

ADMINISTRASI, ANALISIS BUTIR, DAN KAIDAH PENULISAN TES

Khaerudin¹

khaerudin77@yahoo.com

Abstract

Test is a measuring tool used to detect the material which has not been well understood by the students. As a barometer, of course, like the "detector" should be able to detect which parts are healthy, and the portions are ill and require intensive treatment. For the test maker, the analysis is the empirical validity. The item test standard is the item which is equipped with its administration (SK, KD, indicators and number of items), qualitative analysis (the study of items test) up to quantitative analysis. In performing the analysis items, the item writer/maker could analyze qualitatively, in terms of content and form, and quantitative in terms of its statistical characteristics or procedures increases in judgment and procedure improvement empirically. Purpose of study is to assess and examine each item on the matter in order to obtain the grade/the good items before used. In addition, the purpose of items analysis as well to increase the test through revision or dispose of a test that is not effective, and to investigate the diagnostic information on students whether they have / have not understand the material being taught. The good item test is a test that can provide information as precisely as possible in accordance with the purpose of which learners/student can determine which one has or has not mastered/understood the material taught by teacher.

Keyword: Administration, Analysis and Rule Item Writing Tests.

¹ Mahasiswa Program Doktor Universitas Negeri Yogyakarta.

A. Pendahuluan

Pada umumnya hasil belajar dinilai melalui tes, baik tes uraian maupun tes objektif. Pelaksanaan penilaian bisa secara lisan, tulisan, dan tindakan atau perbuatan. Ada beberapa kaidah dalam menyusun soal-soal tes uraian dan tes objektif yang harus diperhatikan agar soal-soal tersebut memenuhi kualitas yang memadai sebagai alat penilaian hasil pembelajaran. Demikian pula pemeriksaan dan skoring hasil penialian melalui tes esai dan tes objektif dengan semua tipenya.

Untuk mengetahui tingkat pencapaian kompetensi, guru dapat melakukan penilaian melalui tes dan non tes. Tes meliputi tes lisan, tertulis (bentuk uraian, pilihan ganda, jawaban singkat, isian, menjodohkan, benar-salah), dan tes perbuatan yang meliputi: kinerja (*performance*), penugasan (projek) dan hasil karya (produk). Penilaian non-tes contohnya seperti penilaian sikap, minat, motivasi, penilaian diri, portfolio, *life skill*. Tes perbuatan dan penilaian non tes dilakukan melalui pengamatan (observasi).

Analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar dapat diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Ada dua jenis analisis butir soal, yakni analisis tingkat kesukaran soal dan daya beda. Menganalisis tingkat kesukaran soal artinya mengkaji soal-soal tes dari segi kesulitannya sehingga dapat diperoleh soal-soal mana yang termasuk mudah, sedang, dan sukar. Sedangkan menganalisis daya pembeda artinya mengkaji soal-soal tes dari segi kesanggupan tes tersebut dalam membedakan siswa yang termasuk ke dalam kategori lemah aatau rendah dan kategori kuat atau tinggi prestasinya.

Ada beberapa hal yang perlu diperhatikan dalam menulis butir instrumen, baik instrumen dalam bentuk skala sikap, skala penilaian, maupun tes diantaranya butir harus langsung mengukur indiaktor, jawaban terhadap butir instrumen dapat mengindikasikan ukuran indikator, butir dapat berbentuk pertanyaan ataupun pernyataan, opsi dari setiap pertanyaan atau pernyataan harus relevan, dan banyaknya opsi menunjukkan panjang skala yang secara konseptual kontinum.

Ada beberapa petunjuk yang perlu diperhatikan dalam membuat butir-butir soal, antara lain: 1) soal yang harus dibuat harus valid dalam arti mampu mengukur tercapai tidaknya TIK yang telah dirumuskan, 2) soal yang dibuat harus dapat dikerjakan dengan menggunakan satu kemampuan spesifik, tanpa dipengaruhi oleh kemampuan lain yang tidak relevan, 3) soal yang dibuat harus terlebih dahulu dikerjakan atau diselesaikan dengan langkah-langkah lengkap sebelum digunakan pada tes yang sesungguhnya, 4) hindari kesalahan ketik, karena hal itu dapat mempengaruhi validitas soal, 5) tetapkan sejak awal kemampuan yang hendak diukur untuk setiap soal, dan 6) berikan petunjuk cara mengerjakan soal secara jelas.

B. Pembahasan

1. Administrasi Tes

Tes merupakan salah satu bentuk instrumen yang digunakan untuk melakukan pengukuran. Tes terdiri atas sejumlah pertanyaan yang memiliki jawaban benar atau salah, atau semua benar atau sebagian benar. Tujuan melakukan tes adalah untuk mengetahui pencapaian belajar atau kompetensi yang telah dicapai peserta didik untuk bidang tertentu. Hasil tes merupakan informasi tentang karakteristik seseorang atau sekelompok orang.²

Tes adalah suatu prosedur yang sistematis untuk mengamati dan mendeskripsikan karakteristik seseorang dengan menggunakan skala numerik atau sistem kategori.³ Tes yang digunakan untuk mengukur prestasi belajar siswa, haruslah memenuhi syarat-syarat tes yang baik agar dapat berfungsi secara tepat dan akurat. Syarat-syarat tes yang baik menurut Suryabrata terdiri atas enam faktor yaitu reliabel, valid, objektif, diskriminatif, komprehensif, dan mudah digunakan.⁴ Tes hasil belajar adalah merupakan

² Djemari Mardapi, *Pengukuran, Penilaian & Evaluasi Pendidikan*, (Yogyakarta: Parama Publishing, 2016), hal. 97.

³ Lee J. Cronbach, *Essentials of Psychological Testing*, (New York: Harper and Row Publisher, 1984), hal. 26.

⁴ Sumadi Suryabrata, *Psikologi Pendidikan*, (Jakarta: PT Raja Grafindo Persada, 2006), hal. 303-306.

salah satu jenis tes yang digunakan untuk mengukur perkembangan atau kemajuan belajar peserta didik, setelah mereka mengikuti proses pembelajaran.⁵

Test sebenarnya adalah merupakan alat ukur yang digunakan untuk mendeteksi materi mana yang belum dikuasai dengan baik oleh siswa. Sebagai barometer, sudah barang tentu ibarat “detektor” harus bisa mendeteksi bagian-bagian mana yang sehat, dan bagian-bagian sakit dan memerlukan perawatan intensif/penyembuhan. Bagi pembuat soal, analisis tersebut merupakan validitas empiris. Karena soal yang standar, adalah soal yang dilengkapi dengan pengadministrasiannya (SK, KD, Indikator dan nomor item), analisis kualitatif (telaah soal) sampai dengan analisis kuantitatif. Analisis kualitatif, yaitu setelah soal ditulis, ditelaah apakah item-item soal yang telah dibuat sudah memenuhi syarat sesuai dengan aspek materi, aspek konstruksi dan aspek bahasa. Selanjutnya soal tersebut dianalisis berdasarkan data empiris atau hasil data uji coba di lapangan.

a. Bentuk Tes

Menurut bentuknya, tes dapat berbentuk tes esai dan tes objektif dalam berbagai variasi. Dalam hubungan ini, Popham dalam I Wayan Koyan menyatakan bahwa bentuk tes tertulis dibedakan menjadi dua golongan besar, yaitu: (1) soal-soal jawaban memilih (*selected-response tests*), yang terdiri dari butir soal pilihan benar-salah (*true-false items*), butir soal pilihan ganda (*multiple-choice items*), dan butir soal menjodohkan (*matching items*); dan (2) soal-soal jawaban tersusun atau terstruktur (*constructed-response tests*), yang terdiri dari butir soal jawaban singkat (*short-answer items*), dan butir soal esai (*essay items*). Sejalan dengan pendapat ini, Wiersma dan Jurs dalam I Wayan Koyan menyatakan bahwa terdapat dua bentuk utama butir tes, yang secara umum disebut tes objektif dan esai, yang masing-masing memiliki format yang bervariasi. Selanjutnya dinyatakan bahwa istilah butir tes objektif secara umum berhubungan dengan butir jawaban pilihan (*selected-*

⁵ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: RajaGrafindo, 2009), hal. 99.

response items). Sedangkan butir tes esai adalah salah satu bentuk dari butir jawaban tersusun (*constructed-response items*).⁶

Bentuk tes yang digunakan di satuan pendidikan dapat dikategorikan menjadi dua, yaitu tes objektif dan tes non objektif. Tes non objektif juga sering disebut dengan tes bentuk esai atau uraian. Tes obyektif di sini dilihat dari cara penskorannya, siapa saja yang memeriksa lembar jawaban akan menghasilkan skor yang sama. Tes yang non objektif adalah yang cara penskorannya dipengaruhi oleh pemberi skor. Dengan kata lain dapat dikatakan bahwa tes yang objektif adalah sistem penskorannya objektif, sedangkan tes yang non objektif sistem penskorannya dipengaruhi oleh subjektivitas pemberi skor.⁷

Bentuk tes objektif yang sering digunakan adalah bentuk pilihan ganda, benar salah, menjodohkan, dan uraian objektif. Tes uraian dapat dibedakan uraian objektif dan uraian non objektif. Tes uraian objektif sering digunakan pada bidang sains dan teknologi atau bidang sosial yang jawaban soalnya sudah pasti, dan hanya satu jawaban yang benar. Tes uraian non objektif sering digunakan pada bidang-bidang ilmu sosial, yaitu yang jawabannya luas dan tidak haanya satu jawaban yang benar, tergantung argumentasi peserta tes.⁸

Ada beberapa bentuk tes yang sering digunakan dalam penilaian tes hasil belajar baik tes objektif maupu tes non objektif. Butir tes objektif erat kaitannya dengan dengan butir jawaban pilihan sedangkan tes non objektif (tes esai) salah satu bentuk butir jawaban tersusun. Pemberian skor tes objektif dengan sistem penskorannya ojektif, sedangkan tes yang non objektif sistem penskorannya dipengaruhi oleh subjektivitas pemberi skor.

⁶ I Wayan Koyan, *Konstruksi Tes*, (Singaraja: Universitas Pendidikan Ganesha, 2012), hal. 14.

⁷ Djemari Mardapi, *Op. Cit.*, hal. 98.

⁸ Djemari Mardapi, *Ibid.*

b. Pelaksanakan Tes

Dalam praktek, pelaksanaan tes hasil belajar dapat diselenggarakan secara tertulis (tes tertulis), dengan secara lisan (tes lisan) dan dengan tes perbuatan.

1) Teknik Pelaksanaan Tes Tertulis

Dalam melaksanakan tes tertulis ada beberapa hal yang perlu mendapat perhatian, yaitu sebagaimana dikemukakan berikut ini.

- a) Pertama, agar dalam mengerjakan soal tes para peserta tes mendapat ketenangan, seyogyanya ruang tempat berlangsungnya tes dipikirkan yang jauh dari keramaian, kebisingan, suara hiruk pikuk dan lalu lalang orang.
- b) Kedua, ruangan tes harus cukup longgar, tidak berdesak-desakan, tempat duduk diatur dengan jarak tertentu yang memungkinkan tercegahnya kerja sama yang tidak sehat di antara *testee*.
- c) Ketiga, ruangan tes sebaiknya memiliki system pencahayaan dan pertukaran udara yang baik.
- d) Keempat, jika dalam ruangan tes tidak tersedia meja tulis atau kursi yang memiliki alas tempat penulis, maka sebelum tes dilaksanakan hendaknya sudah disiapkan alat berupa alat tulis yang terbuat dari *triplex, hardboard* atau bahan lainnya.
- e) Kelima, agar *testee* dapat memulai mengerjakan soal tes secara bersamaan, hendaknya lembar soal-soal tes diletakkan secara terbalik.
- f) Keenam, dalam mengawasi jalannya tes, pengawas hendaknya berlaku wajar.
- g) Ketujuh, sebelum berlangsungnya tes, hendaknya sudah ditentukan lebih dahulu sanksi yang dapat dikenakan kepada *testee* yang berbuat curang.
- h) Kedelapan, sebagai bukti mengikuti tes, harus disiapkan daftar hadir yang harus ditandatangani oleh seluruh peserta tes.

- i) Kesembilan, jika waktu yang ditentukan telah habis, hendaknya testee diminta untuk menghentikan pekerjaannya dan secepatnya meninggalkan ruangan tes.
- j) Kesepuluh, untuk mencegah timbulnya berbagai kesulitan di kemudian hari, pada Berita Acara Pelaksanaan Tes harus dituliskan secara lengkap, berapa orang *testee* yang hadir dan siapa yang tidak hadir, dengan menuliskan identitasnya (nomor urut, nomor induk, nomor ujian, nama dan sebagainya), dan apabila terjadi penyimpangan-penyimpangan atau kelainan-kelainan harus dicatat dalam berita acara pelaksanaan tes tersebut.

2) Teknik Pelaksanaan Tes Lisan

Tes lisan adalah tes dimana *tester* di dalam mengajukan pertanyaan-pertanyaan atau soalnya dilakukan secara lisan, dan testee memberikan jawabannya juga secara lisan pula.⁹ Dari segi persiapan dan cara bertanya, tes lisan dapat dibedakan menjadi dua yakni tes lisan bebas, tes tanpa menggunakan pedoman yang dipersiapkan secara tertulis dan tes lisan pedoman, tentang apa yang ditanyakan kepada peserta didik.¹⁰

Pertanyaan lisan dapat digunakan untuk mengetahui daya serap peserta didik untuk masalah yang berkaitan dengan kognitif yang baru diajarkan. Pertanyaan bisa diajukan di awal pembelajaran, yaitu mengenai konsep atau aplikasi pelajaran yang lalu. Pertanyaan lisan yang diajukan ke kelas harus jelas, dan semua peserta didik harus diberi kesempatan yang sama.

Dalam melakukan pertanyaan dikelas prinsipnya adalah mengajukan pertanyaan, memberi waktu untuk berpikir, kemudian menunjuk peserta didik untuk menjawab pertanyaan. Benar atau salah

⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan.*, (Jakarta: Raja Grafindo Persada, 2003), hal. 75.

¹⁰ M. Chabib Toha, *Teknik Evaluasi Pendidikan*, (Jakarta: Raja Grafindo Persada, 1996), hal. 61.

jawaban peserta didik, sebaiknya jawaban tersebut ditawarkan lagi ke kelas untuk mengaktifkan kelas. Tingkat berpikir untuk pertanyaan lisan di kelas bisa rendah sampai tinggi. Pertanyaan lisan memiliki kebaikan, yaitu melatih peserta didik dalam berkomunikasi secara lisa.¹¹

3) Teknik Pelaksanaan Tes Perbuatan

Tes perbuatan pada umumnya digunakan untuk mengukur taraf kompetensi yang bersifat keterampilan (psiko-motorik), dimana penilaiannya dilakukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai oleh peserta *testee* setelah melaksanakan tugas tersebut.

Dalam melaksanakan tes perbuatan itu, ada beberapa hal yang perlu diperhatikan oleh *tester*: a) Pertama, tester harus mengamati dengan secara teliti, cara yang ditempuh oleh *testee* dalam menyelesaikan tugas yang telah ditentukan., b) Kedua, agar dapat dicapai kadar obyektivitas setinggi mungkin, hendaknya tester jangan berbicara atau berbuat sesuatu yang dapat mempengaruhi *testee* yang sedang mengerjakan tugas tersebut., 3) Ketiga, dalam mengamati *testee* yang sedang melaksanakan tugas itu, hendaknya tester telah menyiapkan instrumen berupa lembar penilaian yang didalamnya telah ditentukan hal-hal apa sajakah yang harus diamati dan diberikan penilaian.¹²

Tes tindakan adalah “tes yang persoalan atau pertanyaan disampaikan dalam bentuk tugas yang harus dikerjakan oleh peserta didik. Alat yang dapat digunakan tes ini adalah berupa observasi atau pengamatan terhadap tingkah laku tersebut, yang hasilnya kemudian diserahkan pada guru.¹³

¹¹ Djemari Mardapi, *Op.Cit.*, hal. 104-105.

¹² Anas Sudijono, *Ibid.*, hal. 151-157.

¹³ W.S. Winkel S.J., *Psikologi Pendidikan dan Evaluasi Belajar*, (Jakarta: Gramedia, 1983), hal. 106.

c. Skor Tes

Pada hakikatnya, penskoran adalah suatu proses pengubahan jawaban dari instrumen menjadi angka-angka yang merupakan nilai kuantitatif dari suatu jawaban terhadap item dalam instrumen. Jadi penskoran merupakan kuantifikasi terhadap jawaban instrumen. Dengan memberikan skor, dapat diperoleh deskripsi tentang seberapa nilai atau harga suatu variabel untuk masing-masing unit analisis.¹⁴

Skoring atau pemberian skor terhadap jawaban yang benar dalam tes objektif, khususnya untuk jenis benar-salah dan pilhan berganda, menggunakan aturan sebagai berikut:

$$S_k = B - \frac{S}{O - 1}$$

S_k adalah skor yang diperoleh

B adalah jawaban yang benar

S adalah jawaban yang salah

O adalah kemungkinan jawaban atau *option*

Untuk jenis benar-salah kemungkinan jawaban atau *option* hanya dua, yakni benar atau salah, sehingga rumusnya bisa disederhanakan menjadi:

$$S_k = B - S$$

Sedangkan dalam melengkapi dan menjodohkan hanya dihitung jawaban yang benar, dengan demikian, rumusnya:

$$S_k = B$$

Setiap jawaban yang bisa dinilai atau diberi skor satu atau lebih, tergantung pada keinginan guru, namun pada umumnya diberi skor satu.¹⁵

Berikut ini adalah beberapa saran umum untuk menghemat waktu skoring dan meningkatkan akurasi penskoran dan konsistensi.

¹⁴ Djaali dan Pudji Muljono, *Pengukuran dalam Bidang Pendidikan*, (Jakarta: Gramedia Widiasarana Indonesia, 2008), hal. 101.

¹⁵ Nana Sudjana, *Penilaian Hasil Proses Belajar Mengajar*, (Bandung: Remaja Rosdakarya, 2010), Cetakan Kelima belas, hal. 54.

- 1) Siapkan sebuah Kunci Jawaban Siapkan kunci jawaban Anda di muka, yang akan menghemat waktu ketika Anda mencetak tes dan akan membantu Anda mengidentifikasi pertanyaan yang perlu dihilangkan. Juga, ketika membangun kunci jawaban, Anda harus mendapatkan ide dari berapa lama itu akan mengambil siswa Anda untuk menyelesaikan tes dan apakah kali ini sesuai untuk engkau slot waktu Anda telah dialokasikan untuk tes.
- 2) Periksa Kunci Jawaban Jika memungkinkan, memiliki seorang rekan memeriksa kunci jawaban Anda untuk mengidentifikasi alternatif jawaban atau masalah potensial.
- 3) Skor membabi buta Cobalah untuk mencetak "membabi buta." Artinya, mencoba untuk menjaga nama siswa dari pandangan untuk mencegah pengetahuan Anda tentang, atau harapan dari skor, siswa dari yang mempengaruhi digunakan.
- 4) Periksa Mesin-Mencetak Jawaban Lembar bertujuan jika digunakan untuk memeriksa setiap lembar jawaban untuk tanda liar, beberapa jawaban, atau tanda yang mungkin terlalu light untuk dijemput oleh mesin gol.
- 5) Periksa skoring Jika mungkin. periksa skoring Anda. kesalahan Scoring dihasilkan dari kesalahan administrasi sering terjadi. Tidak ada alasan untuk berharap bahwa Anda tidak akan membuat kesalahan tersebut.
- 6) Merekam Skor Sebelum kembali kertas-kertas skor kepada siswa, pastikan Anda memiliki mencatat skor mereka dalam buku catatan Anda! (Lupa untuk melakukan hal ini mungkin telah terjadi di setidaknya sekali untuk setiap guru).¹⁶

Pedoman penskoran sangat diperlukan, terutama untuk soal bentuk uraian, agar subjektifitas korektor dapat diperkecil. Pedoman penskoran

¹⁶ Tom Kubiszyn and Gary Borich, *Educational Testing And Measurement: Classroom Application and Practice*, (Unites States of America: Jhon Wiley & Sons. Inc, 2003), 7th.ed, hal. 196-197.

ini merupakan petunjuk yang menjelaskan tentang: batasan atau kata-kata kunci untuk melakukan penskoran terhadap soal bentuk uraian, dan kriteria jawaban yang digunakan untuk melakukan penskoran pada soal bentuk uraian bentuk non-objektif.

2. Analisis Tes

Analisis soal dilakukan untuk mengetahui berfungsi tidaknya sebuah soal. Analisis pada umumnya dilakukan melalui dua cara yaitu analisis kualitatif (*qualitatif control*) dan analisis kuantitatif (*quantitatif control*). Analisis kualitatif sering sebagai validitas logis (*logical validity*) yang dilakukan sebelum soal digunakan untuk melihat berfungsi tidaknya sebuah soal. Analisis soal empiris (*empirical validity*) yang dilakukan untuk melihat lebih berfungsi tidaknya sebuah soal, setelah soal itu diujicobakan kepada sampel yang representatif.

Sama seperti Anda dapat mengharapkan untuk membuat kesalahan penilaian, Anda dapat mengharapkan untuk membuat kesalahan dalam konstruksi tes. Tidak ada tes yang membangun akan menjadi sempurna itu akan mencakup tidak pantas, tidak sah, atau item dinyatakan kekurangan. Dalam sisa bab ini kami akan memperkenalkan Anda ke analisis item teknik yang disebut. analisis item dapat digunakan untuk mengidentifikasi item yang kekurangan beberapa cara. sehingga membuka jalan untuk memperbaiki atau menghilangkan mereka, dengan hasil yang lebih baik tes secara keseluruhan. Kami akan membuat perbedaan antara dua jenis analisis item, kuantitatif dan kualitatif. analisis item kuantitatif mungkin menjadi sesuatu yang baru. Tetapi karena Anda akan melihat, analisis item kualitatif adalah sesuatu yang Anda sudah akrab. Akhirnya, kita akan membahas bagaimana analisis item berbeda untuk norma dan kriteria direferensikan tes, dan kami akan memberikan beberapa norma dirujuk metode analisis dimodifikasi untuk digunakan dengan tes kriteria acuan.¹⁷

¹⁷ Tom Kubiszyn and Gary Borich, *Ibid.*, hal. 197.

Kriteria yang digunakan untuk melakukan telaah butir tes mengikuti pedoman penyusunan tes. Telaah dilakukan terhadap kebenaran konsep, teknik penulisan, dan bahasa yang digunakan.

a. Analisis Butir Kualitatif

Analisis kualitatif sering juga disebut sebagai validitas logis (*logical validity*) yaitu berupa penelaahan yang dimaksudkan untuk menganalisis soal ditinjau dari teknis, isi, dan editorial. Analisis secara teknis dimaksudkan sebagai penelaahan soal berdasarkan prinsip-prinsip pengukuran dan format penulisan soal. Analisis secara isi dimaksudkan sebagai penelaahan khusus yang berkaitan dengan kelayakan pengetahuan yang ditanyakan. Analisis secara editorial dimaksudkan sebagai penelaahan yang khususnya berkaitan dengan keseluruhan format dan keajegan editorial dari soal yang satu ke soal yang lainnya.¹⁸

Analisis kualitatif lainnya dapat juga dikategorikan dari segi materi, konstruksi, dan bahasa. Analisis materi dimaksudkan sebagai penelaahan yang berkaitan dengan substansi keilmuan yang ditanyakan dalam soal serta tingkat kemampuan yang sesuai dengan soal. Analisis konstruksi dimaksudkan sebagai penelaahan yang umumnya berkaitan dengan teknik penulisan soal. Analisis bahasa dimaksudkan sebagai penelaahan soal yang berkaitan dengan penggunaan bahasa Indonesia yang baik dan benar menurut EYD. Melalui analisis kualitatif dapat diketahui berfungsi tidaknya sebuah soal.¹⁹

Telaah item tes atau analisis kualitatif sering disebut analisis teoretik. Item tes yang telah ditulis diperiksa kesesuaiannya dengan kisi-kisi yang diacunya dengan memperhatikan substansi/isi materi, konstruksi, dan bahasa. Telaah item tes dilakukan oleh: 1) bukan oleh penulis item tes, dan 2) dilakukan oleh pakar yang menguasai isi/materi yang diujikan. Dari

¹⁸ Sumarna Surapranata, *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes: Implementasi Kurikulum 2014*, (Jakarta: Remaja Rosdakarya, 2005), Cetakan kedua, hal. 1.

¹⁹ Sumarna Surapranata, *Ibid.*, hal. 2.

kegiatan tersebut, penulis item tes dapat mengetahui validitas isi dari item tes yang disusun.²⁰

Pada analisis kualitatif lebih teknis untuk penelaahan soal berdasarkan prinsip-prinsip pengukuran dan format penulisan soal baik dari segi materi, konstruksi maupun bahasa. Dalam penulisan item perlu diperiksa agar sesuai dengan kisi-kisi. Sehingga validitas item yang sesuai dengan tersebut bisa dipertanggungjawabkan dalam pelaksanaan tes .

b. Analisis Butir Kuantitatif

Analisis soal secara kuantitatif menekankan pada analisis karakteristik internal tes melalui data yang diperoleh secara empiris. Karakteristik internal secara kauntitatif dimaksudkan meliputi parameter soal tingkat kesukaran, daya beda, dan reliabilitas. Khusus soal-soal pilihan ganda, dua tambahan parameter yaitu dilihat dari peluang untuk menebak atau menjawab soal benar berfungsi tidaknya pilihan jawaban, yaitu penyebaran semua alternatif jawaban dari subyek-subyek yang dites.²¹ .

Validitas empiris merupakan analisis kuantitatif, yang meliputi antara lain tingkat kesulitan/kesukaran (TK) item, daya beda (DB) item, daya pengecoh (DP) dan reliabilitas soal, sesuai dengan kaidah prinsip penilaian. Dengan demikian diketahui kualitas soal tersebut. Dari data empiris ini, item-item soal tersebut diketahui termasuk kriteria diterima, direvisi atau ditolak. Item-item soal yang termasuk kriteria baik dimasukkan dalam bank soal.

Sebagai pendeteksi tentunya harus bisa berfungsi dengan baik. Bagaimana mungkin kesimpulan yang dibuat bisa benar, sedangkan alat ukurnya belum sesuai dengan standar. Dan masih sering dijumpai, kita masih salah dalam membedakan antara kumpulan soal dan bank soal. Bank soal, merupakan kumpulan soal-soal yang telah teruji baik validitas kualitatif dan validitas empiris dan dilengkapi dengan administrasi manual

²⁰ Bambang Subali, *Analisis Soal Baik Kualitatif Maupun Kuantitatif*, Kegiatan Workshop *Item Development* Dosen Poltekes Kebidanan Politeknik Kesehatan Surakarta, t.t.

²¹ Sumarna Surapranata, *Ibid.*, hal. 10.

soal tersebut (SK, KD, Indikator, nomor soal dan skor). Sedangkan kumpulan soal, hanya sebatas soal-soal yang dikumpulkan tanpa dilengkapi dengan dokumen seperti pada bank soal.

1) Tingkat Kesukaran

Bermutu atau tidaknya butir-butir item tes hasil belajar pertamanya dapat diketahui dari derajat kesukaran atau taraf kesulitan yang dimiliki oleh masing-masing butir item tersebut. Butir-butir item tes hasil belajar dapat dinyatakan sebagai butir-butir item yang baik, apabila butir-butir item tersebut tidak terlalu sukar dan pula tidak terlalu mudah dengan kata lain derajat kesukaran item itu adalah sedang atau cukup.²²

Selain membantu kita menentukan item untuk menghilangkan dari tes sebelum itu lagi diberikan, analisis item kuantitatif juga memungkinkan kita untuk membuat keputusan lain. Untuk contoh, kita dapat menggunakan analisis item kuantitatif untuk memutuskan apakah item miskeyed, apakah tanggapan untuk item ditandai dengan menebak, atau apakah item tersebut ambigu. Untuk melakukannya, kita hanya perlu mempertimbangkan tanggapan siswa di atas setengah dari kelas.

Tingkat kesukaran soal adalah peluang untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasanya dinyatakan dalam bentuk indeks. Secara klasik indeks tingkat kesukaran ini dinyatakan dalam bentuk proporsi yang besarnya berkisar 0,00 - 1,00.²³ Untuk soal pilihan ganda indeks tingkat kesukaran dihitung dengan rumus

$$\text{Tingkat Kesukaran (TK)} = \frac{\text{Jadwal siswa yang menjawab benar butir soal}}{\text{Jumlah siswa yang mengikuti tes}}$$

²² Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Grafindo, 2009), hal. 370.

²³ Lewis R. Aiken, *Psychological Testing and Assessment*, Eight Edition, (Boston: Allyn and Bacon, 1994), hal. 66.

Untuk mengetahui tingkat kesukaran butir soal bentuk uraian digunakan rumus sebagai berikut:

$$\text{Mean} = \frac{\text{Jadwal skor siswa peserta tes pada suatu soal}}{\text{Jumlah siswa yang mengikuti tes}}$$

$$\text{Tingkat Kesukaran (TK)} = \frac{\text{Mean}}{\text{Skor maksimum yang ditetapkan}}$$

Ada beberapa besar pertimbangan dalam menentukan proporsi jumlah soal kategori mudah, sedang, dan sukar. Pertimbangan pertama adalah adanya keseimbangan, yakni jumlah soal sama untuk kategori tersebut. Artinya, soal mudah, sedang, dan sukar jumlahnya seimbang. Perbandingan antara soal mudah-sedang-sukar bisa dibuat 3-4-3. Artinya, 30% soal kategori mudah, 40% soal kategori sedang, dan 30% soal kategori sukar. Misalnya 60 pertanyaan pilihan ganda terdapat 18 soal kategori mudah, 24 soal kategori sedang, dan 18 soal kategori sukar. Perbandingan lain yang termasuk sejenis dengan proporsi di atas misalnya 3-5-2. Artinya 30% soal kategori mudah, 50% soal kategori sedang, 20% soal kategori sukar.²⁴

Tingkat kesukaran item soal selain dapat digunakan untuk memprediksi alat ukur itu sendiri (soal), juga tentang kemampuan peserta didik dalam memahami materi yang diajarkan guru. Misalnya satu item soal termasuk kategori mudah, maka prediksi terhadap informasi ini adalah seperti berikut:

- 1) Pengecoh item soal itu tidak berfungsi.
- 2) Sebagian besar siswa menjawab benar item soal itu; artinya bahwa sebagian besar siswa telah memahami materi yang ditanyakan.

Bila suatu item soal termasuk kategori sukar, maka prediksi terhadap informasi ini adalah seperti berikut.

- a) Item soal itu "mungkin" salah kunci jawaban.

²⁴ Nana Sudjana, Penilaian Hasil Proses Belajar Mengajar, (Bandung: Remaja Rosdakarya, 2010), Cetakan Kelima belas, hal. 135-136.

- b) Item soal itu mempunyai 2 atau lebih jawaban yang benar.
- c) Materi yang ditanyakan belum diajarkan atau belum tuntas pembelajarannya, sehingga kompetensi minimum yang harus dikuasai siswa belum tercapai.
- d) Materi yang diukur tidak cocok ditanyakan dengan menggunakan bentuk soal yang diberikan.
- e) Pernyataan atau kalimat soal terlalu kompleks dan panjang.

Klasifikasi tingkat kesukaran soal (Puspendik) dapat dicontohkan seperti berikut²⁵:

- 0,00 - 0,30 soal tergolong sukar
- 0,31 - 0,70 soal tergolong sedang
- 0,71 - 1,00 soal tergolong mudah

Untuk pemilihan butir soal

Kriteria	Koefisien	Keputusan
Tingkat Kesukaran	0,30 s/d 0,70	Diterima
	0,10 s/d 0,27	Direvisi
	atau 0,71 s/d 0,90	
	< 0,10 dan > 0,90	Ditolak

2) Daya Beda

Daya pembeda soal adalah kemampuan suatu item soal dapat membedakan antara siswa yang telah menguasai materi yang ditanyakan dan siswa yang tidak/kurang/belum menguasai materi yang

²⁵ Rahman Zulaiha, *Analisis Soal Secara Manual*, (Jakarta: Departemen Pendidikan Nasional Badan Pusat Penelitian dan Pengembangan Pusat Penilaian Pendidikan, 2008), hal. 14.

ditanyakan. Atau dengan kata lain, merupakan indeks perbedaan antara kelompok berkemampuan tinggi dengan berkemampuan rendah.

Daya beda item soal bentuk pilihan ganda adalah dengan menggunakan rumus berikut:

$$DP = \frac{BA - BB}{\frac{1}{2}N} \quad \text{atau} \quad DP = \frac{2(BA - BB)}{N}$$

DP = daya pembeda soal,

BA = jumlah jawaban benar pada kelompok atas,

BB = jumlah jawaban benar pada kelompok bawah,

N = jumlah siswa yang mengerjakan tes.

Di samping itu, dapat dipergunakan rumus korelasi point biserial (r_{pbis}) dan korelasi biserial (r_{bis}) seperti berikut:

$$r_{pbis} = \frac{\bar{X}_b - \bar{X}_s}{SD} \sqrt{pq} \quad \text{dan} \quad r_{bis} = \frac{\bar{Y}_b - \bar{Y}_s}{SD} \cdot \frac{nb \cdot ns}{un\sqrt{n^2 - n}}$$

\bar{X}_b , \bar{Y}_b adalah rata-rata skor warga belajar/siswa yang menjawab benar

\bar{X}_s , \bar{Y}_s adalah rata-rata skor warga belajar siswa yang menjawab salah

SDt adalah simpangan baku skor total

nb dan ns, adalah jumlah siswa yang menjawab benar dan jumlah siswa yang menjawab salah, serta $nb + ns = n$.

p adalah proporsi jawaban benar terhadap semua jawaban siswa

q adalah $1 - p$

U adalah ordinat kurva normal.

Indeks daya beda setiap item soal biasanya juga dinyatakan dalam bentuk proporsi. Semakin tinggi indeks daya pembeda soal berarti semakin mampu soal yang bersangkutan membedakan siswa yang telah memahami materi dengan siswa yang belum memahami materi.

Indeks daya pembeda berkisar antara -1,00 sampai dengan +1,00. Semakin tinggi daya pembeda suatu soal, maka semakin kuat/baik soal itu. Jika daya pembeda negatif (-) atau mempunyai indeks kurang dari nol (< 0), artinya lebih banyak kelompok bawah (siswa yang tidak memahami materi) menjawab benar soal dibanding dengan kelompok atas (siswa yang memahami materi yang diajarkan guru). Atau dengan kata lain item soal tidak atau kurang berfungsi.²⁶

Adapun kriteria indeks daya beda menurut Fernandes dalam I Wayan Koyan adalah seperti berikut ini:

- 0,00 – 0,19 = kurang baik
- 0,20 – 0,39 = cukup baik
- 0,40 – 0,70 = baik
- 0,71 – 1,00 = sangat baik

Jika “D” negatif, soal tersebut sangat buruk dan harus dibuang. Tes yang baik, apabila memiliki D antara 0,15 – 0,20 atau lebih.²⁷

Manfaat daya pembeda item soal adalah seperti berikut ini.

- a) Untuk meningkatkan mutu setiap item soal melalui data empiriknya. Berdasarkan indeks daya pembeda, setiap item soal dapat diketahui apakah item soal itu baik, direvisi, atau ditolak.
- b) Untuk mengetahui seberapa jauh setiap item soal dapat mendeteksi/membedakan kemampuan siswa, yaitu siswa yang telah memahami atau belum memahami materi yang diajarkan guru. Apabila suatu item soal tidak dapat membedakan kedua kemampuan siswa itu, maka item soal itu dapat dicurigai "kemungkinannya" seperti berikut ini:
 - (1) Kunci jawaban item soal itu tidak tepat.

²⁶ Gene V. Glass and Julian C. Stanley, *Statistical Methods in Education and Psychology*, (New Jersey: Prentice Hall, Inc., 1970), hal. 169-170.

²⁷ I Wayan Koyan, *Op. Cit.*, hal. 63-64.

- (2) Item soal itu memiliki 2 atau lebih kunci jawaban yang benar
- (3) Kompetensi yang diukur tidak jelas
- (4) Pengecoh tidak berfungsi
- (5) Materi yang ditanyakan terlalu sulit, sehingga banyak siswa yang menebak
- (6) Sebagian besar siswa yang memahami materi yang ditanyakan berpikir ada yang salah informasi dalam item soalnya

3) Daya Pengecoh (Penyebaran/Distribusi Jawaban).

Penyebaran pilihan jawaban dijadikan dasar dalam penelaahan soal. Hal ini dimaksudkan untuk mengetahui berfungsi tidaknya jawaban yang tersedia. Suatu pilihan jawaban (pengecoh) dapat dikatakan berfungsi apabila pengecoh:

- a) paling tidak dipilih oleh 5 % peserta tes/siswa,
- b) lebih banyak dipilih oleh kelompok siswa yang belum paham materi.

3. Kaidah Penulisan Tes

Ada beberapa prinsip dasar yang perlu dicermati didalam menyusun tes hasil belajar agar tes tersebut dapat mengukur tujuan instruksional khusus untuk mata pelajaran yang telah diajarkan, atau mengukur kemampuan dalam keterampilan peserta didik yang diharapkan, setelah mereka menyelesaikan suatu unit pengajaran tertentu.

Pertama, tes hasil belajar harus dapat mengukur secara jelas hasil belajar (*learning outcomes*) yang telah ditetapkan sesuai dengan tujuan instruksional. Kejelasan mengenai pengukuran hasil belajar yang dikehendaki akan memudahkan guru dalam menyusun butir-butir soal tes hasil belajar.

Kedua, butir-butir tes hasil belajar harus merupakan sampel yang representatif dari populasi bahan pelajaran yang telah diajarkan, sehingga dapat dianggap mewakili seluruh performance yang telah diperoleh selama peserta didik mengikuti suatu unit pengajaran.

Ketiga, bentuk soal yang dikeluarkan dalam tes hasil belajar harus bervariasi, sehingga betul-betul cocok untuk mengukur hasil belajaryang

diinginkan sesuai dengan tujuan tes itu sendiri. Untuk mengukur hasil belajar yang berupa keterampilan misalnya, tidak tepat kalau menggunakan soal-soal yang berbentuk *essay test* yang jawabannya hanya menguraikan dan bukan melakukan atau mempraktekan sesuatu. Demikian pula untuk mengukur kemampuan menganalisis suatu prinsip, tidak cocok jika digunakan butir-butir soal yang berbentuk *objective test* yang pada dasarnya hanya mengungkap daya ingat peserta didik.

Keempat, tes hasil belajar harus didesain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan. Pernyataan tersebut mengandung makna, bahwa desain tes hasil belajar harus disusun relevan dengan kegunaan yang dimiliki oleh masing-masing jenis tes. Desain dari *placement test* (yaitu tes yang digunakan penentuan penempatan siswa dalam suatu jenjang atau jenis program pendidikan tertentu) sudah barang tentu akan berbeda dengan desain dari *formative test* (yaitu tes yang digunakan untuk mencari umpan balik guna memperbaiki proses pembelajaran, baik bagi guru maupun bagi siswa) dan *summative test* (yaitu test yang digunakan untuk mengukur atau menilai sampai dimana pencapaian siswa terhadap bahan pelajaran yang telah diajarkan dan selanjutnya untuk menentukan kenaikan tingkat atau keululusan siswa yang bersangkutan). Demikian pula desain dari *diagnostic test* (yaitu tes yang digunakan untuk mencari sebab-sebab kesulitan belajar siswa, seperti latar belakang psikologis, fisik dan lingkungan sosial ekonomi siswa) tentu akan berbeda pula tiga jenis tes yang telah disebutkan di atas.

Kelima, hasil belajar harus memiliki reliabilitas yang dapat diandalkan. Artinya setelah tes hasil belajar itu dilaksanakan berkali-kali terhadap suyek yang sama. Dengan demikian tes hasil belajar itu hendaknya memiliki keajegan hasil pengukuran yang tidak diragukan lagi.

Keenam, tes hasil belajar disamping harus dapat dijadikan alat pengukur keberhasilan belajar siswa, juga harus dapat dijadikan alat untuk mencari

informasi yang berguna untuk memperbaiki cara belajar siswa dan cara mengajar guru sendiri.²⁸

Ada beberapa langkah yang harus ditempuh dalam menyusun tes hasil atau prestasi belajar yang baku seperti berikut:

a. Menyusun Spesifikasi Tes

Langkah awal dalam mengembangkan tes adalah menetapkan menetapkan spesifikasi tes atau *blue print*, yaitu yang berisi uraian yang menunjukkan keseluruhan karakteristik yang harus dimiliki suatu tes. Spesifikasi yang jelas akan mempermudah dalam menulis soal, dan siapa saja yang menulis soal akan menghasilkan tingkat kesulitan yang relatif sama. Prosedur penyusunan spesifikasi tes adalah sebagai berikut: 1) menentukan tujuan tes, 2) menyusun kisi-kisi tes, 3) menentukan bentuk tes, dan 4) menentukan panjang tes.²⁹

b. Menulis Tes

Banyaknya butir soal yang harus dibuat untuk setiap bentuk soal, untuk setiap pokok bahasan, dan untuk setiap aspek kemampuan yang hendak diukur harus disesuaikan dengan yang tercantum dalam kisi-kisi. Ada beberapa petunjuk yang perlu diperhatikan dalam membuat butir-butir soal, antara lain: 1) soal yang harus dibuat harus valid dalam arti mampu mengukur tercapai tidaknya TIK yang telah dirumuskan, 2) soal yang dibuat harus dapat dikerjakan dengan menggunakan satu kemampuan spesifik, tanpa dipengaruhi oleh kemampuan lain yang tidak relevan, 3) soal yang dibuat harus terlebih dahulu dikerjakan atau diselesaikan dengan langkah-langkah lengkap sebelum digunakan pada tes yang sesungguhnya, 4) hindari kesalahan ketik, karena hal itu dapat mempengaruhi validitas soal, 5) tetapkan sejak awal kemampuan yang

²⁸ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: Raja Grafindo, 2009), hal. 97-98.

²⁹ Djemari Mardapi, *Loc.Cit.*, hal. 98-99.

hendak diukur untuk setiap soal, dan 6) berikan petunjuk cara mengerjakan soal secara jelas.³⁰

Pada umumnya terdapat dua bentuk butir tes, yaitu: 1) butir tes objektif yang menuntut peserta didik untuk memilih jawaban yang benar dari beberapa alternatif atau mengisi satu kata atau kalimat pendek untuk menjawab atau melengkapi pernyataan; dan 2) butir tes subjektif atau esai yang memberi kebebasan kepada peserta didik untuk menyusun dan mengemukakan jawaban yang orisinal.

1) Tes Bentuk Benar Salah

Tes seperti ini terdiri dari kalimat atau pernyataan yang mengandung dua kemungkinan jawab: benar atau salah, dan testee diminta memilih apakah pernyataan-pernyataan tersebut benar atau salah dengan cara tertentu.³¹

2) Bentuk Menjodohkan

Matching adalah tipe pertanyaan yang terdiri atas dari dua kolom, setiap pertanyaan pada kolom pertama harus dijodohkan dengan urutan pada kolom kedua.³² Dalam tes bentuk penjodohan, siswa dituntut untuk menjodohkan, mencocokkan, menyesuaikan, atau menghubungkan antara dua pernyataan yang disediakan. Pernyataan biasanya diletakkan dalam dua lajur, lajur kanan dan lajur kiri. Lajur kiri biasanya berupa pernyataan sedang lajur kanan berupa jawaban.³³

3) Bentuk Pilihan Ganda

Berkaitan dengan tes pilihan ganda ini, Ebel dalam I Wayan Koyan memberikan petunjuk sebagai berikut: a) Susun tes pilihan ganda berdasarkan ide-ide yang penting dan menunjukkan pernyataan yang

³⁰ I Wayan Koyan, *Konstruksi Tes*, (Singaraja: Universitas Pendidikan Ganesha, 2012), hal. 46.

³¹ Sri Esti Wuryani Djiwandono, *Psikologi Pendidikan*, (Jakarta: Grasindo, 2002), hal. 409.

³² Nana Sudjana, *Dasar-dasar Proses Belajar Mengajar*, (Bandung: Sinar Baru Algasindo Offset, 1995), hal. 123.

³³ Burhan Nurgiyantoro, *Penilaian dalam Pengajaran Bahasa dan sastra*, (Yogyakarta: BPFE, 1987), hal. 85.

bermakna, relevan, dan independen, b) pilih topik dan ide, kemudian tulis butir soal pilihan ganda yang mampu memaksimalkan daya beda butir-butir tes tersebut, c) susun draf awal dan adakan revisi, sehingga penggabungan menjadi seperangkat tes akhir menjadi sempurna, d) awali stem pertanyaan dengan pernyataan yang tidak lengkap dan disertai jawaban yang tepat serta dilengkapi dengan jawaban yang salah, tetapi masuk akal, e) susun jawaban yang benar sedemikian rupa atau secara acak tanpa menampakkan adanya petunjuk ke arah jawaban benar tersebut, dan f) pilih susunan pengecoh sedemikian rupa sehingga menjadi salah, tetapi tampak masuk akal, khususnya bagi peserta didik yang bodoh (Ebel, 1972: 191-202).

Dalam kaitan ini, Hopkin dan Antes dalam *I Wayaan Koyan* memberikan petunjuk yang lebih rinci dan praktis dalam menyusun tes pilihan ganda, yaitu: a) definisikan tugas-tugas dalam stem secara jelas, b) tulis alternatif jawaban pada akhir pertanyaan, c) tempatkan sebanyak mungkin kata-kata dalam stem, d) hindari penggunaan kata-kata negatif, e) hindari stem yang mengarah pada alternatif jawaban yang salah atau benar, f) buat alternatif jawaban yang paralel, g) tulis alternatif jawaban secara vertikal, h) hindari jawaban “semua di atas”, i) buat alternatif jawaban sama panjang, j) hilangkan petunjuk ke arah jawaban benar, k) buat pengecoh yang masuk akal, l) usahakan stemnya dalam bentuk pertanyaan, m) kontrol tingkat kesulitan soal sehingga persentase jawaban benar kira-kira separuhnya, n) hindari kemungkinan menebak, o) gunakan jawaban “tidak ada jawaban benar” hanya kalau tidak ada jawaban lain, p) susun alternatif jawaban sesuai dengan abjad atau urutan lainnya, q) letakkan jawaban benar secara acak, dan r) usahakan memiliki empat sampai lima alternatif jawaban (Hopkin dan Antes, 1990: 185-191).

4) Bentuk Tes Esai atau Tes Uraian

Tes esai sering disebut tes subjektif, karena proses pemberian skornya dipengaruhi oleh opini atau penilaian dari pendidik atau

pemeriksa tes tersebut. Seperti telah dikemukakan di atas bahwa tes esai termasuk ke dalam kelompok tes dengan jawaban tersusun (*constructed-response tests*). Jenis tes esai menghendaki peserta didik untuk mengorganisasikan, merumuskan, dan mengemukakan sendiri jawabannya. Dengan perkataan lain bahwa peserta didik tidak memilih jawaban, akan tetapi memberikan jawaban dengan kata-katanya sendiri secara bebas. Oleh karena itu, jawaban peserta didik tersebut hanya bisa diperiksa oleh mereka yang menulis butir tes tersebut atau oleh orang yang ahli atau mengetahui dengan jelas mengenai inti pokok persoalan yang ditanyakan dalam butir tes tersebut. Dalam hubungan ini, Hopkins dan Antes dalam I Wayan Koyan menyatakan bahwa tes esai adalah tes untuk mengembangkan jawaban atau respon peserta didik secara penuh. Keakuratan dan kualitas dari jawaban peserta didik harus dinilai oleh seseorang yang memiliki pengetahuan dan keahlian tentang materi yang diujikan, dalam hal ini biasanya adalah orang yang membuat butir soal tersebut.

Menurut Mehrens dan Lehmann dalam I Wayan Koyan, tes esai dapat digolongkan menjadi dua bentuk, yaitu tes esai jawaban terbuka (*extended-response*) dan jawaban terbatas (*restricted-response*), dan hal ini tergantung pada kebebasan peserta didik untuk mengorganisasikan atau menyusun ide-idenya dan menuliskan jawabannya. Pada tes esai bentuk jawaban terbuka atau jawaban luas dari pertanyaan uraian atau esai, mengizinkan peserta didik untuk mendemonstrasikan kecakapannya untuk: a) menyebutkan atas pengetahuan faktual, b) menilai pengetahuan faktualnya, c) menyusun ide-idenya, dan d) mengemukakan idenya secara logis dan koheren. Sedangkan pada tes esai jawaban terbatas atau terstruktur, peserta didik lebih dibatasi pada bentuk dan ruang lingkup jawabannya, karena secara khusus dinyatakan konteks jawaban yang harus diberikan oleh peserta didik.³⁴

³⁴ I Wayan Koyan, *Op.Cit.*, hal. 20.

5) Bentuk Jawaban Singkat

Soal jawaban singkat adalah soal yang menuntut peserta tes untuk memberikan jawaban singkat berupa kata, frase, nama, tempat, nama tokoh, lambang, atau kalimat yang sudah pasti.

Bentuk soal jawaban singkat sangat tepat digunakan untuk mengukur kemampuan peserta tes yang sangat sederhana. Kemampuan yang diukur dengan jawaban singkat adalah kemampuan menyebutkan istilah, kemampuan menyebutkan fakta, kemampuan menyebutkan prinsip, kemampuan menyebutkan metode atau prosedur, kemampuan menginterpretasi data sederhana, kemampuan memecahkan permasalahan yang berkaitan dengan anak-anak dan kemampuan melengkapi persamaan.

6) Unjuk Kerja/Performans

Dalam kenyataan, tes performa sering diabaikan dalam pengukuran KBM di sekolah, alasannya, mungkin karena tes performa lebih sulit digunakan daripada tes pengetahuan (kognitif), karena memerlukan lebih banyak waktu dalam mempersiapkan dan melaksanakannya, penyekorannya lebih subyektif dan memberatkan, serta guru harus membuat kriteria, yang memberikan gambaran secara khusus 'Apa yang dapat dilakukan dan yang tidak dapat dilakukan oleh setiap individu siswa'.

Tes performa dalam pelaksanaannya dapat dilakukan dalam beberapa alternatif, tahapan/tingkatan realitas mulai dari yang terendah sampai tingkatan tinggi (simulasi) dalam kehidupan nyata. Tentunya hal ini bergantung pada tujuan pengajaran, maupun pertimbangan praktis (waktu, biaya, sarana, ketersediaan perlengkapan, dsb).

7) Portofolio

Menurut Phopam dalam Djemari menyatakan Portofolio adalah kumpulan pekerjaan seseorang dalam mata pelajaran pendidikan. Menurut Marzano & Kendal dalam Djemari bahwa Portofolio cocok digunakan untuk penilaian kelas, tetapi tidak cocok untuk penilaian

dengan skala yang luas. Penilaian dengan portofolio memerlukan kemampuan membaca yang baik. Hal yang penting pada penilaian portofolio adalah mampu mengukur kemampuan membaca dan menulis yang lebih luas, peserta didik menilai kemajuannya sendiri, mewakili karya seseorang.

c. Menelaah Soal Tes

Kriteria yang digunakan untuk melakukan telaah butir tes mengikuti pedoman penyusunan tes. Telaah dilakukan terhadap kebenaran konsep, teknik penulisan, dan bahasa yang digunakan.

d. Melakukan Uji Coba Tes

Tes yang sudah dibuat dan diperbanyak itu akan diujicobakan kepada sejumlah sampel. Sampel uji coba harus memiliki karakteristik yang relatif sama dengan karakteristik peserta tes sesungguhnya. Jumlah sampel uji-coba harus mencukupi, minimal 5 kali jumlah butir soal.

Analisis butir soal, meliputi: analisis validitas, tingkat kesukaran, daya pembeda, dan analisis pengecoh. Soal yang tidak valid, didrop atau diperbaiki. Selanjutnya, dihitung reliabilitasnya untuk memperoleh gambaran tentang kualitas tes tersebut secara empirik.³⁵

e. Menganalisis Butir tes

Pada dasarnya analisis soal secara kualitatif dilaksanakan berdasarkan kaidah penulisan soal. Penelaan soal-soal tersebut dilakukan sebelum alat ukur tersebut digunakan. Faktor yang perlu dicermati dalam penelaan soal secara kualitatif adalah dari aspek isi, materi, konstruksi, bahasa, pedoman pemberian skor, dan kunci jawaban.

f. Memperbaiki Tes

Setelah uji coba dilakukan dan kemudian dianalisis, maka langkah berikutnya adalah melakukan perbaikan-perbaikan tentang bagian soal yang masih belum sesuai dengan yang diharapkan. Langkah ini biasanya

³⁵ I Wayan Koyan, *Ibid.*, hal. 46.

dilakukan atas butir soal, yaitu memperbaiki masing-masing butir soal yang ternyata masih belum baik. Ada kemungkinan beberapa soal sudah baik sehingga tidak perlu direvisi, beberapa butir mungkin perlu direvisi, dan beberapa yang lain mungkin harus dibuang karena tidak memenuhi standar kualitas yang diharapkan.

g. Merakit Tes

Supaya skor tes yang diperoleh dapat dipercaya, diperlukan banyak butir soal. Oleh karena itu, dalam penyajiannya, butir-butir soal perlu dirakit menjadi suatu alat ukur yang terpadu. Hal-hal yang dapat mempengaruhi validitas skor tes seperti urutan nomor soal, pengelompokan bentuk-bentuk soal, kalau dalam satu perangkat tes terdapat lebih dari satu bentuk soal, tata "*lay out*" soal harus diperhatikan dalam perakitan soal menjadi sebuah tes.

h. Melaksanakan Tes

Dalam praktek, pelaksanaan tes hasil belajar dapat diselenggarakan secara tertulis (tes tertulis), dengan secara lisan (tes lisan) dan dengan tes perbuatan.

1) Teknik Pelaksanaan Tes Tertulis

Dalam melaksanakan tes tertulis ada beberapa hal yang perlu mendapat perhatian, yaitu sebagaimana dikemukakan berikut ini. Pertama, agar dalam mengerjakan soal tes para peserta tes mendapat ketenangan, seyogyanya ruang tempat berlangsungnya tes dipikirkan yang jauh dari keramaian, kebisingan, suara hiruk pikuk dan lalu lalang orang.

Kedua, ruangan tes harus cukup longgar, tidak berdesak-desakan, tempat duduk diatur dengan jarak tertentu yang memungkinkan tercegahnya kerja sama yang tidak sehat di antara testee.

Ketiga, ruangan tes sebaiknya memiliki system pencahayaan dan pertukaran udara yang baik.

Keempat, jika dalam ruangan tes tidak tersedia meja tulis atau kursi yang memiliki alas tempat penulis, maka sebelum tes

dilaksanakan hendaknya sudah disiapkan alat berupa alat tulis yang terbuat dari *triplex*, *hardboard* atau bahan lainnya.

Kelima, agar testee dapat memulai mengerjakan soal tes secara bersamaan, hendaknya lembar soal-soal tes diletakkan secara terbalik.

Keenam, dalam mengawasi jalannya tes, pengawas hendaknya berlaku wajar.

Ketujuh, sebelum berlangsungnya tes, hendaknya sudah ditentukan lebih dahulu sanksi yang dapat dikenakan kepada testee yang berbuat curang.

Kedelapan, sebagai bukti mengikuti tes, harus disiapkan daftar hadir yang harus ditandatangani oleh seluruh peserta tes.

Kesembilan, jika waktu yang ditentukan telah habis, hendaknya testee diminta untuk menghentikan pekerjaannya dan secepatnya meninggalkan ruangan tes.

Kesepuluh, untuk mencegah timbulnya berbagai kesulitan di kemudian hari, pada Berita Acara Pelaksanaan Tes harus dituliskan secara lengkap, berapa orang testee yang hadir dan siapa yang tidak hadir, dengan menuliskan identitasnya (nomor urut, nomor induk, nomor ujian, nama dan sebagainya), dan apabila terjadi penyimpangan-penyimpangan atau kelainan-kelainan harus dicatat dalam berita acara pelaksanaan tes tersebut.

2) Teknik Pelaksanaan Tes Lisan

Tes lisan adalah tes dimana tester di dalam mengajukan pertanyaan-pertanyaan atau soalnya dilakukan secara lisan, dan testee memberikan jawabannya juga secara lisan pula.³⁶ Dari segi persiapan dan cara bertanya, tes lisan dapat dibedakan menjadi dua yakni tes lisan bebas, tes tanpa menggunakan pedoman yang dipersiapkan secara

³⁶ Anas Sudijono, *Pengantar Evaluasi Pendidikan.*, (Jakarta: Raja Grafindo Persada, 2003), hal. 75.

tertulis dan tes lisan pedoman, tentang apa yang ditanyakan kepada peserta didik.³⁷

Pertanyaan lisan dapat digunakan untuk mengetahui daya serap peserta didik untuk masalah yang berkaitan dengan kognitif yang baru diajarkan. Pertanyaan bisa diajukan di awal pembelajaran, yaitu mengenai konsep atau aplikasi pelajaran yang lalu. Pertanyaan lisan yang diajukan ke kelas harus jelas, dan semua peserta didik harus diberi kesempatan yang sama.

Dalam melakukan pertanyaan dikelas prinsipnya adalah mengajukan pertanyaan, memberi waktu untuk berpikir, kemudian menunjuk peserta didik untuk menjawab pertanyaan. Benar atau salah jawaban peserta didik, sebaiknya jawaban tersebut ditawarkan lagi ke kelas untuk mengaktifkan kelas. Tingkat berpikir untuk pertanyaan lisan di kelas bisa rendah sampai tinggi. Pertanyaan lisan memiliki kebaikan, yaitu melatih peserta didik dalam berkomunikasi secara lisa.³⁸

3) Teknik Pelaksanaan Tes Perbuatan

Tes perbuatan pada umumnya digunakan untuk mengukur taraf kompetensi yang bersifat keterampilan (psiko-motorik), dimana penilaiannya dilakukan terhadap proses penyelesaian tugas dan hasil akhir yang dicapai oleh peserta testee setelah melaksanakan tugas tersebut.

Dalam melaksanakan tes perbuatan itu, ada beberapa hal yang perlu diperhatikan oleh tester.

Pertama, tester harus mengamati dengan secara teliti, cara yang ditempuh oleh testee dalam menyelesaikan tugas yang telah ditentukan.

³⁷ M. Chabib Toha, *Teknik Evaluasi Pendidikan*, (Jakarta: Raja Grafindo Persada, 1996), hal. 61.

³⁸ Djemari Mardapi, *Loc.Cit.*, hal. 104-105.

Kedua, agar dapat dicapai kadar obyektivitas setinggi mungkin, hendaknya tester jangan berbicara atau berbuat sesuatu yang dapat mempengaruhi testee yang sedang mengerjakan tugas tersebut.

Ketiga, dalam mengamati testee yang sedang melaksanakan tugas itu, hendaknya tester telah menyiapkan instrumen berupa lembar penilaian yang didalamnya telah ditentukan hal-hal apa sajakah yang harus diamati dan diberikan penilaian.³⁹

Tes tindakan adalah “tes yang persoalan atau pertanyaan disampaikan dalam bentuk tugas yang harus dikerjakan oleh peserta didik. Alat yang dapat digunakan tes ini adalah berupa observasi atau pengamatan terhadap tingkah laku tersebut, yang hasilnya kemudian diserahkan pada guru.”⁴⁰

C. Kesimpulan

Bentuk tes yang selama ini digunakan di satuan pendidikan berbentuk tes esai dan tes objektif dalam berbagai variasi. Tes non objektif juga sering disebut dengan tes bentuk esai atau uraian. Tes obyektif di sini dilihat dari cara penskorannya, siapa saja yang memeriksa lembar jawaban akan menghasilkan skor yang sama. Penskoran adalah suatu proses pengubahan jawaban dari instrumen menjadi angka-angka yang merupakan nilai kuantitatif dari suatu jawaban terhadap item dalam instrumen. Jadi penskoran merupakan kuantifikasi terhadap jawaban instrumen.

Analisis pada umumnya dilakukan melalui dua cara yaitu analisis kualitatif (*qualitatif control*) dan analisis kuantitatif (*quantitatif control*). Analisis kualitatif sering sebagai validitas logis (*logical validity*) yang dilakukan sebelum soal digunakan untuk melihat berfungsi tidaknya sebuah soal. Analisis soal empiris (*empirical validity*) yang dilakukan untuk melihat lebih berfungsi tidaknya sebuah soal, setelah soal itu diujicobakan kepada

³⁹ Anas Sudijono, *Pengantar Evaluasi Pendidikan*, (Jakarta: RajaGrafindo, 2009), hal. 151-157.

⁴⁰ W.S. Winkel S.J., *Psikologi Pendidikan dan Evaluasi Belajar*, (Jakarta: Gramedia, 1983), hal. 106.

sampel yang representatif. Analisis kuantitatif meliputi tingkat kesulitan/kesukaran (TK) item, daya beda (DB) item, daya pengecoh (DP) dan reliabilitas soal.

Ada beberapa prinsip dasar yang perlu dicermati didalam menyusun tes hasil belajar agar tes tersebut dapat mengukur tujuan instruksional khusus untuk mata pelajaran yang telah diajarkan, atau mengukur kemampuan dalam keterampilan peserta didik yang diharapkan, setelah mereka menyelesaikan suatu unit pengajaran tertentu.

DAFTAR PUSTAKA

- Aiken, Lewis R. 1994. *Psychological Testing and Assessment*, Eight Edition, Boston: Allyn and Bacon.
- Cronbach, Lee J. 1984. *Essentials of Psychological Testing*. New York: Harper and Row Publisher.
- Djaali dan Pudji Muljono. 2008. *Pengukuran dalam Bidang Pendidikan*. Jakarta: Gramedia Widiasarana Indonesia.
- Djiwandono, Sri Esti Wuryani. 2002. *Psikologi Pendidikan*. Jakarta: Grasindo.
- Glass, Gene V. and Stanley, Julian C. 1970. *Statistical Methods in Education and Psychology*. New Jersey: Prentice Hall, Inc.
- Koyan, I Wayan. 2012. *Konstruksi Tes*. Singaraja: Universitas Pendidikan Ganesha.
- Kubiszyn, Tom and Gary Borich. 2003. *Educational Testing And Measurement: Classroom Application and Practice*. Unites States of America: Jhon Wiley & Sons. Inc.
- Mardapi, Djemari. 2016. *Pengukuran, Penilaian & Evaluasi Pendidikan*. Yogyakarta: Parama Publishing.

- Nurgiyantoro, Burhan. 1987. *Penilaian dalam Pengajaran Bahasa dan Sastra*. Yogyakarta: BPFE.
- Sudijono, Anas. 2009. *Pengantar Evaluasi Pendidikan*. Jakarta: Raja Grafindo.
- Sudjana, Nana. 1995. *Dasar-dasar Proses Belajar Mengajar*. Bandung: Sinar Baru Algasindo Offset.
- Surapranata, Sumarna. 2005. *Analisis, Validitas, Reliabilitas dan Interpretasi Hasil Tes: Implementasi Kurikulum 2014*. Jakarta: Remaja Rosdakarya.
- Suryabrata, Sumadi. 2006. *Psikologi Pendidikan*. Jakarta: PT Raja Grafindo Persada.
- Toha, M. Chabib. 1996. *Teknik Evaluasi Pendidikan*. Jakarta: Raja Grafindo Persada.
- W.S. Winkel S.J. 1983. *Psikologi Pendidikan dan Evaluasi Belajar*. Jakarta: Gramedia.
- Zulaiha, Rahman. 2008. *Analisis Soal Secara Manual*. Jakarta: Departemen Pendidikan Nasional Badan Pusat Penelitian dan Pengembangan Pusat Penilaian Pendidikan.