

BAB II

LANDASAN TEORI

A. Penelitian Relevan

Penelitian tentang evaluasi di bidang pendidikan telah banyak dilakukan oleh beberapa penulis. Penelitian ini juga didasarkan pada penelitian yang telah dilakukan sebelumnya. Ada beberapa hasil penelitian yang mendukung dan dapat dijadikan referensi dalam penelitian ini, yaitu sebagai berikut.

1. Peneliti yang dilakukan oleh Agus Milu Susetyo (UMJ-2019) dengan judul “Analisis Butir Soal Ujian Semester Kelas VIII Mata Pelajaran Bahasa Indonesia”

Pertama, dilihat dari tingkat kesukaran butir soal terdapat 5 jenis tingkat kesukaran di 40 butir pilihan ganda yaitu sangat mudah (3 soal), mudah (9 soal), sedang (22 soal), sukar (3 soal), sangat sukar (3 soal). Sementara itu, untuk soal uraian semuanya (5 soal) berkategori sedang. Kedua, dilihat dari daya pembeda tiap butir terdapat empat jenis yaitu dapat digunakan (9 soal), digunakan dengan revisi (5 soal), revisi (2 soal) dan dibuang (24 soal) untuk tipe soal pilihan ganda. Ketiga, dilihat dari daya pengecoh butir soal pada pilihan ganda terdapat 16 butir soal dengan daya pengecoh baik dan 24 butir soal dengan daya pengecohnya tidak baik.

Persamaan dengan penelitian yang dilakukan oleh penulis yaitu sama-sama menganalisis tentang kualitas butir soal pada tingkat SMP

dengan mata pelajaran Bahasa Indonesia. Sedangkan perbedaannya terletak pada apa yang dianalisis, skripsi Agus Milu Susetyo hanya menganalisis 3 aspek saja di antara lain tingkat kesukaran, daya pembedan, dan efektivitas pengecoh. Sedangkan penulis menganalisis 5 aspek antara lain tingkat kesukaran, daya pembeda, efektivitas pengecoh, validitas, dan reabilitas.

2. Penelitian yang dilakukan oleh Andi Surahma Halik (UIN Alauddin Makassar) pada tahun 2017 yang berjudul *Analisis Butir Soal Ujian Akhir Sekolah (UAS) Mata Pelajaran Matematika Pada Tahun Ajaran 2015/2016 SMP Negeri 36 Makassar*.

Hasil penelitian tersebut diantaranya: (1) Berdasarkan kriteria tingkat kesukaran soal diperoleh soal sangat mudah 2 butir (5%), sedang 24 butir (60%), sukar 12 butir (30%), dan sangat sukar 2 butir (5%), (2) Berdasarkan kriteria daya beda soal diperoleh soal dengan daya beda sangat jelek 3 butir (7,5%), jelek 12 butir (30%), cukup 14 butir (35%), dan baik 11 butir (27,5%) dan (3) Berdasarkan keefektifan pengecoh diperoleh 28 butir (70%) dalam kategori sangat baik, 9 butir (22,5%) dalam kategori baik, dan 3 butir (7,5%) dalam kategori buruk.

Persamaan dengan penelitian yang dilakukan oleh penulis sama-sama menganalisis tentang kualitas butir soal. Teknik pengumpulan data yaitu dokumentasi dan dianalisis dengan bantuan aplikasi Anates V4. Sedangkan perbedaannya terletak pada soal yang dianalisis, skripsi yang dilakukan oleh Andi Surahma Malik menganalisis soal mata pelajaran Matematika tingkat SMP. Sedangkan penulis menganalisis soal Penilaian Tengah Semester (PTS) genap mata pelajaran bahasa Indonesia tingkat SMP kelas VIII. Analisis yang dilakukan oleh Andi Surahma Malik hanya dari

segi tingkat kesukaran, daya beda dan efektivitas pengecoh. Sedangkan penulis menganalisis dari segi validitas, reliabilitas, daya pembeda, tingkat kesukaran dan efektivitas pengecoh.

B. Kajian Teori

Landasan teori dalam penelitian ini menyajikan berbagai teori yang digunakan sebagai acuan pelaksanaan penelitian yang berisi: kurikulum 2013, tes, Penilaian Tengah Semester (PTS), evaluasi pendidikan, karakteristik soal objektif, analisis butir soal, dan kerangka teoritis penelitian.

1. Kurikulum K13

Mulai tahun pelajaran 2013/2014, pemerintah memberlakukan kurikulum 2013 yang merupakan pembaruan dan penyempurnaan kurikulum 2006. Lismina (2018:5) mengungkapkan bahwa kurikulum merupakan program pendidikan yang disediakan oleh lembaga pendidikan untuk peserta didiknya, yang mencakup metode belajar mengajar, evaluasi pendidikan, program pendidikan, bimbingan dan konseling, supervisi, administrasi, serta hal-hal struktural lainnya.

Menurut Arifin (2017:1) kurikulum merupakan salah satu alat untuk mencapai tujuan pendidikan, sekaligus merupakan pedoman dalam pelaksanaan pembelajaran pada semua jenis dan jenjang pendidikan. Kurikulum mempunyai sifat yang dinamis, harus selalu menyesuaikan dengan perkembangan zaman, ilmu pengetahuan dan teknologi, tingkat kecerdasan peserta didik, kultur, serta kebutuhan masyarakat. Kurikulum menjadi pedoman

setiap guru dalam kegiatan belajar mengajar, sehingga pelajaran yang diterima siswa juga sesuai dengan kurikulum yang ada dan materi yang diberikan kepada siswa bisa terarah.

Kurikulum 2013 memiliki beberapa keunggulan menurut Mulyasa (2017:163-4), yaitu: (1) menggunakan pendekatan yang bersifat alamiah (kontekstual) sesuai dengan potensinya masing-masing; (2) berbasis karakter dan kompetensi; dan (3) ada bidang-bidang studi atau mata pelajaran lebih tepat menggunakan pendekatan kompetensi, terutama yang berkaitan dengan keterampilan. Yani (2014:9) menyebutkan tujuan kurikulum bagi guru yaitu dapat dijadikan acuan dalam menyampaikan materi setiap pembelajaran bersama peserta didik. Bagi peserta didik, kurikulum dapat menjadikan motivasi untuk terus belajar, dan bagi pengelola pendidikan, kurikulum dapat dijadikan sebagai indikator tingkat keberhasilan pembelajaran dan kurikulum.

Dari uraian diatas dapat disimpulkan bahwa kurikulum 2013 memiliki tujuan meningkatkan kreativitas guru dalam kegiatan proses pembelajaran agar materi yang diberikan kepada peserta didik dapat diterima dengan baik dan memiliki tujuan pembelajaran yang diinginkan. Jika pada kenyataannya peserta didik belum menguasai ilmu yang dipelajari, maka guru perlu melakukan perbaikan dalam proses pembelajarannya. Jadi, guru harus menguasai kemampuan melakukan penilaian hasil belajar dan membuat keputusan yang terbaik untuk peserta didik agar dapat melaksanakan tugas sebagai guru dengan baik.

2. Evaluasi Pendidikan

Menurut Widoyoko (2009: 5) pengukuran adalah kuantifikasi atau penetapan angka tentang karakteristik atau keadaan individu menurut aturan-aturan tertentu. Keadaan individu ini bisa berupa kemampuan kognitif, afektif dan psikomotor. Sedangkan pengertian penilaian menurut Arifin (2012: 17) adalah suatu proses atau kegiatan yang berkesinambungan untuk pengumpulan informasi tentang proses dan hasil belajar peserta didik untuk membuat keputusan berdasarkan kriteria tertentu. Evaluasi berasal dari bahasa Inggris yaitu *evaluation*. Dalam bahasa Indonesia mempunyai arti penilaian. Wand dan Brown (1977) dalam Sudijono (2012: 1) menjelaskan bahwa evaluasi adalah proses dalam menentukan nilai dari sesuatu.

Pengukuran, penilaian dan evaluasi pendidikan bersifat hirarki. Evaluasi didahului dengan penilaian, sedangkan penilaian didahului dengan pengukuran. Pengukuran diartikan sebagai kegiatan membandingkan hasil pengamatan dengan kriteria. Penilaian merupakan kegiatan menafsirkan dan medeskripsikan hasil pengukuran, sedangkan evaluasi merupakan penetapan nilai atau implikasi perilaku.

Menurut Ratnawulan dan Rusdiana (2015: 4) evaluasi pendidikan adalah penilaian terhadap kinerja pendidikan yang telah berjalan guru memperoleh informasi yang nantinya akan digunakan untuk memperbaiki hal-hal yang memang perlu diperbaiki pada kinerja pendidikan. Sedangkan Menurut Arikunto (2015 : 3) evaluasi merupakan sebuah proses pengumpulan data untuk menentukan sejauh mana, dalam hal apa, dan bagian mana tujuan pendidikan

sudah tercapai. Jika belum, bagaimana yang belum dan apa sebabnya. Terdapat tiga istilah dalam evaluasi pendidikan, yaitu; evaluasi, pengukuran, dan penilaian.

Fungsi dari evaluasi pendidikan secara umum sebagai suatu tindakan atau proses setidaknya –tidaknya memiliki tiga macam fungsi pokok, yaitu (1) mengukur kemajuan, (2) menunjang penyusunan rencana, dan (3) memperbaiki atau melakukan penyempurnaan kembali. Evaluasi pendidikan memiliki kegunaan diantaranya yaitu terbukanya kemungkinan bagi evaluator guna memperoleh informasi tentang hasil-hasil yang telah dicapai dalam rangka pelaksanaan program pendidikan, dapat diketahuinya relevansi antara program pendidikan yang telah dirumuskan dengan tujuan yang hendak dicapai, dapat dilakukannya usaha perbaikan, penyesuaian dan penyempurnaan program pendidikan (Sudijono 2012: 7).

Mengenai penilaian, Arikunto (2015: 18-9) menyatakan bahwa ada 4 tujuan atau fungsi penilaian yaitu (1) penilaian berfungsi selektif, misalnya 14 digunakan untuk memilih siswa yang seharusnya mendapat beasiswa atau kenaikan kelas; (2) penilaian berfungsi diagnostik, misalnya untuk mengetahui kelemahan siswa atau suatu permasalahan yang dialami oleh siswa beserta sebab-sebabnya; (3) penilaian berfungsi sebagai penempatan, misalnya untuk menentukan sekelompok siswa yang mempunyai hasil penilaian yang sama; dan (4) penilaian berfungsi sebagai pengukur keberhasilan, fungsi ini dimaksudkan untuk mengetahui sejauh mana suatu program berhasil diterapkan.

Berdasarkan pendapat yang dikemukakan, dapat disimpulkan bahwa

evaluasi pendidikan adalah suatu proses untuk mengukur, menilai dan akan mengetahui seberapa tujuan pendidikan yang diperoleh, dan tercapai tidaknya tujuan pendidikan yang diinginkan. Oleh karena itu, perencanaan evaluasi harus ditetapkan pada waktu menyusun satuan pembelajaran, agar dapat disesuaikan dengan tujuan instruksionalnya dan hasilnya dapat sesuai dengan yang diharapkan.

3. Penilaian Tengah Semester (PTS)

Pelaksanaan Penilaian Tengah Semester (PTS) dilakukan dengan tes, untuk menentukan hasil belajar peserta didik selama setengah semester dengan materi yang sesuai Kompetensi Dasar (KD). Istilah Penilaian Tengah Semester (PTS) merupakan pengganti dari Ulangan Tengah Semester (UTS) yang dilaksanakan pada tengah semester genap. Penilaian Tengah Semester (PTS) adalah kegiatan yang diberikan oleh sekolah kepada peserta didik untuk mengukur pencapaian materi yang sesuai dengan kompetensi dasar setelah melaksanakan kegiatan pembelajaran.

4. Tes

Kata tes berasal dari bahasa Prancis Kuno yang berarti piring untuk menyisihkan logam-logam mulia yang dimaksud di sini adalah dengan menggunakan alat berupa piring akan dapat diperoleh jenis-jenis logam mulia yang bernilai tinggi. Dalam perkembangannya dan seiring kemajuan zaman tes berarti ujian atau percobaan. Ada beberapa istilah yang memerlukan penjelasan sehubungan dengan uraian di atas, yaitu *test*, *testing*, *tester*, dan *testee*, yang masing-masing mempunyai pengertian sendiri-sendiri. *Test* adalah alat atau

prosedur yang dipergunakan dalam rangka pengukuran dan penilaian. *Testing* berarti saat dilaksanakannya pengukuran dan penilaian ayai saat pengambilan tes. *Tester* artinya orang yang melaksanakan tes atau orang yang dieserahi untuk melaksanakan pengambilan tes terhadap responden. *Testee* adalah pihak yang sedang dikenai tes (Ismail 2020: 15).

Tes pada hakikatnya adalah suatu alat yang berisi serangkaian tugas yang harus dikerjakan atau soal-soal yang harus dijawab oleh peserta didik untuk mengukur suatu aspek perilaku tertentu. Artinya, fungsi tes adalah sebagai alat ukur. Dalam tes prestasi belajar, aspek perilaku yang hendak diukur adalah tingkat kemampuan peserta didik (Arifin, 2012: 6-7). Tes yang dilaksanakan terdiri dari soal pilihan ganda dan soal isian. Salah satu alat yang digunakan sebagai sarana untuk penilaian hasil belajar yaitu tes. Apabila setelah dilaksanakan tes dan menunjukkan hasil belajar peserta yang baik, maka pembelajaran yang dilaksanakan guru bisa dibilang berhasil, namun kemampuan siswa juga berpengaruh dalam hal ini.

Tes merupakan suatu alat ukur yang paling sering digunakan oleh guru untuk mengukur hasil belajar peserta didik. Tes biasanya berisi tentang serangkaian tugas yang harus dikerjakan atau soal-soal yang harus dijawab oleh peserta didik untuk mengukur suatu aspek tertentu. Tes sebagai alat ukur perlu dirancang secara khusus sesuai dengan tujuan pembelajaran dan perlu dipersiapkan dengan sebaik-baiknya sesuai dengan kaidah-kaidah dalam penyusunan. Proses suatu pengukuran sangat diperlukan tes dengan kualitas yang baik, sebab baik buruknya kualitas tes akan menentukan data yang

dihasilakan.

a. Kriteria Tes yang Baik

Menurut Arikunto (2015: 72) tes dikatakan baik sebagai alat pengukur apabila memenuhi persyaratan sebagai berikut:

1) Validitas

Sebuah data atau informasi dikatakan valid jika sesuai dengan keadaan nyatanya. Jika data yang dihasilkan oleh instrumen yang valid, sesuai dengan kenyataan, maka instrumen yang digunakan tersebut juga valid, karna tes bisa dikatakan valid apabila tes tersebut dapat tepat mengukur apa yang hendak diukur.

2) Reliabilitas

Kata reliabilitas berasal dari kata *reliable* yang artinya dapat dipercaya. Tes dikatakan dapat dipercaya jika memberikan hasil yang tetap apabila diteskan berkali-kali. Tes dikatakan reliabel apabila hasil-hasil tes tersebut menunjukkan ketetapan. Dengan kata lain, yang dimaksud dari reliabilitas adalah ketetapan dari sebuah instrumen.

3) Objektivitas

Sebuah tes yang dikatakan memiliki objektivitas apabila dalam melaksanakan tes itu tidak ada faktor subjektif yang mempengaruhi. Hal ini terutama terhadai pada sistem skorsingnya.

4) Praktikabilitas

Tes dikatakan memiliki praktikabilitas yang tinggi apabila tes tersebut bersifat praktis dan mudah mengadministrasikannya. Tes yang praktis adalah tes

yang mudah dikerjakan, mudah pemikirannya, dan dilengkapi dengan petunjuk yang jelas.

5) Ekonomis

Sebuah alat atau instrumen dikatakan ekonomis apabila dalam pelaksanaannya tidak memerlukan biaya yang mahal, waktu dan tenaga yang banyak.

b. Bentuk-bentuk Tes

Ismail (2020: 17-27) mengemukakan bahwa bentuk tes yang digunakan di lembaga pendidikan dilihat dari segi sistem penskorannya dapat dikategorikan menjadi dua, yaitu:

1) Tes Objektif

Tes objektif biasa disebut "*short answer*" atau "*new type*". Tes objektif adalah soal yang terdiri dari item-item yang dapat dijawab dengan jalan memilih salah satu alternative yang benar dari sejumlah alternatif yang tersedia atau mengisi jawaban yang benar dengan beberapa perkataan dan symbol. Tes objektif terdiri dari atas beberapa bentuk yaitu:

a) Benar salah (*True-Falses, Or Yes-NO*)

Tes tipe benar-salah adalah tes yang isinya terdiri dari pernyataan yang disertai dengan alternatif jawaban. Yaitu jawab atau pertanyaan yang benar dan salah. Tes ini diminta untuk melingkari atau menandai huruf "B" jika pernyataan itu benar, dan "S" jika pernyataan itu salah.

b) Pilihan ganda (*Multi Choice*)

Soal bentuk pilihan ganda termasuk dalam bentuk tes objektif.

Penilaiannya dapat dilakukan secara objektif, karena dalam butir soal sudah mengandung jawaban yang harus dipilih oleh peserta tes. Siapapun yang mengoreksi jawaban tes objektif hasilnya akan sama, karena kunci jawaban sudah pasti dan jelas. Mengenai jumlah alternatif jawaban sebenarnya, tidak ada aturan baku. Semakin banyak alternatif jawaban, semakin bagus (Arifin, 2017:138).

Soal pilihan ganda adalah soal yang menuntut orang itu memberikan pilihan jawaban yang sudah tersedia dan menurutnya benar. Soal pilihan ganda ini sudah tersedia jawabannya 4-5, namun hanya satu jawaban yang benar dan harus dipilih. Sudjana (2014:48) menjelaskan, “Soal pilihan ganda adalah bentuk tes yang mempunyai satu jawaban yang benar atau paling benar”. Tes pilihan ganda terdiri dari pertanyaan atau pernyataan yang sifatnya belum selesai, sehingga dalam penyelesaiannya harus memilih salah satu alternatif jawaban yang telah disediakan pada setiap butir soal.

c) Menjodohkan (*Matching*)

Tes ini memiliki dua kolom yang parallel, dimana masing-masing kolom berisi uraian-uraianm, keterangan-ketrangan atau biasa disebut dengan *statement*. Disini murid ditugaskan untuk menjodohkan masing-masing keterangan yang berapada pada kolom sebelah kanan ke kolom sebelah kiri.

2) Tes Subjektif

Tes subjektif atau juga bisa disebut dengan tes esai yang merupakan salah satu jenis tes hasil belajar yang jawabannya menuntut seseorang untuk mengingat dan menuliskan kembali materi yang sudah disampaikan dengan

menggunakan bahasa sendiri namun inti dari jawaban itu sama saja. Tes esai memiliki jawaban yang berupa uraian. Bentuk-bentuknya berupa:

a) Tes uraian bebas

Tes uraian ini memberikan kebebasan kepada peserta tes untuk mengorganisasikan dan mengekspresikan pikiran dan gagasannya dalam menjawab soal tes. Jawaban peserta tes ini bersifat terbuka, bebas, fleksibel.

b) Tes uraian terbatas

Tes uraian ini memberikan batasan-batasan atau rambu-rambu kepada peserta tes dalam menjawab soal tes. Batasan atau rambu tersebut mencakup format, usi, dan ruang lingkup jawaban. Batasan yang dimaksud meliputi konteks jawaban yang diinginkan, jumlah dan luas jawaban yang diminta.

5. Karakteristik Soal Objektif

Tes objektif (*objective test*) yang juga dikenal dengan istilah tes jawaban pendek, tes “ya-tidak” dan tes model baru (*short answer test*), adalah salah satu jenis tes hasil belajar yang terdiri dari butir-butir soal yang dapat dijawab oleh *teste* dengan jalan memilih salah satu (atau lebih) di antara beberapa kemungkinan jawaban yang telah dipasangkan pada masing-masing items; atau dengan jalan menuliskan (mengisikan) jawabannya berupa kata-kata atau simbol-simbol tertentu pada tempat atau ruang yang telah disediakan untuk masing-masing butir item yang bersangkutan (Sudijono, 2015:106).

Menurut Ismail (2020: 17-27) Tes objektif biasa disebut “*short answer*” atau “*new type*”. Tes objektif adalah soal yang terdiri dari item-item yang dapat dijawab dengan jalan memilih salah satu alternative yang benar dari

sejumlah alternatif yang tersedia atau mengisi jawaban yang benar dengan beberapa perkataan dan symbol. Menurut Arikunto (2015:181-92) tes objektif terdiri dari empat macam, yakni tes benar-salah, tes pilihan ganda, tes menjodohkan, dan tes isian. Berdasarkan macam tes objektif tersebut, penelitian ini dibatasi pada bentuk tes pilihan ganda.

Tes pilihan ganda adalah tes yang sering dipakai oleh sekolah baik untuk ulangan harian, penilaian tengah semester, dan penilaian tengah semester. Tes ini setiap butirnya memiliki banyak alternatif jawaban yang bisa dipilih oleh peserta didiknya, namun jawaban tidak diperkenankan terlalu banyak, agar peserta didik tidak bingung untuk memilih jawaban yang tepat. Purwanto (2017: 41) menyatakan bahwa ada beberapa syarat khusus untuk menyusun soal dengan bentuk tes pilihan ganda, antara lain:

- a. Pernyataan dari setiap soal harus jelas merumuskan sebuah masalah.
- b. Kalimat antara pernyataan dan alternatif jawaban hendaknya tidak terlalu panjang.
- c. Alternatif jawaban haruslah homogen.

Berdasarkan penjelasan tersebut, tes pilihan ganda merupakan bentuk tes yang terdiri dari pertanyaan yang sifatnya belum selesai, sehingga peserta didik harus memilih salah satu alternatif jawaban yang telah disediakan pada setiap butir soal.

Tes objektif bentuk pilihan ganda (*multiple choice item*) memiliki banyak variasi atau model. Sudijono (2015: 120-130) menyatakan bahwa ada sembilan model variasi tes onjektif bentuk pilihan ganda, yaitu:

a. Model melengkapi lima pilihan

Model melengkapi lima pilihan ini pada umumnya terdiri atas kalimat pokok (=item) yang berupa pernyataan yang belum lengkap, diikuti oleh lima kemungkinan jawab (alternatif) yang dapat melengkapi pernyataan tersebut. Tugas peserta didik disini ialah memilih salah satu dari lima kemungkinan jawaban tersebut, yang menurut peserta didik paling benar. Contoh soal:

Bahasa laporan hendaknya tidak memenuhi kriteria berikut....

- a. Logis
- b. Baku
- c. Bahasa yang rancau
- d. Mudah dipahami
- e. jelas

(kunci jawaban: C)

b. Model asosiasi dengan lima pilihan

Model asosiasi dengan lima pilihan ini terdiri dari lima judul/istilah/pengertian yang diberi tanda hirif abjad di depannya, dan diikuti oleh beberapa pernyataan yang diberi nomer urut di depannya. Untuk tiap pernyataan tersebut peserta didik diminta memilih salah satu judul/istilah/pengertian yang berhuruf abjad yang menurut keyakinan peserta didik itu benar. Contoh soal:

Untuk soal nomer 1 sampai 5, pasangkan ciri yang terdapat di belakang huruf dengan pernyataan yang terdapat pada soal!

- a. Pantun
- B. Puisi
- C. Novel
- D. Cerpen
- E. Drama

Soal:

- 1. Terdiri dari 400 halaman atau jumlah kata lebih dari 35.000 kata.
- 2. Terdiri dari empat baris setiap baitnya.
- 3. Cerita fiktif.
- 4. Diksi yang dipakai biasanya bersifat kiasan, padat, dan indah.
- 5. Memiliki naskah atau teks yang di dalamnya berbentuk dialog.

(kunci jawaban: 1. C 2. A 3. D 4. B 5. E)

c. Model melengkapi berganda

Butir soal sejenis ini pada dasarnya sama dengan *multiple choice* item model melengkapi lima pilihan, yaitu terdiri atas satu kalimat pokok yang tidak (belum) lengkap, diikuti dengan beberapa kemungkinan jawaban. Perbedaannya adalah bahwa pada butir soal jenis ini, kemungkinan jawaban bisa satu, dua, tiga, atau empat. Berikut contoh soalnya:

- (1) Puisi memiliki ciri diksi yang dipakai biasanya bersifat kiasan, padat, dan indah.
- (2) Novel memiliki ciri terdiri dari 400 halaman atau jumlah kata lebih dari 35.000 kata.
- (3) Cerpen memiliki cerita yang fiktif.
- (4) Pantun memiliki ciri terdiri dari empat baris setiap baitnya.

Pilihlah:

- a. Bila (1), (2), dan (3) betul
- b. Bila (1) dan (3) betul
- c. Bila (2) dan (4) betul
- d. Bila semuanya betul
(kunci jawaban D)

d. Model analisis hubungan antarhal

Model Analisis Hubungan Antarhal Model ini terdiri dari satu kalimat pernyataan yang diikuti oleh satu kalimat keterangan. Berikut contoh soalnya:

Cerpen memiliki kata yang tidak lebih dari 10.000 kata

SEBAB

Cerpen hanya bercerita secara pendek saja dan jika melebihi 10.000 kata maka tidak disebut dengan cerpen lagi.

Pilihlah:

- a. Pernyataan benar, alasan benar, dan keduanya menunjukkan hubungan sebab akibat.
 - b. Pernyataan benar, alasan benar, tetapi keduanya tidak menunjukkan hubungan sebab akibat.
 - c. Pernyataan benar, tetapi alasan salah.
 - d. Pernyataan salah, tetapi alasan benar.
 - e. Pernyataan dan alasan salah.
- (kunci jawaban: A)

e. Model analisis kasus

Butir soal jenis ini merupakan tiruan keadaan yang sebenarnya. Jadi seolah-olah peserta didik dihadapkan dengan keadaan suatu kasus. Dari kasus tersebut, kepada peserta didik ditanyakan mengenai berbagai hal dan kunci jawaban-jawaban itu tergantung pada tahu atau tidaknya peserta didik dalam memahami kasus tersebut. Contoh soalnya:

Untuk menjawab soal berikut ini, pahami kasus secara cermat, kemudian jawablah!

Ibu Andi merupakan anak ke dua dari tiga bersaudara. kakak dari ibu Andi mempunyai dua orang anak. Bernama Ahmad dan Ahman. Maka andi memanggil Ahmad dan Ahman dengan silsilah apa?

- a. Kakak sepupu
- b. Adik sepupu
- c. Pakdhe
- d. Lilik

(kunci jawaban A)

f. Model hal kecuali

Model ini mencantumkan tiga macam kategori (yakni A, B, dan C), lalu di bawahnya terdapat lima hal atau keadaan (yakni 1, 2, 3, 4, dan 5), empat di antaranya cocok dengan satu hal yang dicantumkan dalam kategori. Berikut contoh soalnya:

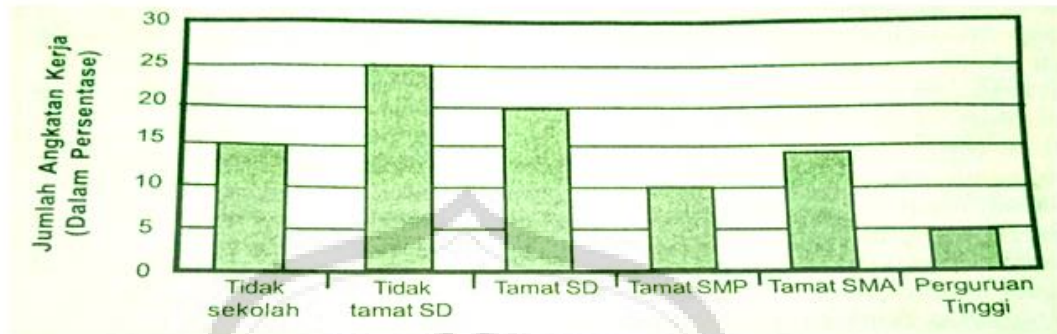
Kategori manakah yang berhubungan erat dengan empat hal tersebut, dan pilihlah hal yang tidak termasuk kelompok hal dimaksud di bawah ini!

- a. Jenis puisi lama yang tiap baitnya terdiri atas empat baris serta memiliki sampiran dan isi.
- b. Genre (jenis) sastra yang menggambarkan gerak kehidupan manusia.
- c. Salah satu jenis prosa yang isi ceritanya bukan kejadian nyata dan hanya dibuat-buat.
- d. Ragam sastra yang terikat oleh unsur-unsurnya, seperti irama, rima, baris dan bait.

- 1. Cerpen
- 2. Biografi
- 3. Drama
- 4. Pantun
- 5. Puisi

(kunci jawaban: D.2)

g. Model pemakaian gambar/grafik/diagram/peta



Pada model ini, terdapat gambar/grafik/diagram/peta yang selanjutnya pertanyaan diberikan sesuai dengan gambar/grafik/diagram/peta tersebut. Contoh soal:

Pernyataan berikut yang sesuai dengan grafik di atas ialah

- Jumlah angkatan kerja yang simpulan SD lebih tinggi dari pada tidak simpulan SD.
- Persentase angkatan kerja dari sekolah tinggi tinggi tidak mencapai 10%.
- Angkatan kerja yang simpulan Sekolah Menengah Pertama mencapai 10%.
- Jumlah angkatan kerja dari yang tidak sekolah meduduki peringkat paling tinggi.

(kunci jawaban: c)

Soal pilihan ganda juga memiliki kelebihan dan kekurangan. Menurut Ratnawulan dan Rusdiana (2015: 215) menyatakan bahwa kelebihan dari bentuk sial pilihann ganda antara lain:

Kelebihan bentuk soal pilihan ganda:

- Tes pilihan ganda memliki karakteristik yang baik untuk suatu alat pengukur hasil belajar siswa.
- Item tes pilihan ganda yang di konstruksi dengan intensif dapat mencakup hampir seluruh bahan pembelajaran yang diberikan oleh guru di kelas.

- c. Item tes pilihan ganda adalah tepat untuk mengukur penguasaan informasi para siswa yang hendak dievaluasi.

Kekurangan bentuk soal pilihan ganda:

- a. Mengonstruksi item tes betul salah pada umumnya diperlukan waktu yang lebih lama jika dibandingkan dengan pembuatan tes esai.
- b. Penggunaan pertanyaan alternatif lebih memungkinkan peserta didik mengira-ngira jawaban.

Berdasarkan uraian tersebut, dapat disimpulkan bahwa soal pilihan ganda bersifat objektif memiliki karakteristik koreksi yang sama. Maksudnya adalah jawaban bisa dikoreksi oleh siapapun dan tidak harus dikoreksi oleh pembuat soal pilihan ganda karena kunci jawaban dari soal ini sudah jelas.

6. Analisis Butir Soal

Sunarti dan Rahmawati (2014: 135) mengungkapkan bahwa analisis butir soal adalah suatu kegiatan yang wajib dilakukan oleh guru dalam pengembangan instrument penilaian. Analisis butir soal sangat digunakan untuk memperbaiki soal jika ternyata dalam pembuatannya masih dijumpai kekurangan atau kesalahan. Analisis butir soal merupakan kegiatan mengkaji pertanyaan-pertanyaan tes, agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai (Sudjana 2014: 135).

Kegiatan analisis butir soal ini mempunyai tujuan yaitu mengkaji dan menelaah setiap butir soal, agar diperoleh soal yang bermutu sebelum digunakan. Selain itu, analisis butir soal juga bertujuan untuk mengetahui informasi diagnostic peserta didik apakah mereka sudah atau belum memahami materi yang diberikan

oleh bapak ibu guru dalam proses pembelajaran berlangsung. Manfaat analisis butir soal menurut Kusaeri dan Suprananto (2012: 164) antara lain:

- a. Menentukan soal-soal yang cacat atau tidak berfungsi dengan baik.
- b. Meningkatkan butir soal melalui tiga komponen analisis yaitu tingkat kesukaran, daya pembeda, dan pengecoh soal.

Departemen Pendidikan Nasional (2008: 8) mengungkapkan bahwa analisis butir soal secara kuantitatif merupakan penelaah yang didasarkan pada data empiris dari butir soal yang telah diajukan. Analisis butir soal mempunyai dua pendekatan yaitu pendekatan secara klasik dan modern. Penelitian ini membahas tentang analisis butir soal secara klasik sesuai dengan fokus penelitian. Analisis butir soal klasik yaitu proses penelaah butir soal melalui informasi dari jawaban peserta didik guna meningkatkan mutu butir soal yang bersangkutan dengan menggunakan teori tes klasik. Aspek yang perlu diperhatikan dalam analisis butir soal secara klasik yaitu: validitas, reliabilitas, tingkat kesukaran, daya pembeda dan efektivitas pengecoh.

- a. Tingkat kesukaran

Tingkat kesukaran yaitu peluang peserta didik untuk menjawab benar suatu soal pada tingkat kemampuan tertentu yang biasa dinyatakan dengan indeks. Arikunto (2015: 222) menyatakan bahwa soal yang baik adalah soal yang tidak terlalu mudah atau terlalu sukar. Soal yang terlalu mudah tidak merangsang siswa untuk mempertinggi usaha memecahkannya. Sebaliknya soal yang terlalu sukar akan menyebabkan siswa menjadi putus asa dan tidak mempunyai semangat untuk mencoba lagi karena di luar jangkauannya.

Hal serupa juga dikemukakan oleh Sudijono (2012 : 370), bermutu tidaknya butir soal tes hasil belajar dapat ditentukan dari derajat kesukaran atau tingkat kesukaran butir soal tersebut. Bilangan yang menunjukkan sukar-mudahnya suatu soal disebut indeks kesukaran.

Indeks kesukaran dalam istilah evaluasi diberi simbol P (p besar), singkatan dari kata “proporsi”. Dengan demikian, soal dengan $P = 0,70$ lebih mudah daripada soal dengan $P = 0,20$. Sebaliknya, soal dengan $P = 0,30$ lebih sukar daripada soal dengan $P = 0,80$. Penelitian ini menggunakan ketentuan indeks kesukaran menurut Daryanto (2010 : 182).

No	Indeks Kesukaran	Kategori Soal
1	0,00 – 0,30	Sukar
2	0,31-0,70	Sedang
3	0,71- 1,00	Mudah

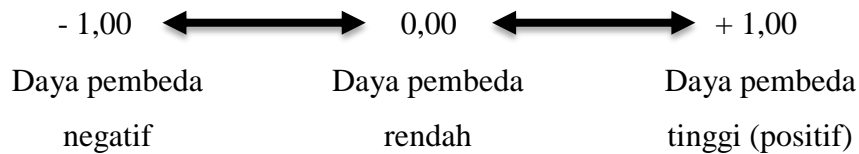
Tabel 1. Kategori Indeks Kesukaran Soal

Berdasarkan uraian di atas, dapat disimpulkan bahwa analisis tingkat kesukaran merupakan langkah prosedur untuk mengukur bagaimana tingkat kesukaran dari tiap butir soal, apakah termasuk dalam kategori soal yang mudah, sedang, atau sukar. Soal yang baik tidak terlalu mudah ataupun terlalu sukar, namun soal yang baik memiliki tingkat kesukaran sedang di indeks sekitar 0,30 sampai 0,70.

b. Daya pembeda

Arikunto (2015: 211) mengungkapkan bahwa daya pembeda soal adalah kemampuan sesuatu soal untuk membedakan antara siswa yang pandai (berkemampuan tinggi) dengan siswa yang memang mempunyai kemampuan

kurang (berkemampuan rendah). Terdapat tiga titik pada daya pembeda, yaitu:



Kriteria yang digunakan untuk menentukan indeks daya beda dan kualitas butir soal menurut Arikunto (2015: 232), seperti dalam tabel 2.4.

No.	Indeks Daya Beda	Kategori Soal
1.	0,00 – 0,20	Jelek
2.	0,21 – 0,40	Cukup
3.	0,41 – 0,70	Baik
4.	0,71 – 1,00	Baik Sekali
5.	Negatif (-)	Semuanya tidak baik

Tabel 2. Kategori Daya Pembeda Soal

Berdasarkan uraian di atas, dapat disimpulkan bahwa butir soal yang baik adalah butir soal yang mempunyai indeks diskriminasi 0,41 sampai 0,70. Apabila indeks diskriminasi item bertanda negatif berarti mempunyai arti soal tersebut lebih banyak dijawab oleh peserta didik kelompok bawah daripada peserta didik kelompok atas.

c. Efektivitas pengecoh

Pengecoh merupakan pilihan atau alternatif jawaban yang salah. Tujuan dari pengecoh ini yaitu membuat peserta didik merasa bimbang untuk memilih jawaban yang benar. Pengecoh dikatakan baik apabila telah memiliki daya rangsang atau daya tarik yang membuat peserta didik merasa bimbang dan ragu-

ragu, sehingga akhirnya mereka terkecoh dan memilih pengecoh sebagai jawaban yang betul. Menurut Arikunto (2015: 234) mengatakan bahwa suatu pengecoh dapat berfungsi dengan baik apabila paling sedikit dipilih oleh 5% pengikut tes.

Analisis butir soal merupakan kegiatan mengkaji pertanyaan-pertanyaan tes, agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai (Sudjana, 2014 : 135). Kegiatan analisis butir soal dapat dilakukan secara kualitatif dan/atau secara kuantitatif.

d. Validitas

Validitas merupakan sejauh mana suatu soal tes atau skala dalam menjalankan fungsi pengukurannya (Azwar 2015: 8). Validitas tes yaitu ketepatan pengukuran yang dimiliki oleh butir soal dalam mengukur apa yang seharusnya diukur lewat butir soal tersebut. Suatu alat ukur disebut valid bilamana alat ukur tersebut memiliki isi yang layak mengukur objek yang harus diukur dan sesuai dengan kriteria tertentu.

Sudijono (2012: 184) mengatakan bahwa butir soal dapat dikatakan valid apabila skor soal tersebut terbukti memiliki korelasi positif yang signifikan dengan skor totalnya. Skor total berkedudukan sebagai variabel terikat (*dependent viable*), sedangkan skor butir soal berkedudukan sebagai variabel bebas (*independent variable*).

Sudijono (2012: 185) mengungkapkan bahwa pada tes objektif hanya ada dua kemungkinan jawaban, yaitu jawaban benar dan salah. Setiap butir yang dijawab dengan benar pada umumnya akan diberi nilai 1 dan untuk jawaban salah diberi nilai 0. Jenis data seperti ini, yaitu: betul-salah, ya-tidak atau sejenis dengan

itu, dalam dunia ilmu statistic dikenal dengan nama data dikret muui atau data dikotomik.

Hasil dari perhitungan koefisien korelasi biserial kemudian dibandingkan dengan r_{tabel} pada taraf signifikansi 5% dan disesuaikan dengan jumlah peserta tes. Apabila $r_{hitung} > r_{tabel}$ maka soal tersebut valid (Arikunto, 2013: 92).

e. Reliabilitas

Kata reliabilitas berasal dari kata *reliable* yang artinya dapat dipercaya. Reliabilitas mengacu pada ketepatan hasil yang diperoleh dari suatu pengukuran. Sudjana (2014: 16) menyatakan “reliabilitas alat penilaian adalah ketetapan atau keajegan alat tersebut dalam menilai yang dinilai”. Artinya, kapanpun alat penilaian tersebut digunakan akan memberikan hasil yang relatif sama.

Tujuan dalam menghitung reliabilitas skor tes adalah untuk menentukan tingkat ketepatan dan keajegan skor tes. Semakin tinggi koefisien reliabilitas (mendekati 1), semakin tinggi pula keajegan/ketetapannya (Departemen Pendidikan Nasional 2008 : 16). Sudjana (2014: 209) dalam pemberian interpretasi terhadap koefisien reliabilitas tes pada umumnya digunakan patokan sebagai berikut :

1. Apabila r_{11} sama dengan atau lebih besar daripada 0,70 berarti tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan telah memiliki reliabilitas yang tinggi (*reliable*).
2. Apabila r_{11} lebih kecil daripada 0,70 berarti bahwa tes hasil belajar yang sedang diuji reliabilitasnya dinyatakan belum memiliki reliabilitas yang tinggi (*un-reliable*).

7. Kerangka Pikir

Dari uraian yang telah penulis paparkan di atas maka dapat dikatakan bahwa guru harus memenuhi pemilihan soal yang baik dengan syarat dapat dilihat dari beberapa hal antara lain adalah validitas atau sering disebut dengan apa yang akan diukur. Keterpercayaan (reliabilitas) mengacu bahwa suatu alat ukur/tes akan menghasilkan suatu yang kurang lebih sama apabila dikenakan pada setiap individu meskipun berbeda pelaksanaannya. Tingkat kesukaran mengacu bahwa suatu tes tidak terlalu sukar dan juga tidak terlalu mudah. Daya beda menunjukkan bahwa sebuah butir soal harus bisa membedakan antara siswa yang berkemampuan tinggi dan yang berkemampuan rendah. Efektivitas pengecoh mengisyaratkan bahwa semua pilihan jawaban (opsi) harus ada yang memilih dan jawaban yang betul harus lebih banyak dipilih oleh siswa yang berkemampuan tinggi.

Oleh karena itu ketika butir soal sudah dapat dikatakan baik (memenuhi kriteria) maka butir-butir soal itu dapat disimpan di bank soal dan dapat digunakan lagi. Sedangkan butir-butir soal yang belum dapat dikatakan baik dapat direvisi kembali dan butir-butir soal yang tidak memenuhi kriteria harus dihapus.