dan buku teks secara nasional dijadikan sebagai dasar dalam menyeleksi bahan dan tujuan tes (Tuckman dalam Nurgiantoro, 2001: 63).

Menurut Arikunto (2018: 105), tes standar memiliki beberapa kegunaan, antara lain:

- a. Mengetahui perbandingan antara prestasi belajar individu dengan kelompok.
- b. Mengetahui perbandingan tingkat prestasi siswa pada berbagai keterampilan, baik individu maupun kelompok.
- c. Mengetahui perbandingan prestasi siswa antar sekolah atau kelas.
- d. Menganalisis perkembangan siswa dalam jangka waktu tertentu.

Berdasarkan bentuknya tes dapat dibedakan menjadi dua macam yakni; tes subjektif dan tes objektif.

1. Tes Subjektif

Menurut Ebel (dalam Nurgiantoro, 2001: 71) tes subjektif mengharuskan siswa untuk menghubungkan antara fakta, konsep,

lalu mengorganisasikannya dalam kepaduan yang masuk akal dan mengungkapkan hasil pemikiran tersebut dalam bentuk tulisan.

Tes esai merupakan salah satu bentuk dari tes subjektif karena kegiatan penilaiannya bersifat subjektif (Nurgiantoro, 2001: 71). Dalam pengerjaan tes esai, siswa diberikan kebebasan untuk menata serta menjabarkan jawabannya dengan gaya bahasanya sendiri dalam lingkup yang relatif dibatasi (Tuckman dalam Nurgiantoro, 2001: 71).

Menurut Arikunto (2018: 108), tes esai merupakan salah satu jenis tes yang mengharuskan siswa memberikan jawaban berupa pembahasan atau uraian kata dengan gaya bahasanya sendiri. Bentuk pertanyaan tes esai biasanya diawali dengan kata yang memiliki makna penjabaran seperti: uraikan, jelaskan, mengapa, bagaimana, dan sejenisnya. Tes esai mengharuskan siswa agar mampu mengelola, menafsirkan, dan menghubungkan definisi-definisi yang sudah dipelajari. Tes esai mengharuskan siswa agar mengingat dan mengenali informasi yang telah dipelajari, serta memiliki kreativitas yang tinggi.

Kelebihan tes subjektif menurut Arikunto (2018: 109) sebagai berikut.

- a. Persiapan dan penyusunan yang mudah.
- b. Mengurangi kesempatan untuk berpikir curang dan mengandalkan keberuntungan.

- c. Mendongkrak siswa untuk menyatakan opini dan menyusun kalimatnya dengan baik.
- d. Memberi peluang pada siswa untuk mengekspresikan pemikiran mereka dengan gaya bahasanya mereka sendiri.
- e. Memungkinkan penilaian seberapa jauh siswa memahami masalah yang diuji.

Kekurangan tes subjektif menurut Arikunto (2018: 109) sebagai berikut.

- a. Reliabilitas dan validitasnya rendah karena sulit untuk menentukan dengan pasti pengetahuan mana yang telah benar-benar dipahami oleh siswa.
- b. Kurang representatif karena jumlah soal terbatas.
- c. Rentan terhadap pengaruh unsur subjektif dalam proses penilaian.
- d. Memerlukan lebih banyak pertimbangan individual dari penilai dalam proses pemeriksaannya.
- e. Memerlukan waktu relatif lama guna mengoreksi dan tidak dapat didelegasikan pada orang lain.

2. Tes Objektif

Tes objektif atau dapat disebut juga sebagai tes jawaban singkat ialah tes yang mengharuskan siswa untuk memberikan tanggapan singkat, meskipun mereka memilih kode-kode tertentu

yang mewakili opsi pilihan jawaban yang sudah disediakan (Nurgiantoro, 2001: 76).

Manfaat tes objektif menurut Arikunto (2018:110-111) adalah sebagai berikut.

- a. Isi dan panjang materi lebih terwakili, lebih objektif, dan dapat menghindari campur tangan subjektivitas siswa atau guru.
- b. Mudah dan cepat dalam proses verifikasinya.
- c. Orang lain dapat mewakili pemeriksaannya.
- d. Dalam pemeriksaan tidak ada komponen objektif yang berpengaruh terhadapnya.

Kekurangan tes objektif menurut Arikunto (2018: 111) sebagai berikut.

- a. Tahap penyiapan dan penyusunan soal lebih sulit.
- b. Soal-soal lebih menekankan pada fase ingatan dan daya mengenali kembali suatu materi saja.
- c. Besarnya peluang untuk mengerjakan curang.
- d. Lebih terbukanya peluang kerja sama antar siswa.

Terdapat beberapa macam penilaian objektif yakni; tes benar-salah, tes pilihan ganda, tes isian, dan tes penjumlahan. Berikut penjelasan terkait macam tes objektif tersebut.

- a. Tes Benar-Salah (*true-false*)

Tes benar-salah merupakan jenis penilaian yang berisi pernyataan dan terdiri atas dua alternatif jawaban: benar atau

salah. Dalam menjalani tes ini, siswa harus benar-benar memahami pernyataan yang diberikan (Nurgiantoro, 2001: 79).

b. Tes Pilihan Ganda (*multiple choice*)

Pada dunia pendidikan tes pilihan ganda merupakan salah satu bentuk tes yang paling umum digunakan. Item tes pilihan ganda digunakan untuk menilai definisi suatu pengetahuan dalam materi pembelajaran yang sudah jelas (Sukardi, 2015: 125). Tes ini terdiri dari pernyataan yang belum selesai, diikuti dengan beberapa opsi yang bisa melengkapinya. Di antara opsi-opsi tersebut, hanya ada satu opsi yang benar, sementara lainnya merupakan distraktor (Nurgiantoro, 2001: 83).

c. Tes Isian (*completion*)

Tes isian adalah jenis penilaian objektif yang susunannya terdiri dari pernyataan atau kalimat dengan bagian tertentu yang sengaja dihapus atau tidak diselesaikan. Bagian yang hilang ini adalah elemen penting yang harus dijawab oleh siswa (Nurgiantoro, 2001: 89).

d. Tes Penjodohan (*matching*)

Tes penjodohan adalah salah satu jenis tes yang meminta siswa untuk mencocokkan dua set pernyataan yang diberikan. Pernyataan-pernyataan ini biasanya disusun dalam dua kolom: kolom kiri berisi pernyataan atau pertanyaan utama,

sementara kolom kanan berisi jawaban yang tepat dengan pernyataan di kolom kiri (Nurgiantoro, 2001: 91).

b. Teknik Non Tes

Teknik non tes ialah metode penilaian yang bertujuan untuk mengumpulkan data tentang kondisi siswa dengan tidak menggunakan alat tes tradisional. Teknik ini bertujuan untuk memperoleh informasi yang berkaitan secara tidak langsung dengan perilaku kognitif (Nurgiantoro, 2001: 54).

Menurut Arikunto (2018: 41-44), ada berbagai jenis alat penilaian non tes, termasuk skala bertingkat, kuesioner, daftar periksa, wawancara, observasi, dan riwayat hidup. Berikut penjelasan dari masing-masing teknik non tes berdasarkan pemikiran Arikunto (2018: 41-44).

1) Skala Bertingkat (*Rating Scale*)

Skala merupakan representasi nilai dalam bentuk angka berdasarkan suatu evaluasi. Angka-angka pada skala disusun secara berurutan dari yang terendah hingga tertinggi, sehingga disebut sebagai skala bertingkat.

2) Kuesioner (*Questionair*)

Kuesioner merupakan susunan daftar pertanyaan yang harus diisi oleh individu yang hendak diukur (responden). Tujuan dari kuesioner adalah untuk memperoleh informasi mengenai kondisi atau data pribadi, pengalaman, pengetahuan, sikap, atau pendapat dari responden.

3) Daftar Cocok (*Check List*)

Daftar cocok adalah kumpulan pertanyaan, yang biasanya singkat, di mana responden yang dievaluasi menandai tanda centang (☐) di tempat yang telah disediakan.

4) Wawancara (*Interview*)

Salah satu cara yang dapat digunakan untuk memperoleh jawaban responden melalui sesi tanya jawab di mana responden tidak diberi kesempatan untuk mengajukan pertanyaan ialah dengan melakukan wawancara. Berdasarkan pelaksanaannya, wawancara terbagi menjadi dua jenis:

- a. Wawancara Bebas, di mana responden bebas menyampaikan pendapatnya tanpa dibatasi oleh pedoman yang ditetapkan oleh pewawancara.
- b. Wawancara Terpimpin, di mana pewawancara mengajukan pertanyaan yang telah disiapkan sebelumnya, dan responden hanya dapat memilih jawaban dari yang telah disiapkan.

5) Pengamatan (*Observation*)

Pengamatan ialah teknik yang melibatkan observasi mendalam dan pencatatan yang sistematis. Ada tiga jenis observasi:

- a. Observasi Partisipan, di mana pengamat ikut serta langsung dalam kegiatan kelompok yang diobservasi.
- b. Observasi Sistematis, di mana aspek-aspek yang diobservasi telah diidentifikasi dan dikategorikan secara sistematis.

c. Observasi Eksperimental, di mana pengamat tidak terlibat dalam kegiatan kelompok yang diobservasi.

6) Riwayat Hidup

Riwayat hidup merupakan deskripsi mengenai perjalanan hidup seseorang. Dari riwayat hidup ini, dapat disimpulkan kepribadian, kebiasaan, dan sikap individu yang dinilai.

Dapat disimpulkan bahwa riwayat hidup berisi tentang deskripsi singkat mengenai informasi kehidupan seseorang dalam aspek identitas diri, pendidikan, kepribadian, pengalaman, sikap, dan kebiasaannya.

3. Analisis Kualitas Butir Soal

Kegiatan penting dalam pengembangan tes untuk memperoleh soal yang berkualitas serta sesuai dengan kriteria serta tujuan pembelajaran yang telah ditetapkan adalah dengan melakukan analisis kualitas butir soal. Tujuan analisis soal, sebagaimana diungkapkan oleh Kusaeri (2013: 102), adalah untuk mengevaluasi setiap butir soal agar dapat memastikan kualitasnya sebelum digunakan, melakukan perbaikan untuk meningkatkan kualitas soal, serta membuang butir soal yang tidak layak. Tuckman (dalam Nurgiantoro, 2001: 136) mengungkapkan maksud dari analisis butir soal adalah untuk memastikan konsistensi setiap butir soal dengan keseluruhan tes. Menurut Oller (1979: 245), analisis soal adalah langkah penting dan perlu dalam persiapan tes pilihan ganda yang baik. Analisis pada butir soal dilakukan guna mengetahui tingkat kesukaran, daya

pembeda, dan efektivitas pengecoh (*distractor*). Berikut penjelasan dari ketiga aspek kualitas butir soal tersebut.

a. Tingkat Kesukaran

Menurut Oller (dalam Nurgiantoro, 2001: 138), tingkat kesukaran ialah dasar petunjuk untuk mengetahui seberapa mudah atau sulit suatu butir soal untuk siswa yang dites. Butir soal yang tingkat kesukarannya sedang maka dapat dikatakan baik dan layak digunakan. Apabila terlalu mudah atau terlalu sulit maka soal tersebut tidak efektif untuk digunakan karena tidak mampu untuk menunjukkan perbedaan antara siswa dengan kemampuan tinggi dengan siswa berkemampuan rendah. Soal seperti ini tidak dapat memberi informasi yang berguna mengenai perbedaan kemampuan antar individu (Noll dalam Nurgiantoro, 2001: 138).

Siswa tidak akan ter tuntut untuk berusaha lebih keras dalam menyelesaikan suatu butir soal, jika soal tersebut terlalu mudah. Sebaliknya, siswa juga akan mudah menyerah serta putus asa jika soal terlalu sulit karena merasa soal tersebut melebihi batas kemampuan mereka (Arikunto, 2018: 232).

Angka atau bilangan yang bertujuan untuk menunjukkan sukar dan mudahnya suatu soal disebut dengan indeks kesukaran (*difficulty indeks*). Besaran indeks kesukaran dapat berkisar dari 0,00 hingga 1,0. Soal dikatakan terlalu sukar jika indeks kesukarannya 0,0. Sebaliknya,

soal dikatakan terlalu mudah jika indeks kesukarannya 1,0 (Arikunto, 2018: 232).

Rumus untuk menghitung indeks kesukaran (P) menurut Arikunto (2018: 233) sebagai berikut.

$$P = \frac{B}{JS}$$

P = Indeks kesukaran

B = banyaknya siswa yang menjawab soal itu dengan betul

JS = Jumlah seluruh siswa peserta tes

Indeks kesukaran menurut Arikunto (2018: 235) sebagai berikut.

Soal dengan P 0,00 sampai 0,30 adalah soal sukar

Soal dengan P 0,31 sampai 0,70 adalah soal sedang

Soal dengan P 0,71 sampai 1,00 adalah soal mudah

Oller (dalam Nurgiantoro, 2001: 138) menyatakan jika butir soal yang memiliki indeks kesukaran berkisar dari 0,15 - 0,85 maka soal tersebut dapat dikatakan layak. Apabila indeks di luar itu maka dapat dipastikan bahwa butir soal terlalu mudah atau terlalu sulit, sehingga perlu adanya revisi atau perbaikan.

Rumus untuk menghitung tingkat kesulitan butir soal menurut Oller (dalam Nurgiantoro, 2001: 138) sebagai berikut:

$$IF = \frac{FH + FL}{N}$$

Keterangan :

IF = (*item facility*) indeks tingkat kesulitan yang dicari.

FH = (*frequency high*) jumlah jawaban betul kelompok tinggi

FL = (*frequency low*) jumlah jawaban betul kelompok rendah

N = Jumlah siswa kedua kelompok

Jadi, untuk mengetahui tingkat kesukaran suatu butir soal dapat diketahui dengan cara menjumlahkan jumlah jawaban benar kelompok tinggi dan kelompok rendah, kemudian dibagi dengan jumlah keseluruhan siswa.

b. Daya Pembeda

Menurut Oller (dalam Nurgiantoro, 2001: 140), daya pembeda (*item discrimination*) ialah kemampuan butir soal dalam membedakan siswa dengan kemampuan tinggi dan siswa kemampuan rendah. Indeks -1,00 hingga 1,00 digunakan untuk mengetahui besar kecilnya daya pembeda butir soal. Semakin besar suatu indeks atau dekat dengan 1,00 berarti soal tersebut semakin baik, karena berpeluang besar dapat melihat perbedaan kelompok tinggi dengan kelompok rendah. Apabila indeks bernilai negatif maka siswa dari kelompok rendah dapat menjawab dengan benar lebih banyak jumlahnya daripada kelompok tinggi.

Menurut Arikunto (2018: 235), daya pembeda soal ialah kemampuan soal untuk melihat perbedaan antara siswa yang berpengetahuan tinggi dan siswa yang berpengetahuan rendah. Angka

yang menggambarkan besarnya daya pembeda disebut indeks diskriminasi (D). Indeks diskriminasi (D) berkisar dari 0,00 hingga 1,00.

Arikunto (2018: 236) menjelaskan cara menentukan daya pembeda (Nilai D) sebagai berikut:

- 1) Untuk kelompok kecil: Semua peserta tes dibagi menjadi dua kelompok yang sama besar, yaitu 50% kelompok atas dan 50% kelompok bawah. Peserta tes diurutkan berdasarkan skor, dari yang tertinggi hingga terendah, lalu dibagi menjadi dua kelompok.
- 2) Untuk kelompok besar: Karena faktor biaya dan waktu untuk analisis, biasanya hanya diambil kedua kutubnya saja, yaitu 27% skor tertinggi sebagai kelompok atas (JA) dan 27% skor terendah sebagai kelompok bawah (JB).

J_A = Jumlah kelompok atas

J_B = Jumlah kelompok bawah

Rumus untuk mencari indeks diskriminasi (D) menurut Arikunto (2018: 238) sebagai berikut.

$$D = \frac{B_A}{J_A} - \frac{B_B}{J_B} = P_A - P_B$$

Keterangan :

J = Jumlah peserta tes

J_A = Banyaknya peserta kelompok atas

J_B = Banyaknya peserta kelompok bawah

B_A = Banyaknya peserta kelompok atas yang menjawab soal dengan benar

B_B = Banyaknya peserta kelompok bawah yang menjawab soal dengan benar

P_A = Proporsi peserta kelompok atas yang menjawab benar

P_B = Proporsi peserta kelompok bawah yang menjawab benar

Rumus untuk menghitung daya pembeda butir soal menurut Oller (Nurgiantoro, 2001: 140) sebagai berikut:

$$ID = \frac{FH - FL}{n}$$

Keterangan :

ID = (*item discrimination*) indeks daya beda yang dicari

FH = Jumlah jawaban betul kelompok tinggi

FL = Jumlah jawaban betul kelompok rendah

n = Jumlah subjek kelompok tinggi atau rendah

Jadi, untuk mengetahui daya pembeda pada butir soal dapat diketahui dengan cara mengurangi jumlah jawaban betul kelompok tinggi dengan jumlah jawaban betul kelompok rendah, kemudian dibagi jumlah subjek kelompok tinggi atau rendah.

c. Pengecoh (*Distractor*)

Distraktor adalah pola jawaban atau pilihan yang ada dalam soal pilihan ganda. Distraktor digunakan sebagai jawaban pengecoh untuk membuat peserta tes tertarik dalam memilihnya. Menurut Arikunto (2018: 244), pengecoh dianggap tidak efektif jika peserta

ujian tidak memilih pengecoh sama sekali. Nama pengecoh yang tidak pernah dipilih dalam penilaian disebut "Omit". Namun demikian, pengecoh dapat dianggap berhasil jika dapat menarik perhatian siswa yang tidak mengerti konsep atau materi yang diuji. Suatu distraktor dianggap berfungsi dengan efektif apabila minimal 5% siswa memilih pengecoh tersebut. Berikut contoh tabel kontingensi 2×5 sebaran pola pilihan jawaban.

Tabel 2.1 Kontingensi 2×5 sebaran pola pilihan jawaban

Kelompok/Pilihan	a*	b	c	d	om	Jumlah
Kelompok Atas	2	1	9	2	1	15
Kelompok Bawah	1	4	5	4	1	15
Jumlah	3	5	14	6	2	30

Berdasarkan tabel di atas dapat diketahui penjelasan sebagai berikut.

- 1) Pemilih jawaban a terdapat 3 orang, 2 orang kelompok atas (At) dan 1 orang kelompok bawah (Bw).
- 2) Pemilih jawaban b terdapat 5 orang, 1 orang kelompok atas (At) dan 4 orang kelompok bawah (Bw).
- 3) Pemilih jawaban c terdapat 14 orang, 9 orang kelompok atas (At) dan 5 orang kelompok bawah (Bw).
- 4) Pemilih jawaban d terdapat 6 orang, 2 orang kelompok atas (At) dan 4 orang kelompok bawah (Bw).
- 5) Yang tidak memilih (omit) terdapat 2 orang, masing-masing 1 orang dari kelompok atas dan kelompok bawah.

Dari hasil tabel di atas, dapat diketahui butir soal tersebut tidak baik, karena hampir semua dari jumlah siswa tersesat sehingga memilih opsi pilihan c. Opsi pilihan c memiliki daya tarik yang tinggi seolah-olah pilihan jawaban itulah yang benar. Pada kenyataannya pilihan jawaban yang tepat adalah pilihan a.

Menurut Quaigrain dan Arhin (2017) butir soal pilihan ganda dapat dikatakan berkualitas baik didasarkan pada ketersediaan distraktor yang baik pula karena mampu membedakan antara siswa yang berpengetahuan dan tidak berpengetahuan.

Adapun kriteria pengecoh (*Distractor*) menurut Arifin (2010: 280) adalah sebagai berikut.

- | | |
|-------------------------------|----------------|
| a. Lebih dari 200% | : Sangat Jelek |
| b. 0% - 25% atau 176% - 200% | : Jelek |
| c. 26% - 50% atau 151% - 175% | : Kurang Baik |
| d. 51% - 75% atau 126% - 150% | : Baik |
| e. 76% - 125% | : Sangat Baik |

4. Kajian Penelitian Yang Relevan

Penelitian-penelitian terdahulu yang relevan dengan penelitian ini terdiri dari beberapa penelitian. Yang pertama penelitian oleh Yanuar Bagas Arwansyah, Kodrat Eko Putro Setiawan, dan Ramadhan Kusuma Yuda pada tahun 2021 yang berjudul “Analisis Butir Soal Penilaian Akhir Tahun (PAT) Bahasa Indonesia Kelas XI SMA Negeri 1 Polanharjo Klaten”. Penelitian tersebut menyatakan bahwa soal PAT yang diteliti

memiliki validitas yang tinggi, reliabilitas yang tinggi, tingkat kesukaran rendah, daya pembeda baik, dan fungsi pengecoh sangat baik.

Kedua, penelitian oleh Vina Nur Indah Sari, Asep Purwo Yudi Utomo, dan Sumarwati pada tahun 2022 yang berjudul “Kualitas Soal Bahasa Indonesia di SMP Muhammadiyah 1 Pontianak: Analisis Butir Soal”. Hasil penelitian tersebut menyatakan bahwa tingkat kesukaran butir soal yang rendah, daya pembeda sedang, dan pengecoh tidak layak.

Ketiga, penelitian oleh Wirdatul Jannah, Elvrin Septyanti, dan Zulhafizh pada tahun 2023 yang berjudul “Analisis Kualitas Butir Soal Bahasa Indonesia SMP Amal Mulia Boarding School Pekanbaru pada Tes Sumatif Tahun Ajaran 2022/2023”. Hasil penelitian tersebut menyatakan bahwa kualitas butir soal yang diteliti sangat baik dan secara kualitatif kualitas soal PTS tidak berhubungan dengan hasil belajar siswa.

Keempat, penelitian oleh Ira Yoshita Cahyaningrum, Anies Fuady, dan Sunismi pada tahun 2023 yang berjudul “Analisis Butir Soal Sumatif Akhir Semester Ganjil Mata Pelajaran Matematika Kelas VII dengan Berbantuan Aplikasi Software Anates”. Hasil penelitian menyatakan bahwa butir soal yang diteliti memiliki reliabilitas yang cukup, daya pembeda cukup baik, dan pengecoh yang kurang efektif.

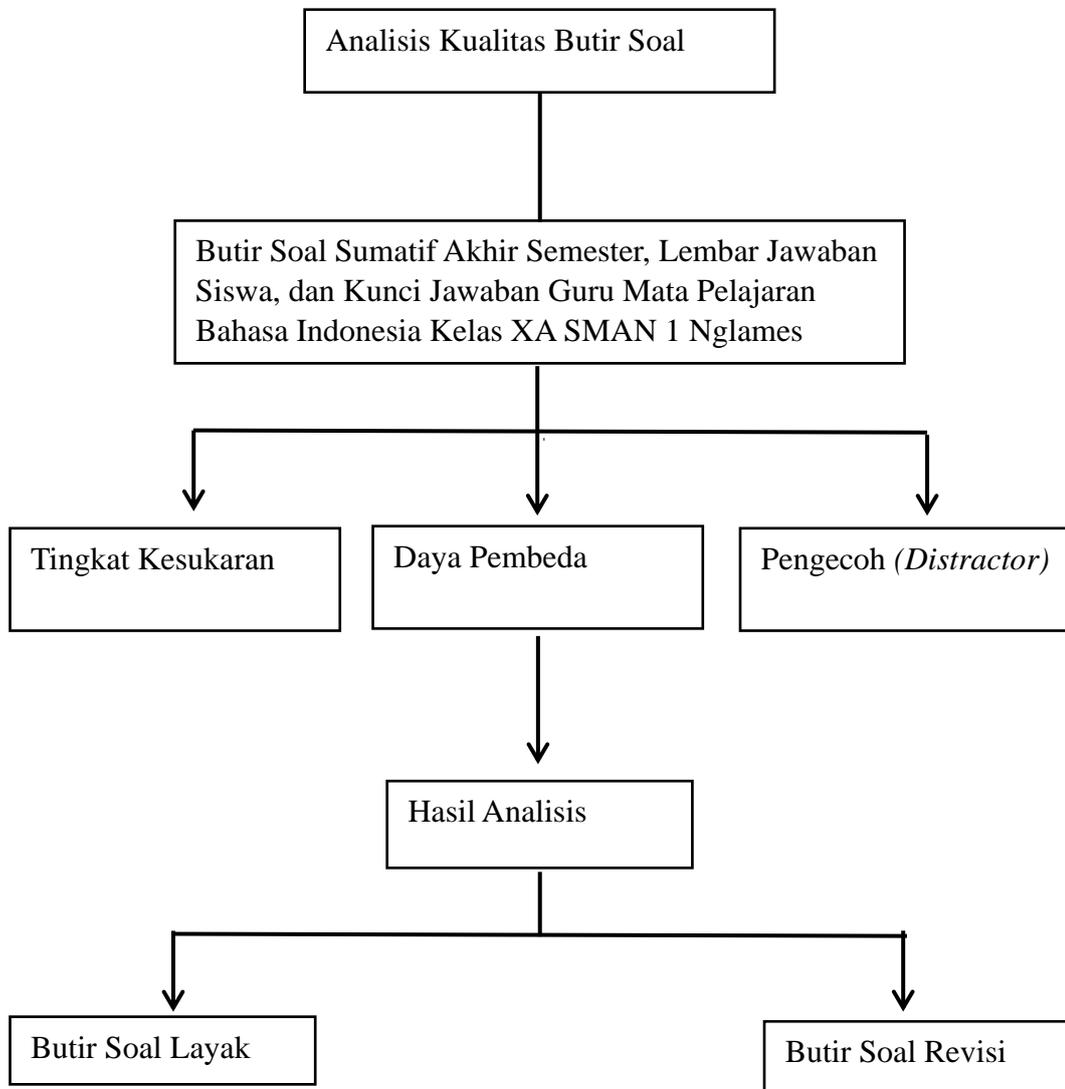
Kelima, penelitian oleh Ratni pada tahun 2023 yang berjudul “Analisis Tingkat Kesukaran, Daya Pembeda dan Efektivitas Pengecoh Soal Ulangan Bahasa Indonesia di Sekolah Dasar”. Ditunjukkan dari hasil

penelitian bahwa butir soal yang diteliti memiliki tingkat kesukaran rendah, daya pembeda cukup baik, dan diistraktor yang efektif.

B. Kerangka Berpikir

Kegiatan belajar mengajar tidak dapat dipisahkan dengan asesmen atau penilaian pembelajaran. Tujuan dilakukannya asesmen atau penilaian dalam pembelajaran adalah supaya guru mengerti apakah capaian pembelajaran telah tercapai. Salah satu bentuk dari asesmen atau penilaian adalah tes. Tes ialah alat yang berguna untuk menguji pengetahuan siswa dalam memahami suatu materi. Tes Sumatif Akhir Semester yang berbentuk pilihan ganda adalah tes yang dipakai pada penelitian ini.

Tes Sumatif Akhir Semester ialah tes yang dilakukan di akhir semester baik semester ganjil maupun semester genap. Tujuan tes Sumatif Akhir Semester ini adalah untuk mengetahui keterampilan siswa dalam memahami materi pembelajaran selama satu semester. Tes yang digunakan harus berkualitas baik, maka dari itu perlu dilakukan analisis butir soal terlebih dahulu. Setelah melakukan analisis butir soal, maka bisa diketahui tingkat kesukaran, daya pembeda, dan keefektifan diistraktor (pengecoh). Berdasarkan hasil analisis nanti akan diperoleh butir soal yang layak dan butir soal yang tidak layak.



Bagan 2.1 Kerangka Berpikir