

BAB I

PENDAHULUAN

1.1. Latar Belakang Masalah

Pendidikan adalah usaha sadar dan terencana untuk mewujudkan suasana belajar dan proses pembelajaran agar peserta didik secara aktif mengembangkan potensi dirinya untuk memiliki kekuatan spiritual keagamaan, pengendalian diri, kepribadian, kecerdasan, akhlak mulia, serta keterampilan yang diperlukan dirinya, masyarakat, bangsa dan negara. Pendidikan merupakan upaya untuk mengubah sesuatu menjadi lebih baik. Melalui pendidikan, manusia mendapatkan pengetahuan tentang hal baru yang bernilai sangat baik sehingga mereka dapat menjadi seseorang yang memiliki karakter yang baik.

Pendidikan yang berkualitas tinggi tidak terlepas dari kurikulum yang dirancang dengan sangat baik. Kurikulum mengandung makna dimensi gagasan, perencanaan, pelaksanaan, dan evaluasi untuk memperoleh predikat pencapaian akhir. Kurikulum harus selalu dinamis mengikuti perubahan zaman berdasarkan landasan pengembangan yang benar. Kurikulum memiliki empat hal yang harus didiskusikan, keempat hal tersebut adalah capaian kompetensi siswa, materi pembelajaran, proses pembelajaran, dan penilaian atau asesmen (Nugroho, 2018: 8).

Selanjutnya, belajar adalah suatu proses usaha yang dilakukan seseorang untuk memperoleh suatu perubahan tingkah laku yang baru secara keseluruhan, sebagai hasil pengalamannya sendiri dalam interaksi dengan lingkungannya (Slameto, 2015: 2). Tujuan belajar sebenarnya adalah referensi dalam pelaksanaan proses pembelajaran. Untuk menentukan pencapaian tujuan pembelajaran diperlukan pengevaluasian hasil pembelajaran.

Guru memiliki peran penting dalam kegiatan belajar. Tugas utama guru adalah merencanakan, melaksanakan, dan mengevaluasi pembelajaran. Di sebuah sistem pembelajaran, evaluasi merupakan salah satu komponen penting yang perlu diketahui efektivitas belajarnya. Hasilnya digunakan untuk menjamin merancang evaluasi yang berkualitas. Dalam desain kurikulum terdapat tahapan evaluasi. Penilaian (asesmen) merupakan bagian dari evaluasi pencapaian siswa dan guru dalam mengajar. Asesmen akan lebih baik jika dilengkapi dengan portofolio sehingga penilaian akan lebih utuh. Selain itu, evaluasi dapat digunakan sebagai timbal balik untuk guru dalam mengoreksi dan menyelesaikan program dan kegiatan pembelajaran. Ini relevan dengan Permendikbud Nomor 54 Tahun 2013 tentang Standar Kompetensi Lulusan (SKL):

“untuk mengetahui ketercapaian dan kesesuaian antara Standar Kompetensi Lulusan dan lulusan dari masing-masing satuan pendidikan dan kurikulum yang digunakan pada satuan pendidikan tertentu perlu dilakukan monitoring dan evaluasi secara berkala dan berkelanjutan dalam setiap periode. Hasil yang diperoleh dari monitoring dan evaluasi digunakan sebagai bahan masukan bagi penyempurnaan Standar Kompetensi Lulusan di masa yang akan datang.”

Berdasarkan pasal 58 ayat (1) Undang-Undang Republik Indonesia Nomor 20 Tahun 2003 tentang Sisdiknas (Sistem Pendidikan Nasional):

“evaluasi terhadap hasil belajar siswa yang dilakukan oleh guru untuk mengamati proses, kemajuan dan revisi hasil belajar siswa terus menerus.”

Kemudian menurut Arikunto (2016: 25), evaluasi adalah akumulasi data untuk mengukur tujuan yang ada tercapai.

Definisi yang lebih luas dikemukakan oleh dua ahli lainnya, Rea-Dickins dan Germaine (2003), definisi tambahannya adalah evaluasi merupakan bagian intrinsik dari proses belajar mengajar. Ini penting bagi guru karena dapat memberikan banyak informasi untuk digunakan pada latihan di kelas di masa depan, untuk perencanaan kursus, dan untuk pengelolaan tugas belajar dan siswa.

Oleh karena itu, evaluasi perlu dilakukan oleh guru pada waktu yang sama dalam proses belajar mengajar. Tujuannya adalah memberi informasi apakah pengajaran dan pembelajaran telah berhasil atau tidak. Di sisi lain guru akan mengetahui kelemahan pengajaran yang ia lakukan di kelas jadi guru akan merevisi pengajarannya, mengatur perencanaan pelajaran, memberi tugas, dan mengendalikan kelas dengan baik.

Ada dua teknik dalam mengevaluasi hasil belajar siswa di sekolah, yaitu teknik tes dan teknik non tes. Menurut Arifin (2016: 118), tes adalah teknik atau prosedur yang digunakan untuk mengimplementasikan aktivitas pengukuran, di samping ada banyak pertanyaan atau tugas yang harus dilakukan oleh siswa untuk mengukur perilaku seorang siswa. Sementara, non tes adalah teknik atau prosedur yang tidak menggunakan pertanyaan atau tugas.

Tes lebih sempit cakupannya dibandingkan pengukuran, dan pengukuran lebih sempit dibandingkan evaluasi. Tes dibangun berdasarkan teori pengukuran tertentu. Tanpa bantuan teori pengukuran, maka pembuatan tes dapat dikatakan tidak mungkin. Bagaimana pertanyaan-pertanyaan dalam tes harus dibuat, validitas, dan realibilitas tes yang pada saat sekarang diukur berdasarkan teori *psychometric*, mencerminkan peranan teori pengukuran yang sangat besar dan penting. Pengukuran dalam psikometrik tidak lagi merupakan bagian integral ataupun hanya merupakan salah satu langkah yang mungkin dipergunakan dalam kegiatan evaluasi (Arifin, 2016: 7).

Selanjutnya, tes adalah instrumen atau alat pengukuran yang sering digunakan guru mengukur hasil belajar siswa (Arifin, 2016: 8). Guru akan

mengetahui sejauh mana tujuan pembelajaran dari hasil tes siswa. Oleh karenanya, tes yang akan digunakan harus diatur dengan benar. Selain itu, tes juga sudah mewakili semua materi pembelajaran. Untuk itu guru harus mengevaluasi tes perangkat yang telah dilakukan, seperti dengan analisis butir tes untuk menentukan kualitas tes yang telah dibuat.

Kemudian, tes yang baik harus teruji dalam dua tahap pengujian, yaitu secara teoritik dan empirik. Untuk menguji apakah tes baik secara empirik maka harus dilakukan uji coba untuk menstandarkan tes sebagai tes yang baik dengan melakukan analisis tes (Purwanto, 2016: 96). Analisis kualitas tes merupakan suatu tahap yang harus ditempuh untuk mengetahui derajat kualitas tes, baik tes secara keseluruhan maupun butir soal yang menjadi bagian dari tes tersebut. Dalam penilaian hasil belajar, tes diharapkan dapat menggambarkan sampel perilaku dan menghasilkan nilai yang objektif serta akurat. Jika tes yang digunakan guru kurang baik, maka hasil yang diperoleh pun tentunya kurang baik. Hal ini dapat merugikan peserta didik itu sendiri. Artinya, hasil yang diperoleh peserta didik menjadi tidak objektif dan tidak adil. Oleh sebab itu, tes yang digunakan guru harus memiliki kualitas yang lebih baik dilihat dari berbagai segi. Tes hendaknya disusun sesuai dengan prinsip dan prosedur penyusunan tes. Setelah digunakan perlu diketahui apakah suatu tes yang digunakan termasuk baik atau kurang baik, maka perlu dilakukan analisis kualitas tes.

Analisis kualitas tes berkaitan pada dua hal pokok, yaitu validitas dan reliabilitas. Kedua hal ini sekaligus merupakan karakteristik alat ukur yang baik. Para ahli banyak mengemukakan tentang karakteristik tersebut. R.L. Thorndike dan H.P. Hagen dalam Arifin (2016: 246) mengemukakan bahwa:

“there are many specific considerations entering into the evaluation of a test, but we shall consider them... under three main headings. These are, respectively, validity, reliability, and practicality.” Hal ini berarti terdapat banyak pertimbangan khusus yang bisa digunakan ke dalam evaluasi suatu tes. Namun, hanya akan dibatasi pada tiga karakteristik saja, yaitu validitas, reliabilitas, dan kepraktisan.

Meski demikian, menurut Feldt dan Brennan (1989) mengatakan bahwa selain valid dan reliabel tes yang baik juga tergantung dari jumlah butir-butir soal berkategori baik yang terdapat dalam tes. Semakin banyak butir soal yang baik, semakin baiklah perangkat tesnya. Sebaliknya, semakin sedikit jumlah butir soal yang baik, semakin bobrok kualitas tes itu.

Selanjutnya, menurut Arifin (2016: 68), ada delapan karakteristik instrumen butir soal yang baik, yaitu:

- 1) Valid, suatu instrumen dapat dianggap valid jika instrumen butir soal tersebut benar-benar mengukur apa yang akan diukur,
- 2) Reliabel, jika hasilnya nyata,
- 3) Relevan, jika relevan dengan standar kompetensi,

- 4) Representatif, jika instrumen mewakili semua materi pembelajaran,
- 5) Praktis, jika instrumen mudah digunakan,
- 6) Diskriminatif, instrumen harus diatur untuk menunjukkan perbedaan objek,
- 7) Spesifik, jika instrumen itu khusus untuk objek tertentu, dan
- 8) Proporsional, instrumen harus memiliki proporsional tingkat kesulitan (sulit, menengah, dan mudah).

Sebelum pengujian kualitas tes dilakukan, terlebih dulu diperiksa mutu butir-butirnya dengan melakukan analisis butir. Sudjana (2017: 135) mendefinisikan analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai. Analisis butir soal antara lain bertujuan untuk mengadakan identifikasi soal-soal yang baik, kurang baik, dan soal-soal yang tidak baik (jelek).

Sejalan dengan itu, kegiatan menganalisis butir soal merupakan suatu kegiatan yang harus dilakukan guru untuk meningkatkan mutu soal yang telah ditulis. Kegiatan ini merupakan proses pengumpulan, peringkasan, dan penggunaan informasi dari jawaban siswa untuk membuat keputusan tentang setiap penilaian (Nitko dalam Ratnawulan dan Rusdiana, 2015: 168).

Merujuk dari pendapat para ahli di atas, dapat disimpulkan bahwa analisis butir soal merupakan suatu proses yang dilakukan untuk menyelidiki, meneliti dan meninjau pertanyaan tes yang dibuat guna untuk mendapatkan kualitas pertanyaan yang memadai.

Dalam kenyataan di lapangan, tidak banyak guru yang melakukan analisis butir tes. Padahal kegiatan menganalisis butir tes merupakan kegiatan yang harus dilakukan untuk meningkatkan kualitas guru yang telah dilakukannya ujian tertulis. Sesuai dengan penelitian yang dilakukan oleh Susanto, Rinaldi, dan Novalia (2015) bahwa berdasarkan fakta di lapangan analisis butir soal masih jarang dilakukan oleh pendidik. Beberapa alasan yang menyebabkan pendidik tidak melakukan analisis butir soal antara lain: (1) pendidik merasa terbebani dalam proses analisis butir soal sehingga tidak melakukannya, (2) pendidik meyakini bahwa kualitas soal tes yang dibuat sudah baik sehingga tidak melakukan penelaahan lebih lanjut. Menurut ketua MGMP mata pelajaran Matematika kota Bandar Lampung, soal yang digunakan untuk ujian semester dibuat oleh guru yang memang ahli di bidangnya dan tidak dibuat berdasarkan analisis yang telah dilakukan. Analisis butir soal dilakukan oleh sekolah atau masing-masing guru, karena menganalisis butir soal secara keseluruhan sekota Bandar Lampung membutuhkan biaya yang besar dan waktu yang cukup banyak. Hal yang sama juga ditunjukkan pada tingkat perguruan tinggi dalam penelitian Iskandar dan Rizal (2018) yang menyatakan hasil observasi pendahuluan pada lokasi penelitian ditemukan bahwa soal yang digunakan dalam penerimaan calon mahasiswa baru tidak melalui analisis secara empirik (uji validitas, reliabilitas, tingkat kesukaran, daya beda, dan pengecoh). Hal tersebut diungkapkan salah seorang petugas pelaksana tes. Sehingga bisa dipastikan bahwa

informasi yang dikumpulkan dari tes yang diberikan, mengandung bias atau tidak sesuai dengan apa yang diinginkan.

Berdasarkan observasi yang dilakukan peneliti melalui wawancara dengan salah satu guru matematika di SMA Swasta Darussalam Medan pada tanggal 17 Juni 2020 sebelum dilakukannya Ujian Akhir Semester pada mata pelajaran matematika wajib terdapat beberapa permasalahan. Adapun permasalahan pertama yang ditemukan peneliti adalah guru belum pernah melakukan analisis butir soal secara menyeluruh dan analisis yang dilakukan hanya untuk mengetahui seberapa besar tingkat kesukarannya yaitu berapa banyak soal yang sukar, sedang, ataupun mudah. Bahkan syarat utama tes tersebut dikatakan sah dan dapat diandalkan, yakni adanya analisis validitas, reliabilitas dan lainnya jarang dilakukan. Dikarenakan guru hanya melakukan analisis tingkat kesukaran sehingga analisis tes dilakukan setelah dilaksanakan ujian akhir semester, sedangkan penganalisisan tes sebelum diujikan kepada siswa jarang dilakukan. Hal ini tentu dapat mengindikasikan bahwa kualitas soal belum diketahui, karena mengabaikan unsur validitas dan reliabilitas.

Permasalahan kedua yaitu, tim penyusun soal di sekolah tersebut tidak ada. Penyusunan soal dilakukan oleh guru yang mengampu kelasnya masing-masing, sehingga guru tersebut bertanggung jawab membuat atau menyusun soal untuk kelas tersebut dan materi yang diujikan berasal dari pembelajaran yang telah diajarkan. Hal ini dapat menimbulkan sedikitnya dilakukan perubahan soal ujian di tahun ajaran berikutnya.

Permasalahan yang terakhir adalah guru hanya melakukan beberapa perubahan butir soal, sekitar 5 (lima) dari 20 (dua puluh) butir soal yang diujikan pada ujian semester tahun ajaran berikutnya. Hal ini dikarenakan materi yang diajarkan guru tersebut sama untuk tiap tahun ajaran dan tingkatan kelas. Apabila materi yang diajarkan berbeda maka berbeda pula materi yang diujikan.

Berdasarkan permasalahan-permasalahan sebelumnya, diperlukan adanya analisis butir tes dan sebelum tes digunakan hendaknya dianalisis terlebih dahulu agar memenuhi syarat sebagai alat evaluasi yang memiliki kualitas tinggi. Hal ini dapat mempermudah guru untuk mengetahui siswa mana sudah mengerti materi yang diberikan oleh guru atau tidak dan hasil yang diperoleh peserta didik menjadi objektif dan adil.

Analisis butir tes dapat dilakukan secara kualitatif dan kuantitatif. Analisis secara teoritis atau analisis kualitatif dapat dilakukan sebelum maupun setelah dilaksanakan uji coba. Cara analisisnya adalah dengan cara mencermati butir-butir soal yang telah disusun dilihat dari kesesuaian dengan kompetensi dasar dan indikator yang diukur serta pemenuhan persyaratan baik dari ranah materi, konstruksi dan bahasa. Butir-butir soal yang akan dianalisis dapat berupa butir soal bentuk uraian, butir soal bentuk melengkapi, dan butir soal bentuk pilihan ganda (*multiple choice*). Analisis butir tes secara kuantitatif diarahkan untuk menelaah tingkat validitas butir soal, reliabilitas, daya pembeda, tingkat kesukaran, dan

khusus tipe soal pilihan ganda perlu juga ditelaah efektifitas fungsi distraktor. Validitas tes membahas apakah tes itu dapat mengukur apa yang akan diukur. Analisis validitas tes dapat dilakukan dari dua segi, yaitu: dari segi tes sebagai suatu totalitas dan dari segi itemnya, sebagai bagian tak terpisahkan dari tes secara totalitas (Ratnawulan dan Rusdiana, 2015: 191-192).

Ada beberapa alasan mengapa analisis butir tes perlu dilaksanakan. Menurut Zainul dan Nasution (2005: 23) alasan-alasan ini, antara lain:

- 1) Untuk mengetahui kekuatan dan kelemahan butir tes, sehingga dapat dilakukan seleksi dan revisi butir soal.
- 2) Untuk memberikan informasi tentang spesifikasi butir soal secara lengkap, sehingga akan lebih memudahkan pembuat soal dalam menyusun perangkat soal yang akan memenuhi kebutuhan ujian dalam bidang dan tingkat tertentu.
- 3) Untuk segera dapat mengetahui masalah yang terkandung dalam butir soal, seperti: butir soal yang ambigu, kesalahan menempatkan kunci jawaban, soal yang terlalu sukar dan terlalu mudah, atau soal yang mempunyai daya beda rendah. Masalah ini bila diketahui dengan segera akan memungkinkan bagi pembuat soal untuk membuat keputusan tentang apakah butir soal yang bermasalah itu akan digugurkan atau direvisi guna menentukan nilai siswa.
- 4) Untuk dijadikan alat guna menilai butir soal yang akan disimpan dalam kumpulan soal.
- 5) Untuk mendapatkan informasi tentang butir soal sehingga memungkinkan untuk menyusun beberapa perangkat soal yang paralel. Penyusunan perangkat seperti ini sangat bermanfaat bila akan melakukan ujian ulang atau mengukur kemampuan beberapa kelompok peserta tes dalam waktu yang berbeda.

Selanjutnya, kegiatan analisis butir tes memiliki banyak manfaat (Anastasi dan Urbina dalam Ratnawulan dan Rusdiana, 2015: 168), di antaranya adalah:

- 1) Dapat membantu para pengguna tes dalam evaluasi tes yang digunakan,
- 2) Sangat relevan bagi penyusunan tes informal dan lokal seperti tes yang disiapkan guru untuk siswa di kelas,
- 3) Mendukung penulisan butir soal yang efektif,
- 4) Secara materi dapat memperbaiki tes di kelas,
- 5) Meningkatkan validitas soal dan reliabilitas.

Selain analisis butir tes memiliki efektivitas untuk mengetahui kualitas butir soal tes, soal ujian akhir semester juga memiliki peranan penting untuk mengukur ketercapaian siswa akan kompetensi yang diajarkan, terlebih lagi sebagai salah satu penentuan kenaikan kelas bagi siswa. Beberapa faktor kemungkinan penyebab guru belum melaksanakan analisis soal ujian akhir semester adalah keterbatasan waktu dan biaya dalam menganalisis soal ujian akhir semester, kurangnya pengalaman dan pengetahuan mengenai analisis soal.

Berdasarkan keseluruhan uraian diatas, penulis tertarik untuk melakukan penelitian dengan menganalisis soal ujian akhir semester pada dua tahun ajaran

terakhir dengan menggunakan analisis butir soal berdasarkan validitas butir soal, reliabilitas, tingkat kesukaran, daya pembeda dan daya pengecoh (*distractor*) pada mata pelajaran matematika wajib yang dianalisis secara manual dan dengan bantuan aplikasi Ms. Excel. Dalam hal ini penulis akan melakukan penelitian dengan judul “**Analisis Butir Soal Ujian Akhir Semester (UAS) Matematika Wajib Kelas X dan XI di SMA Swasta Darussalam Medan**”.

1.2. Ruang Lingkup Masalah

Ada banyak hal pada analisis data yang dapat diterapkan untuk pengujian instrumen. Mereka terkait dengan validitas, reliabilitas, praktis, keaslian, *washback*, tingkat kesukaran, daya pembeda, daya pengecoh (*distractor*). Namun, peneliti memutuskan untuk membatasi aspek penelitian ini. Penelitian ini hanya fokus dalam menganalisis validitas butir soal, reliabilitas, tingkat kesukaran, daya pembeda, dan daya pengecoh (*distractor*) soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI di SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.

1.3. Rumusan Masalah

Berdasarkan ruang lingkup masalah di atas, yang menjadi rumusan masalah dalam penelitian ini adalah:

1. Bagaimana validitas butir soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir?
2. Bagaimana reliabilitas soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir?
3. Bagaimana tingkat kesukaran soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir?
4. Bagaimana indeks daya pembeda soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir?
5. Bagaimana daya pengecoh (*distractor*) soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir?

1.4. Batasan Masalah

Untuk menghindari pembahasan yang terlalu luas, penelitian ini akan difokuskan pada analisis butir soal ujian akhir semester mata pelajaran matematika

wajib kelas X dan XI di SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir, yang dianalisis secara kuantitatif berdasarkan:

1. Validitas butir soal
2. Reliabilitas
3. Tingkat kesukaran
4. Daya pembeda
5. Daya pengecoh (*distractor*)

Dalam analisis butir soal kita akan mengetahui tentang kualitas butir soal ujian akhir semester pada mata pelajaran matematika wajib kelas X dan XI di SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.

1.5. Tujuan Penelitian

Berdasarkan rumusan masalah di atas, maka tujuan dari penelitian ini adalah untuk mendeskripsikan:

1. Validitas butir soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.
2. Reliabilitas soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.
3. Tingkat kesukaran soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.
4. Indeks daya pembeda soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.
5. Daya pengecoh (*distractor*) soal ujian akhir semester mata pelajaran matematika wajib kelas X dan XI SMA Swasta Darussalam Medan selama dua tahun ajaran terakhir.

1.6. Manfaat Penelitian

Penelitian ini diharapkan dapat memberikan manfaat bagi semua pihak sebagai berikut:

1. Secara teoritis, diharapkan dapat memberikan kontribusi bagi ilmu pengetahuan dan dunia pendidikan khususnya dalam evaluasi hasil belajar agar soal yang dibuat nantinya telah memenuhi standar kualitas yang baik.
2. Bagi guru, sebagai bahan pertimbangan kepada guru mata pelajaran matematika wajib khususnya mengenai soal yang baik berdasarkan validitas butir soal, reliabilitas, tingkat kesukaran, tingkat daya beda, dan tingkat daya pengecoh (*distractor*).

3. Bagi peneliti, sebagai bahan menambah wawasan dan pengetahuan peneliti sehingga peneliti dapat menerapkan pengetahuan yang didapatkan saat memasuki dunia kerja.
4. Dapat dijadikan sebagai referensi untuk penelitian lebih lanjut pada mata pelajaran lain atau sejenis.

1.7. Definisi Operasional

Untuk menghindari terjadinya perbedaan penafsiran terhadap istilah-istilah yang terdapat dalam rumusan masalah dalam penelitian ini, perlu dikemukakan definisi operasional sebagai berikut:

1. Analisis butir soal atau analisis item adalah pengkajian pertanyaan-pertanyaan tes agar diperoleh perangkat pertanyaan yang memiliki kualitas yang memadai.
2. Validitas adalah ketepatan suatu alat ukur dalam melakukan fungsi ukurnya. Tes sebagai alat ukur hasil belajar dikatakan valid apabila tes tersebut dapat tepat mengukur hasil belajar yang hendak diukur.
3. Reliabilitas adalah derajat konsistensi hasil pengukuran. Seperangkat soal dikatakan memiliki reliabilitas tinggi apabila menunjukkan hasil yang sama atau hampir sama ketika soal di tes secara berulang pada kelompok yang sama di waktu yang berbeda.
4. Tingkat kesukaran adalah peluang menjawab benar suatu soal pada tingkat kemampuan tertentu siswa. Soal yang baik memiliki tingkat kesukaran yang sedang.
5. Daya pembeda adalah kemampuan soal dalam membedakan siswa yang pintar dengan yang kurang pintar. Dalam hal ini, siswa yang menguasai materi pembelajaran dengan yang tidak menguasai materi pembelajaran.
6. Pengecoh (*distactor*) yang juga dikenal dengan istilah penyesat atau penggoda adalah pilihan jawaban yang bukan merupakan kunci jawaban. Pengecoh diadakan untuk menyesatkan siswa agar tidak memilih kunci jawaban pengecoh dan menggoda siswa yang kurang begitu memahami materi pelajaran untuk memilihnya.
7. Ujian Akhir Sekolah (UAS) adalah bagian dari bentuk evaluasi yang bertujuan untuk mengukur dan menilai kompetensi siswa sehingga guru bisa menentukan apakah siswa dapat melanjutkan pembelajaran pada tingkat yang lebih tinggi atau perlu adanya pengujian kembali.