BAB 2

TINJAUAN TEORI

2.1 Instrumen Tes Hasil Belajar

Menurut Hamzah (2014: 92) instrumen arti sederhananya adalah seperangkat alat ukur berupa tulisan, materi, lisan yang dipakai untuk mengukur sesuatu. Pengelompokan alat ukur atas dasar prosedur pengumpulan datanya terbagi menjadi dua kelompok yaitu tes dan nontes. Menurut Yusuf (2015: 93) tes adalah suatu prosedur yang spesifik dan sistematis untuk mengukur tingkah laku seseorang; atau suatu pengukuran yang bersifat objektif mengenai tingkah laku seseorang, sehingga tingkah laku tersebut dapat digambarkan dengan bantuan angka, skala atau dengan sistem kategori, sedangkan menurut Tondowala (2012: 5) tes adalah sebuah alat atau instrumen yang dijadikan sebagai sarana dalam mengukur suatu prilaku tertentu pada peserta didik dengan menggunakan skala pengukuran tertentu. Di samping itu, menurut Sudijono (2013: 67) tes adalah cara (yang dapat dipergunakan) atau prosedur (yang perlu ditempuh) dalam rangka pengukuran dan penilaian dibidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas baik berupa pertanyaan-pertanyaan (yang harus dijawab), atau perintah-perintah (yang harus dikerjakan) oleh peserta didik, sehingga (atas dasar data yang diperoleh dari hasil pengukuran tersebut) dapat dihasilkan nilai yang melambangkan tingkah laku atau prestasi peserta didik dengan standar nilai tertentu. Dari beberapa pendapat tersebut, maka dapat disimpulkan bahwa tes adalah sebuah alat atau instrumen yang disusun secara sistematis dan dijadikan sebagai sarana untuk mengukur tingkat pemahaman seseorang dengan menggunakan skala pengukuran tertentu.

Menurut Aini (2017: 6) sebuah tes dapat mengukur aspek-aspek kognitif (pengetahuan), afektif (sikap), dan psikomotor (keterampilan). Hal yang akan diukur adalah tingkat penguasaan peserta didik terhadap materi yang telah diajarkan. Dalam pembelajaran, tes digunakan untuk meningkatkan pembelajaran. Dengan melalui tes, seorang guru dapat memperoleh informasi mengenai berhasil atau tidaknya peserta didik dalam menguasai tujuan-tujuan pembelajaran yang telah ditetapkan.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

Hasil belajar merupakan wujud pencapaian peserta didik; sekaligus merupakan gambaran keberhasilan pendidik dalam membelajarkan peserta didik. Menurut Yusuf (2015: 181) tes hasil belajar merupakan instrumen pengukuran dan penilaian untuk menentukan tingkat pencapaian peserta didik dalam belajar sesuai dengan karakteristik individualitas masing-masing. Di samping itu, menurut Purwanto (2013: 33) tes hasil belajar adalah tes yang dipergunakan untuk menilai hasil-hasil pelajaran yang telah diberikan oleh guru kepada murid-muridnya, atau oleh dosen kepada mahasiswa, dalam jangka waktu tertentu. Jadi, instrumen tes hasil belajar adalah alat (ukur) yang digunakan untuk menentukan pencapaian hasil belajar peserta didik, sedangkan instrumen tes hasil belajar matematika merupakan alat ukur yang dipakai dalam pembelajaran matematika, untuk menilai dan mengevaluasi sampai sejauh mana proses pembelajaran matematika mencapai sasarannya. Menurut Sudjana (2011: 12) suatu alat penilaian (tes) dikatakan mempunyai kualitas yang baik apabila alat tersebut memliki atau memenuhi dua hal, yakni ketepatannya atau validitasnya dan ketetapannya atau keajegannya atau reliabilitasnya.

Menurut Sudijono (2013: 97-99) ada beberapa prinsip dasar yang perlu dicermati dalam menyusun tes hasil belajar agar tes tersebut dapat mengukur tujuan instruksional khusus untuk mata pelajaran yang telah diajarkan, atau mengukur kemampuan dan keterampilan peserta didik yang diharapkan, setelah mereka menyelesaikan unit pengajaran tertentu. prinsip dasar tersebut adalah sebagai berikut:

1. Tes hasil belajar harus dapat mengukur secara jelas hasil belajar (*learning outcomes*) yang telah ditetapkan sesuai dengan tujuan instruksional.
2. Butir-butir soal tes hasil belajar harus merupakan sampel yang representatif dari populasi bahan pelajaran yang telah diajarkan.
3. Bentuk soal yang dikeluarkan dalam tes hasil belajar harus dibuat bervariasi, sehingga betul-betul cocok untuk mengukur hasil belajar yang diinginkan sesuai dengan tujuan tes itu sendiri.
4. Tes hasil belajar harus didesain sesuai dengan kegunaannya untuk memperoleh hasil yang diinginkan.
5. Tes hasil belajar harus memiliki reliabilitas yang dapat diandalkan.
6. Tes hasil belajar disamping harus dapat dijadikan alat pengukur keberhasilan siswa, juga harus dapat dijadikan alat untuk mencari informasi yang berguna untuk memperbaiki cara belajar siswa dan cara mengajar guru itu sendiri.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

Menurut Yusuf (2015: 200) langkah-langkah penyusunan tes hasil belajar adalah sebagai berikut:

1. Menetapkan tujuan penilaian.
2. Mengembangkan spesifikasi/*blueprint*.
3. Mengembangkan ruang lingkup ujian.
4. Memilih tipe item/butir soal.
5. Mempersiapkan penyusunan item/butir soal.
6. Menyusun item/butir soal.
7. Me-*review* item.
8. Mengadministrasikan tes/pelaksanaan tes.
9. Menskor dan mengolah tes.
10. Menginterpretasikan tes.
11. Menggunakan tes untuk perbaikan kegiatan pembelajaran dan "grading".

Menurut Djaali (dalam Sudaryono, dkk, 2013: 65-66) prosedur pengembangan instrumen tes hasil belajar, adalah:

1. Menetapkan tujuan tes. Tes prestasi belajar dapat dibuat untuk bermacam-macam tujuan, seperti: (a) tes yang bertujuan untuk mengadakan ujian nasional atau ujian lain yang sejenis; (b) tes yang bertujuan untuk mengadakan seleksi, misalnya SNMPTN; (c) tes yang bertujuan untuk mediagnosis kesulitan belajar siswa yang dikenal dengan tes diagnosis.
2. Analisis kurikulum. Analisis kurikulum bertujuan untuk menentukan bobot setiap pokok bahasan yang akan dijadikan dasar dalam menentukan jumlah item atau butir soal untuk setiap pokok bahasan soal objektif atau bobot soal untuk bentuk uraian, dalam membuat kisi-kisi tes.
3. Analisis buku pelajaran dan sumber materi belajar lainnya. Analisis ini mempunyai tujuan yang sama dengan analisis kurikulum, yaitu menentukan bobot setiap pokok bahasan. Namun, dalam analisis buku pelajaran menentukan bobot setiap pokok bahasan berdasarkan jumlah halaman materi yang termuat dalam buku pelajaran atau sumber materi lainnya.
4. Membuat kisi-kisi. Manfaat kisi-kisi adalah untuk menjamin sampel soal yang baik, dalam arti mencakup semua pokok bahasan secara proporsional. Agar item-item atau butir-butir tes mencakup keseluruhan materi (pokok bahasan atau sub pokok bahasan) secara proporsional, maka sebelum menulis butir-butir tes terlebih dahulu kita harus membuat kisi-kisi sebagai pedoman.
5. Penulisan Tujuan Instruksional Khusus (TIK). Penulisan TIK harus sesuai dengan ketentuan yang telah ditetapkan. TIK harus mencerminkan tingkah laku siswa, oleh karena itu harus dirumuskan secara operasional, dan secara teknis menggunakan kata-kata operasional.
6. Penulisan soal. Setelah kisi-kisi dalam bentuk tabel spesifikasi telah tersedia, maka kita akan membuat butir-butir soal. Beberapa petunjuk yang perlu diperhatikan dalam membuat butir-butir soal (khususnya matematika misalnya) adalah:

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

a. Soal yang dibuat harus valid (validitas konstruk) dalam arti mampu mengukur tercapai tidaknya tujuan pembelajaran yang telah dirumuskan.
b. Soal yang dibuat harus dapat dikerjakan dengan menggunakan satu kemampuan spesifik, tanpa dipengaruhi kemampuan lain yang tidak relevan.
c. Soal yang dibuat harus terlebih dahulu dikerjakan atau diselesaikan dengan langkah-langkah lengkap sebelum digunakan pada tes yang sesungguhnya.
d. Menetapkan sejak awal aspek kemampuan yang hendak diukur untuk setiap soal matematika yang dibuat.
e. Dalam membuat soal matematika, hindari sejauh mungkin kesalahan-kesalahan ketik betapapun kecilnya, karena hal itu akan memengaruhi validitas soal.
f. Memberikan petunjuk mengerjakan soal secara lengkap dan jelas untuk setiap bentuk soal matematika dalam suatu tes.

7. Telaah soal (*face validity*). Soal-soal yang dibuat masih mungkin terjadi kekurangan atau kekeliruan yang menyangkut aspek kemampuan spesifik yang diukur, bahasa yang digunakan, kesalahan ketik dan sebagainya. Untuk itu sebelum diperbanyak maka soal terlebih dahulu harus ditelaah oleh teman sejawat yang memahami materi tes maupun teknik penulisan soal untuk meneliti validitas permukaan dari soal yang dibuat.
8. Reproduksi tes terbatas. Tes yang sudah jadi diperbanyak dalam jumlah yang cukup menurut jumlah sampel uji-coba atau jumlah peserta yang akan mengerjakan tes tersebut dalam suatu kegiatan uji coba tes.
9. Uji-coba tes. Tes yang sudah diperbanyak itu akan diuji-cobakan pada sejumlah sampel yang telah ditentukan. Sampel uji-coba harus mempunyai karakteristik yang kurang lebih sama dengan karakteristik peserta tes yang sesungguhnya.
10. Analisis hasil uji-coba. Berdasarkan data hasil uji-coba dilakukan analisis, terutama analisis butir soal yang meliputi validitas butir, tingkat kesukaran, dan fungsi pengecoh. Berdasarkan validitas butir soal tersebut diadakan seleksi soal dengan menggunakan kriteria validitas tertentu.
11. Revisi soal. Soal-soal yang valid berdasarkan kriteria validitas empirik dikonfirmasikan dengan kisi-kisi. Apabila soal-soal tersebut sudah memenuhi syarat dan telah mewakili semua materi yang akan diujikan, soal-soal tersebut selanjutnya dirakit menjadi sebuah tes, tetapi apabila soal-soal yang valid belum memenuhi syarat berdasarkan hasil konfirmasi dengan kisi-kisi, dapat dilakukan perbaikan terhadap soal yang diperlukan.
12. Merakit soal menjadi tes. Urutan soal dalam suatu tes dilakukan menurut tingkat kesukaran, yaitu dari soal yang mudah sampai soal yang sulit.

Dalam penelitian ini, peneliti menggunakan langkah-langkah sebagai berikut dalam mengembangkan sebuah tes:

1. Menetapkan tujuan tes.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

2. Analisis kurikulum yang bertujuan untuk menentukan bobot setiap pokok bahasan yang akan dijadikan dasar dalam menentukan jumlah butir soal untuk setiap pokok bahasan.
3. Analisis buku pelajaran dan sumber materi belajar lainnya yang bertujuan untuk menentukan bobot setiap pokok bahasan berdasarkan jumlah halaman materi yang termuat dalam buku pelajaran atau materi ajar lainnya.
4. Membuat kisi-kisi
5. Penulisan tujuan instruksional yang dirumuskan secara operasional dan secara teknis menggunakan kata-kata operasional.
6. Penulisan soal.
7. Telaah soal yang dilakukan oleh validator.
8. Reproduksi tes terbatas menurut jumlah sampel uji-coba.
9. Uji coba tes.
10. Analisis hasil uji coba yang meliputi validitas, reliabilitas, dan analisis butir soal.
11. Revisi soal berdasarkan hasil rekomendasi analisis butir soal.
12. Merakit soal menjadi tes.

Menurut Permendikbud Nomor 23 tahun 2016 prosedur penilaian proses belajar dan hasil belajar olehpendidik dilakukan dengan urutan:

1. Menetapkan tujuan penilaian dengan mengacu pada RPP yang telah disusun.
2. Menyusun kisi-kisi penilaian.
3. Membuat instrumen penilaian berikut pedoman penilaian.
4. Melakukan analisis kualitas instrumen.
5. Melakukan penilaian.
6. Mengolah, menganalisis, dan menginterpretasikan hasil penilaian.
7. Melaporkan hasil penilaian.
8. Memanfaatkan laporan hasil penilaian.

Bentuk tes hasil belajar yang sering digunakan oleh guru di sekolah sebagai alat pengukur perkembangan dan kemajuan siswa, yaitu tes objektif dengan bentuk pilihan ganda (*multiple choice*). Menurut Arikunto (2012: 183) tes pilihan ganda (*multiple choice test*) terdiri atas suatu keterangan atau pemberitahuan tentang suatu pengertian yang belum lengkap. Untuk melengkapinya harus memilih satu dari beberapa kemungkinan jawaban yang telah disediakan, sedangkan Sudijono (2013: 118) mengemukakan bahwa tes objektif bentuk pilihan ganda, yaitu salah satu bentuk tes objektif yang terdiri atas pertanyaan atau pernyataan yang sifatnya belum selesai, dan untuk menyelesaikannya harus dipilih salah satu dari beberapa kemungkinan jawaban yang telah disediakan pada tiap-tiap butir soal yang bersangkutan. Jadi, tes objektif bentuk pilihan ganda adalah tes yang terdiri atas

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

pernyataan/pertanyaan, dan dua atau lebih alternatif jawaban yang akan melengkapi suatu pernyataan/pertanyaan tersebut.

Menurut Sudijono (2013: 133-135) adapun kelebihan yang dimiliki oleh tes objektif adalah:

1. Tes objektif sifatnya lebih representatif dalam hal mencakup dan mewakili materi yang telah diajarkan kepada peserta didik atau telah diperintahkan kepada peserta didik untuk mempelajarinya.
2. Tes objektif lebih memungkinkan bagi tester untuk bertindak lebih objektif, baik dalam mengoreksi lembar-lembar jawaban soal, menentukan bobot skor maupun dalam menentukan nilai hasil tesnya.
3. Mengoreksi hasil tes objektif jauh lebih mudah dan cepat dibanding mengoreksi hasil tes uraian.
4. Berbeda dengan tes uraian, maka tes objektif memberikan kemungkinan kepada orang lain untuk ditugasi atau dimintai bantuan guna mengoreksi hasil tes tersebut.
5. Butir-butir soal pada tes objektif, jauh lebih mudah dianalisis, baik analisis dari segi derajat kesukarannya, daya pembedanya, validitas maupun reliabilitasnya.

Menurut Sukardi (2015: 125-126) adapun kelebihan tes pilihan ganda adalah diantaranya:

1. Tes pilihan ganda memiliki karakteristik yang baik untuk suatu alat pengukur hasil belajar siswa. Karakter yang baik tersebut yaitu lebih fleksibel dalam implementasi evaluasi dan efektif untuk mengukur tercapai tidaknya tujuan belajar mengajar.
2. Item tes pilihan ganda yang dikonstruksi dengan intensif dapat mencakup hampir seluruh bahan pembelajaran yang diberikan oleh guru dikelas.
3. Item tes pilihan ganda adalah tepat untuk mengukur penguasaan informasi para siswa yang hendak dievaluasi.
4. Item tes pilihan ganda dapat mengukur kemampuan intelektual atau kognitif, afektif, dan psikomotorik siswa.
5. Dengan menggunakan kunci jawaban yang sudah disiapkan secara terpisah, jawaban siswa dapat dikoreksi dengan lebih mudah.
6. Hasil jawaban siswa yang diperoleh dari tes pilihan ganda dapat dikoreksi bersama, baik oleh guru maupun siswa dengan situasi yang lebih kondusif.
7. Item tes pilihan ganda yang sudah dibuat terpisah antara lembar soal dan lembar jawaban, dapat dipakai secara berulang-ulang.

Jadi, secara umum dapat disimpulkan bahwa kelebihan tes pilihan ganda diantaranya adalah:

1. Representatif dalam hal mencakup dan mewakili materi yang telah diajarkan kepada peserta didik.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

2. Fleksibel dan efektif untuk suatu alat pengukur hasil belajar peserta didik.
3. Penilaian lebih objektif.
4. Mengoreksi hasil tes objektif jauh lebih mudah dan cepat.
5. Butir-butir soal pada tes objektif, jauh lebih mudah dianalisis, baik analisis dari segi derajat kesukarannya, daya pembedanya, validitas maupun reliabilitasnya.
6. Dapat mengukur kemampuan intelektual atau kognitif, afektif, dan psikomotorik peserta didik.
7. Tepat untuk mengukur penguasaan informasi para siswa yang hendak dievaluasi.

Menurut Purwanto (2009: 39) kelemahan dari tes pilihan ganda adalah sebagai berikut:

1. Kurang memberi kesempatan untuk menyatakan isi hati atau kecakapan yang sesungguhnya karena siswa tidak membuat kalimat.
2. Memungkinkan siswa berbuat coba-coba dalam menjawabnya. Untuk menghindari kemungkinan ini guru harus dapat menyusun tesnya dengan teliti dan baik sehingga tes atau pertanyaan-pertanyaan itu benar-benar dapat merangsang berpikir siswa.
3. Menyusun tes ini tidak mudah, memerlukan ketelitian dan waktu yang agak lama.
4. Kurang ekonomis karena memakan biaya dan kertas yang banyak.

Menurut Sudijono (2013: 135-136) kelemahan dari tes objektif antara lain adalah:

1. Menyusun butir-butir soal tes objektif adalah tidak semudah seperti halnya menyusun tes uraian. Bukan hanya karena jumlah butir-butir soalnya cukup banyak, menyiapkan kemungkinan jawaban yang harus dipasangkan pada setiap butir item pada tes objektif itu juga bukan merupakan pekerjaan yang ringan.
2. Tes objektif pada umumnya kurang dapat mengukur atau mengungkap proses berpikir yang tinggi atau mendalam. Tes objektif lebih banyak mengungkap daya ingat atau hafalan ketimbang mengungkap tingkat kedalaman berpikir peserta didik terhadap materi yang diujikan. Kelemahan ini terutama disebabkan karena untuk memberikan jawaban pada tes objektif yang pada umumnya jawabannya pendek-pendek, peserta didik tidak dituntut berpikir secara mendalam.
3. Dengan tes objektif, terbuka kemungkinan bagi peserta didik untuk bermain spekulasi, tebak terka, adu untung dalam memberikan jawaban soal. Ini dapat terjadi, sebab bagi peserta didik yang sekalipun sebenarnya tidak tahu jawabannya, namun karena pada setiap butir soal sudah dipasang kemungkinan-kemungkinan jawabanya, maka tidak ada kesulitan sama sekali bagi peserta didik untuk menebak salah satu di antara kemungkinan jawaban yang telah tersedia. Kalau saja jawaban yang sifatnya menebak ternyata betul, maka tes tersebut akan menjadi alat pengukur yang diragukan daya ketepatan mengukurnya.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

4. Cara memberikan jawaban soal pada tes objektif, dimana dipergunakan simbol-simbol huruf yang sifatnya seragam, seperti: A, B, C, D, dan E atau B – S dan sebagainya, maka hal seperti ini dapat membuka peluang bagi peserta didik untuk melakukan kerja sama yang tidak sehat dengan sesama peserta didik lainnya. Misalnya dengan menggunakan kode-kode berupa gerakan tangan, gelengan kepala atau dengan cara-cara lainnya.

Jadi, secara umum dapat disimpulkan bahwa kelemahan tes pilihan ganda diantaranya adalah:

1. Memungkinkan siswa berbuat coba-coba dalam menjawabnya.
2. Menyusun tes memerlukan ketelitian dan waktu yang agak lama.
3. Kurang ekonomis.
4. Kurang dapat mengukur atau mengungkap proses berpikir yang tinggi atau mendalam.
5. Membuka peluang bagi peserta didik untuk melakukan kerja sama.

Menurut Sudjana (2011: 50-53) pedoman pembuatan tes bentuk pilihan ganda adalah:

1. Pokok soal (*stem*) yang merupakan permasalahan harus dirumuskan dengan jelas.
2. Perumusan pokok soal dan alternatif jawaban hendaknya merupakan pernyataan yang diperlukan saja.
3. Untuk setiap soal hanya ada satu jawaban yang benar atau yang paling benar.
4. Pada pokok soal (*stem*) sedapat mungkin dicegah perumusan pernyataan yang bersifat negatif.
5. Alternatif jawaban (*option*) harus logis dan pengecoh harus berfungsi.
6. Usahakan agar tidak ada “petunjuk” untuk jawaban yang benar.
7. Usahakan untuk tidak menggunakan *option* yang berbunyi “semua jawaban di atas salah” atau “semua jawaban di atas benar”.
8. Usahakan agar *option* homogen, baik dari segi isi maupun dari segi struktur kalimat.
9. Apabila *option* berbentuk angka, susunlah secara berurutan dari angka terkecil ke angka terbesar atau sebaliknya.

Menurut Sukardi (2015: 127) aturan penyusunan item tes pilihan ganda yakni sebagai berikut:

1. Pokok persoalan (*stem of item*) sebaiknya mengandung permasalahan atau *problem* yang dinyatakan dalam satu paragraf atau dalam bentuk pertanyaan.
2. Item tes pilihan ganda dengan empat jawaban, banyak digunakan untuk mengukur hasil pembelajaran siswa, dari empat jawaban tersebut hanya satu jawaban benar, sisanya atau tiga lainnya, disebut sebagai jawaban alternatif salah.
3. Jawaban benar dalam satu tes, direkomendasikan untuk diatur secara random pada semua *item*.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

4. Kata-kata yang tidak relevan, sebaiknya dihilangkan dari *stem*, agar ruang untuk pernyataan atau pertanyaan pada setiap *item* menjadi lebih jelas.
5. Hindari memberi kata-kata pada *item* yang mengandung pentunjuk (*clues*) yang mengarah pada jawaban benar, baik yang tersirat maupun tersurat.
6. Penataan jawaban sebaiknya diatur dengan posisi dalam bentuk kolom, tidak dalam bentuk paragraf.
7. Kalimat pada setiap *stem* sebaiknya menggunakan kalimat positif, kecuali jika guru atau evaluator sangat perlu menggunakan kalimat negatif.
8. Semua pilihan jawaban sebaiknya direncanakan, memiliki panjang atau jumlah kata yang sama, dan tidak mengandung petunjuk jawaban benar.

Kaidah-kaidah dalam penulisan soal pilihan ganda dapat disimpulkan sebagai berikut:

1. Pokok soal (*stem*) yang merupakan permasalahan harus dirumuskan secara jelas.
2. Perumusan pokok soal dan alternatif jawaban hendaknya merupakan pernyataan yang diperlukan saja.
3. Untuk setiap soal hanya ada satu jawaban yang benar atau yang paling benar.
4. Pada pokok soal sedapat mungkin dicegah perumusan pernyataan yang bersifat negatif.
5. Alternatif jawaban (*option*) sebaiknya logis dan pengecoh harus berfungsi.
6. Diusahakan agar tidak ada petunjuk untuk jawaban yang benar.
7. Diusahakan untuk mencegah penggunaan jawaban terakhir yang berbunyi “semua alternatif jawaban di atas benar” atau “semua pilihan jawaban di atas salah”.
8. Diusahakan agar alternatif jawaban homogen, baik dari segi isi atau materi maupun panjang-pendeknya kalimat (pernyataan).
9. Apabila alternatif jawaban berbentuk angka (bilangan), susunlah secara berurutan mulai angka terkecil di atas dan terbesar di bawah.
10. Di dalam pokok soal diusahakan tidak menggunakan ungkapan atau kata-kata yang bersifat tidak tentu, seperti kebanyakan, seringkali, kadang-kadang, dan sejenisnya.
11. Diusahakan agar jawaban butir soal yang satu tidak tergantung dari jawaban butir soal yang lain.
12. Dalam merakit soal harus diusahakan agar jawaban yang benar (kunci jawaban) tidak terpola, tetapi letaknya tersebar dan ditentukan secara *random*.
13. Penulisan alternatif jawaban diusahakan tersusun dari atas ke bawah, tidak ke samping. Hal ini dimaksudkan agar siswa tidak mudah lelah dalam membacanya, meskipun hal ini cukup menyita ruang dalam kertas. Dengan demikian salah satu faktor pembiasan alat evaluasi dapat ditekan.

Menurut Slameto (dalam Sukardi, 2015: 129-130) mengatakan bahwa penskoran *item* tes pilihan ganda dapat dibedakan menjadi dua, yaitu:

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

1. Penskoran dengan memperhitungkan jawaban *item* yang salah. Pada cara ini sebagian guru mempertimbangkan jawaban salah diperhitungkan dan digunakan sebagai denda untuk mengurangi jawaban yang benar. Penilaian model ini mengikuti formula seperti berikut:

$$N = B - \frac{S}{n-1}$$

   Keterangan:
   N: Nilai
   B: Jumlah jawaban benar
   S: Jumlah jawaban salah
   N: Banyak pilihan jawaban
2. Penskoran dengan tidak memperhitungkan jawaban salah. Dengan kata lain, jawaban salah tidak memengaruhi nilai pada jawaban benar. Skor akhir dari *item* tes pilihan ganda sama dengan jumlah jawaban benar.

2.2 Validitas Instrumen Tes

Menurut Borg dan Gall (dalam Purwanto, 2009: 114) validitas merupakan derajat sejauh mana tes mengukur apa yang ingin diukur, sedangkan menurut Sukardi (2015: 31) validitas instrumen tes adalah derajat yang menunjukkan dimana suatu tes mengukur apa yang hendak diukur. Validitas suatu instrumen tes mempunyai beberapa makna penting diantaranya seperti berikut:

1. Validitas berhubungan dengan ketepatan interpretasi hasil tes atau instrumen tes untuk grup individual dan bukan instrumen itu sendiri.
2. Validitas diartikan sebagai derajat yang menunjukkan kategori yang bisa mencakup kategori rendah, menengah, dan tinggi.
3. Prinsip suatu tes valid, tidak universal. Validitas suatu tes yang perlu diperhatikan adalah bahwa tes tersebut hanya valid untuk suatu tujuan tertentu saja.

Jadi, validitas adalah sejauh mana ketepatan dan kecermatan suatu alat ukur dalam melakukan fungsi ukurnya. Suatu skala atau instrumen pengukur dapat dikatakan mempunyai validitas yang tinggi apabila instrumen tersebut menjalankan fungsi ukurnya, atau memberikan hasil ukur yang sesuai dengan maksud dilakukannya pengukuran tersebut. Saran dan masukan dari para ahli dijadikan acuan untuk melakukan telaah terhadap butir soal. Menurut Kunandar (2014: 235-236) melalui kegiatan telaah soal, diharapkan akan menghasilkan soal yang baik, dari aspek materi, konstruksi, dan bahasa. Dalam rangka memudahkan kegiatan telaah soal bisa menggunakan format berikut:

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

Tabel 1. Telaah Soal Pilihan Ganda

Berilah tanda centang pada kolom "ya" atau "tidak" sesuai dengan kriteria telaah dari butir soal yang ditelaah.

Nomor Soal:			
No.	**Kriteria Penelaahan**	**Ya**	**Tidak**
A.	**MATERI**		
1.	Soal sesuai dengan indikator		
2.	Pengecoh berfungsi		
3.	Mempunyai jawaban benar/paling benar		
B.	**KONSTRUKSI**		
4.	Pokok soal dirumuskan secara jelas dan tegas		
5.	Rumusan soal dan rumusan jawaban merupakan pernyataan yang diperlukan saja		
6.	Pokok soal tidak memberi petunjuk ke arah jawaban yang benar		
7.	Pokok soal tidak mengandung pernyataan yang bersifat negatif ganda		
8.	Pilihan jawaban homogen dan logis ditinjau dari segi materi		
9.	Panjang rumusan jawaban relatif sama		
10.	Pilihan jawaban tidak mengandung pernyataan "semua pilihan jawaban di atas salah" atau "semua pilihan jawaban di atas benar"		
11.	Pilihan angka yang berbentuk angka atau waktu disusun berdasarkan urutan besar kecilnya nilai angka tersebut aau kronologis waktunya		
12.	Gambar, grafik, tabel, diagram, dan sejenisnya jelas dan berfungsi		
13.	Butir soal tidak bergantuk pada soal sebelumnya		
C.	**BAHASA**		
14.	Soal menggunakan bahasa yang sesuai dengan kaidah Bahasa Indonesia		
15.	Bahasa yang digunakan komunikatif		
16.	Tidak menggunakan bahasa yang berlaku setempat		
17.	Pilihan jawaban tidak mengulang kata/frase yang bukan merupakan satu kesatuan pengertian		
Catatan:			

Peneliti menggunakan modifikasi format telaah soal di atas sebagai lembar validasi yang akan diisi oleh validator.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

2.3 Analisis Butir Soal

Menurut Sudjana (2011: 135) analisis butir soal atau analisis *item* adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai, sedangkan menurut Mudjijo (dalam Ariandani, 2013: 8-9) analisis butir soal merupakan suatu kegiatan yang dapat membantu meningkatkan kebaikan suatu tes dengan melihat tingkat kebaikan tiap-tiap butir soalnya. Jadi, analisis butir soal adalah suatu kegiatan analisis untuk menentukan tingkat kebaikan butir-butir soal yang terdapat dalam suatu tes sehingga informasi yang dihasilkan dapat kita pergunakan untuk memperbaiki butir soal dan tes tersebut. Suatu tes yang baik berarti tes tersebut memiliki butir-butir soal yang baik pula, oleh karena itu pengujian tingkat kebaikan suatu soal tes tidak terlepas dari analisis kebaikan butir-butir soalnya.

Menurut Aini (2017: 13) dalam pengembangan sebuah tes, perlu adanya analisis butir tes. Pada analisis butir, butir akan dilihat karakteristiknya dan dipilih butir-butir yang baik. Butir yang baik adalah butir yang karakteristiknya memenuhi syarat sebagaimana kriteria karakteristik butir yang baik.

Menurut Thorndike (dalam Purwanto, 2009: 118) analisis terhadap soal-soal (*items*) tes yang telah dijawab murid-murid mempunyai dua tujuan penting, yaitu:

1. Jawaban-jawaban soal itu merupakan informasi diagnostik untuk meneliti pelajaran dari kelas itu dan kegagalan-kegagalan belajarnya, serta selanjutnya untuk membimbing ke arah cara belajar yang lebih baik.
2. Jawaban-jawaban terhadap soal-soal yang terpisah dan perbaikan (*review*) soal-soal yang didasarkan atas jawaban-jawaban itu merupakan basis bagi penyiapan tes-tes yang lebih baik untuk tahun berikutnya.

Jadi, tujuan khusus dari analisis soal tes adalah mencari soal tes mana yang baik dan mana yang tidak baik, dan mengapa *item* atau soal itu dikatakan baik atau tidak baik. Dengan mengetahui soal-soal yang tidak baik itu selanjutnya kita dapat mencari kemungkinan sebab-sebab mengapa *item* itu tidak baik. Dengan membuat analisis soal, sedikitnya kita dapat mengetahui tiga hal penting yang dapat diperoleh dari tiap soal, yaitu:

1. Sampai dimana tingkat atau taraf kesukaran soal itu (*difficulty level of an item*).

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

2. Apakah soal itu mempunyai daya pembeda (*discriminating power*) sehingga dapat membedakan kelompok siswa.
3. Apakah semua alternatif jawaban (*options*) menarik jawaban-jawaban, ataukah ada yang demikian tidak menarik sehingga tidak perlu dimasukkan ke dalam soal.

Menurut Sudijono (2013: 369) identifikasi setiap butir *item* tes hasil belajar itu dilakukan dengan harapan akan menghasilkan berbagai informasi berharga, yang pada dasarnya akan memberikan umpan balik (*feed back*) guna melakukan perbaikan, pembenahan, dan penyempurnaan kembali terhadap butir-butir *item* yang telah dikeluarkan dalam tes hasil belajar, sehingga pada masa-masa yang akan datang tes hasil belajar yang disusun atau dirangcang oleh *tester* (guru, dosen, dan lain-lain) itu betul-betul dapat menjalankan fungsinya sebagai alat pengukur hasil belajar yang memiliki kualitas tinggi. Analisis butir soal tersebut terdiri atas:

1. Analisis Kualiatif

Menurut Mudjijo (dalam Ariandani, 2013: 10) analisis butir soal secara kualitatif artinya menguji tingkat kebaikan suatu butir soal mengenai isi dan bentuk serta kesesuaian butir soal tersebut terhadap tujuan pembelajaran (indikator) yang telah ditetapkan dalam kurikulum, atas dasar penelitian seorang ahli bidang studi. Data kualitatif berasal dari dua sumber, yaitu: (1) hasil penilaian dan masukan dari ahli (validator), dan (2) saran dan masukan dari *reviewer*. Ahli yang menjadi validator, yaitu ahli bidang studi.

Analisis kualitatif bertujuan untuk mengetahui apakah butir soal yang digunakan mengukur tujuan pembelajaran (indikator) tertentu yang telah ditetapkan dalam kurikulum. Butir soal yang tidak secara akurat mengukur tujuan pembelajaran (indikator) tertentu yang telah ditetapkan merupakan butir soal yang mubazir. Bahkan bila butir soal tersebut diikut sertakan dalam pertimbangan untuk menentukan keputusan bagi siswa, maka keputusan tersebut dapat berakibat fatal dimasa yang akan datang bagi siswa.

2. Analisis Kuantitatif

Menurut Mudjijo (dalam Ariandani, 2013: 10) analisis kuantitatif berarti menguji tingkat kebaikan butir soal melalui teknik statistik. Data kuantitatif diperoleh dari hasil uji coba produk yang diperoleh dalam bentuk skala nominal.

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

Data diperoleh dari lembar respon peserta didik. Peserta didik yang menjawab benar untuk tiap butir soal mendapat skor 1, sedangkan yang salah mendapat skor 0. Analisis kuantitatif ini terdiri dari:

1) Tingkat Kesukaran

Menurut Sukardi (2015: 136) tingkat kesulitan atau disebut juga indeks kesulitan *item* adalah angka yang menunjukkan proporsi siswa yang menjawab betul dalam satu soal yang dilakukan dengan menggunakan tes objektif. Tingkat kesulitan tes *item* pada umumnya ditunjukkan dengan presentase siswa yang memperoleh jawaban item benar. Menurut Sudjana (2011: 135) untuk memperoleh kualitas soal yang baik diperlukan adanya keseimbangan dari tingkat kesulitan soal tersebut. Keseimbangan yang dimaksudkan adalah adanya soal-soal yang termasuk mudah, sedang, dan sukar secara proporsional. Tingkat kesukaran soal dipandang dari kesanggupan atau kemampuan siswa dalam menjawabnya, bukan dilihat dari sudut pandang pembuat soal.

Menurut Sudijono (2013: 370) bermutu tidaknya butir-butir *item* tes hasil belajar pertama-tama 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* tes tersebut tidak terlalu sukar dan tidak pula terlalu mudah dengan kata lain derajat kesukaran item itu adalah sedang atau cukup.

Menurut Witherington (dalam Sudijono, 2013:371) angka indeks kesukaran *item* itu besarnya berkisar antara 0,00 sampai dengan 1,00. Artinya, angka indeks kesukaran itu paling rendah adalah 0,00 dan paling tinggi adalah 1,00. Angka indeks kesukaran sebesar 0,00 merupakan petunjuk bagi tester bahwa butir *item* tersebut termasuk dalam kategori *item* yang terlalu sukar, sebab disini seluruh peserta didik tidak dapat menjawab item dengan betul (yang dapat menjawab dengan betul $= 0$). Sebaliknya, apabila angka indeks kesukaran *item* itu adalah 1,00 hal ini menganduk makna bahwa butir item yang bersangkutan adalah termasuk dalam kategori item yang terlalu mudah, sebab disini seluruh peserta didik dapat menjawab dengan betul butir item yang bersangkutan (yang dapat menjawab butir item$= 100\% = 100 \div 100 = 1{,}00$).

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

2) Daya Pembeda

Barnard (dalam Sukardi, 2015: 138) daya pembeda adalah angka atau koefisien yang memberikan informasi tentang pembeda secara individual, termasuk membedakan antara siswa yang pencapaiannya tinggi dengan siswa yang pencapaiannya rendah dalam suatu tes pencapaian hasil belajar.

Menurut Sudijono (2013: 387) daya pembeda *item* itu dapat diketahui dengan melihat besar kecilnya angka indeks diskriminasi *item*. Angka indeks diskriminasi *item* adalah sebuah angka atau bilangan yang menunjukkan besar kecilnya daya pembeda (*discriminatory* power) yang dimiliki oleh sebutir *item*. *Discriminatory power* pada dasarnya dihitung atas dasar pembagian peserta didik kedalam dua kelompok, yaitu kelompok atas dan kelompok bawah. Adapun cara menentukan dua kelompok itu bisa bervariasi; misalnya; dapat menggunakan median sehingga pembagian menjadi dua kelompok itu terdiri atas 50% peserta didik kelompok atas dan 50% peserta didik kelompok bawah; dapat juga dengan hanya mengambil 20% dari pesera didik yang termasuk dalam kelompok atas dan 20% lainnya diambilkan dari peserta didik yang termasuk dalam kelompok bawah; dapat juga menggunakan angka presentase lainnya. Namun, pada umumnya para pakar bidang evaluasi pendidikan lebih banyak menggunakan presentase sebesar 27% dari peserta didik yang termasuk dalam kelompok atas dan 27% lainnya diambilkan dari peserta didik yang termasuk dalam kelompok bawah. Hal ini disebabkan berdasarkan bukti-bukti empirik penganbilan subyek sebanyak 27% peserta didik kelompok atas dan 27% peserta didik kelompok bawah itu telah menunjukkan kesensitifannya, atau dengan kata lain cukup dapat diandalkan.

Pada penelitian ini jumlah siswa uji coba sebanyak 65 orang. Menurut Arikunto (2012: 227) mengemukakan bahwa untuk kelompok kecil (jumlah siswa kurang dari 100) seluruh *testee* dikelompokkan menjadi dua, yakni 50% kelompok siswa kemampuannya tinggi dan 50% kelompok siswa kemampuannya rendah.

3) Fungsi Distraktor

Menurut Sudijono (2013: 409-411) tes objektif bentuk pilihan ganda untuk setiap butir *item* yang dikeluarkan dalam tes hasil belajar telah dilengkapi dengan beberapa kemungkinan jawaban, atau yang disebut dengan istilah *option* atau

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

alternatif. *Option* atau alternatif itu jumlahnya berkisar antara tiga sampai dengan lima buah, dan dari kemungkinan-kemungkinan jawaban yang terpasang pada setiap butir *item* itu, salah satu di antaranya adalah jawaban yang benar; sedangkan sisanya adalah jawaban yang salah yang dikenal dengan istilah *distractor* (pengecoh). Tujuan utama dari pemasangan distraktor pada setiap butir item itu adalah agar dari sekian banyak peserta didik yang mengikuti tes hasil belajar ada yang tertarik atau terangsang untuk memilihnya, sebab mereka menyangka bahwa distraktor yang mereka pilih itu merupakan jawaban betul. Semakin banyak peserta didik yang terkecoh, maka kita dapat menyatakan bahwa distraktor yang terpasang pada item itu makin dapat menjalankan fungsinya dengan baik. Menurut Kunandar (2014: 241) pengecoh dapat dikatakan berfungsi apabila pengecoh paling sedikit dipilih oleh 2,5% peserta tes dan lebih banyak dipilih oleh kelompok bawah dari peserta tes.

4) Validitas Butir Soal (*Item*)

Menurut Arikunto (2012: 90) jika seorang peneliti atau guru mengetahui bahwa validitas soal tes misalnya terlalu rendah atau rendah saja, maka selanjutnya ingin mengetahui butir-butir tes manakah yang menyebabkan soal secara keseluruhan tersebut jelek karena memiliki validitas rendah, untuk keperluan inilah dicari validitas butir soal. Sehingga, perlu dilakukan telaah dan tindakan yang tepat pada soal-soal yang menyebabkan tes tersebut menjadi jelek. Menurut Yusuf (2015: 65) validitas instrumen dapat diketahui dengan jalan mencari korelasi instrumen dengan kriterium, atau melakukan analisis butir (*item*). Menurut Jihad dan Haris (2012: 179-180) dalam penentuan tingkat validitas butir soal digunakan korelasi *product moment Pearson* dengan mengkorelasikan antara skor yang didapat siswa pada suatu butir soal dengan skor total yang didapat.

5) Reliabilitas

Menurut Arifin (2013: 258) reliabilitas adalah tingkat atau derajat konsistensi dari suatu instrumen. Reliabilitas tes berkenaan dengan pertanyaan, apakah suatu tes teliti dan dapat dipercaya sesuai dengan kriteria yang telah ditetapkan. Suatu tes dapat dikatakan reliabel jika selalu memberikan hasil yang sama bila diteskan pada kelompok yang sama pada waktu atau kesempatan yang

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :

berbeda. Tes yang reliabel adalah apabila koefisien reliabilitasnya tinggi dan kesalahan baku pengukurannya (*standard error of measurement*) rendah, sedangkan menurut Yusuf (2015: 74) suatu alat ukur dikatakan reliabel, apabila alat ukur itu diujikan kepada objek atau subjek yang sama secara berulang-ulang, hasilnya akan tetap sama, konsisten, stabil atau relatif sama (tidak berbeda secara statistik). Prosedur mengukur konsistensi internal tanpa membagi tes ke dalam dua bagian bertujuan untuk menilai konsistensi antar item atau homogenitas item-item dalam satu tes evaluasi yang direncanakan. Untuk tes bentuk soal objektif mencari reliabilitasnya menggunakan formula Kuder-Richardson yang merupakan termasuk dalam pendekatan *single test-single trial* yang berarti penentuan reliabilitas tes tersebut dilakukan dengan hanya menggunakan satu jenis alat pengukur, dan pelaksanaan pengukuran itu hanya dilakukan sebanyak satu kali saja (Sudijono, 2013: 213-214).

Perpustakaan Universitas Islam Riau
Dokumen ini adalah Arsip Milik :