

BAB I

KONSEP DASAR PENILAIAN BERBASIS KOMPETENSI

A. Pengertian Penilaian

Penilaian adalah proses sistematis meliputi pengumpulan informasi (angka atau deskripsi verbal), analisis, dan interpretasi untuk mengambil keputusan. Sedangkan penilaian pendidikan adalah proses pengumpulan dan pengolahan informasi untuk menentukan pencapaian hasil belajar peserta didik.

Untuk itu, diperlukan data sebagai informasi yang diandalkan sebagai dasar pengambilan keputusan. Dalam hal ini, keputusan berhubungan dengan sudah atau belum berhasilnya peserta didik dalam mencapai suatu kompetensi. Jadi, penilaian merupakan salah satu pilar dalam pelaksanaan Kurikulum Tingkat Satuan Pendidikan (KTSP) yang berbasis kompetensi.

Penilaian merupakan suatu proses yang dilakukan melalui langkah-langkah perencanaan, penyusunan alat penilaian, pengumpulan informasi melalui sejumlah bukti yang menunjukkan pencapaian hasil belajar peserta didik, pengolahan, dan penggunaan informasi tentang hasil belajar peserta didik. Penilaian dilaksanakan melalui berbagai bentuk antara lain: penilaian unjuk kerja (*performance*), penilaian sikap, penilaian tertulis (*paper and pencil test*), penilaian proyek, penilaian melalui kumpulan hasil kerja/karya peserta didik (*portfolio*), dan penilaian diri.

Penilaian hasil belajar baik formal maupun informal diadakan dalam suasana yang menyenangkan, sehingga memungkinkan peserta didik menunjukkan apa yang dipahami dan mampu dikerjakannya. Hasil belajar seorang peserta didik tidak dianjurkan untuk dibandingkan dengan peserta didik lainnya, tetapi dengan hasil yang dimiliki peserta didik tersebut sebelumnya. Dengan demikian peserta didik tidak merasa dihakimi oleh guru tetapi dibantu untuk mencapai apa yang diharapkan.

B. Prinsip Penilaian

Dalam melaksanakan penilaian mempertimbangkan prinsip-prinsip sebagai berikut.

1. Memandang penilaian dan kegiatan pembelajaran secara terpadu.
2. Mengembangkan strategi yang mendorong dan memperkuat penilaian sebagai cermin diri.
3. Melakukan berbagai strategi penilaian di dalam program pembelajaran untuk menyediakan berbagai jenis informasi tentang hasil belajar peserta didik.
4. Mempertimbangkan berbagai kebutuhan khusus peserta didik.
5. Mengembangkan dan menyediakan sistem pencatatan yang bervariasi dalam pengamatan kegiatan belajar peserta didik.
6. Menggunakan cara dan alat penilaian yang bervariasi. Penilaian dapat dilakukan dengan cara tertulis, lisan, produk portofolio, unjuk kerja, proyek, dan pengamatan tingkah laku.
7. Melakukan penilaian secara berkesinambungan untuk memantau proses, kemajuan, dan perbaikan hasil, dalam bentuk: ulangan harian, ulangan tengah semester, ulangan akhir semester, dan ulangan kenaikan kelas. Ulangan harian dapat dilakukan bila sudah menyelesaikan satu atau beberapa indikator atau satu kompetensi dasar (KD), ulangan tengah semester dilakukan bila telah menyelesaikan beberapa KD atau satu stándar kompetensi (SK), ulangan akhir semester dilakukan setelah menyelesaikan semua KD atau SK semester bersangkutan, sedangkan ulangan kenaikan kelas dilakukan pada akhir semester genap dengan menilai semua SK semester ganjil dan genap, dengan penekanan pada semester genap.
8. Penilaian kompetensi pada uji kompetensi melibatkan pihak sekolah dan Institusi Pasangan/Asosiasi Profesi, dan pihak lain terutama DU/DI. Idealnya, lembaga yang menyelenggarakan uji kompetensi ini independen; yakni lembaga yang tidak dapat diintervensi oleh unsur atau lembaga lain.

Agar penilaian objektif, pendidik harus berupaya secara optimal untuk (1) memanfaatkan berbagai bukti hasil kerja peserta didik dari sejumlah penilaian, (2) membuat keputusan yang adil tentang penguasaan kompetensi peserta didik dengan mempertimbangkan hasil kerja (karya).

C. Kegunaan Penilaian

Kegunaan penilaian antara lain sebagai berikut:

1. Memberikan umpan balik bagi peserta didik agar mengetahui kekuatan dan kelemahan dirinya dalam proses pencapaian kompetensi.
2. Memantau kemajuan dan mendiagnosis kesulitan belajar yang dialami peserta didik sehingga dapat dilakukan pengayaan dan remedial.
3. Untuk umpan balik bagi pendidik/guru dalam memperbaiki metode, pendekatan, kegiatan, dan sumber belajar yang digunakan.
4. Memberikan informasi kepada orang tua dan komite sekolah tentang efektivitas pendidikan.
5. Memberi umpan balik bagi pengambil kebijakan (Dinas Pendidikan Daerah) dalam meningkatkan kualitas penilaian yang digunakan.

D. Fungsi Penilaian

Penilaian memiliki fungsi untuk:

1. Menggambarkan sejauhmana peserta didik telah menguasai suatu kompetensi.
2. Mengevaluasi hasil belajar peserta didik dalam rangka membantu memahami dirinya, membuat keputusan tentang langkah berikutnya, baik untuk perencanaan program belajar, pengembangan kepribadian, maupun untuk penjurusan (sebagai bimbingan).
3. Menemukan kesulitan belajar, kemungkinan prestasi yang bisa dikembangkan peserta didik, dan sebagai alat diagnosis yang membantu pendidik/guru menentukan apakah seseorang perlu mengikuti remedial atau pengayaan.

4. Menemukan kelemahan dan kekurangan proses pembelajaran yang sedang berlangsung guna perbaikan proses pembelajaran berikutnya.
5. Pengendali bagi pendidik/guru dan sekolah tentang kemajuan perkembangan peserta didik.

E. Jenis-Jenis Penilaian

Berdasarkan Peraturan Pemerintah No. 19 Tahun 2005 tentang Standar Nasional Pendidikan dan Permendiknas No. 20 Tahun 2007 tentang Standar Penilaian Pendidikan, jenis penilaian dan bentuk pengadministrasiannya diuraikan seperti tabel berikut.

Tabel Jenis-jenis Penilaian

Penilai	No	Jenis Penilaian	Unsur yang terlibat	Ruang lingkup materi	Bentuk Administrasi Penilaian	
					Produktif	Normatif dan Adaptif
Pendidik	1	Ulangan Harian (Penilaian pro-ses akhir KD/tatap muka)	Guru	KD	KHS	KHS
	2	Ulangan Tengah Semester (Penilaian akhir beberapa KD atau akhir sebuah SK)	Guru (Internal/QA) dan Unsur Eksternal/ QC	Beberapa KD atau SK	KHS/Skill Passport	KHS
	3	Ulangan Akhir Semester Ganjil (komprehensif, seluruh kompe-tensi dalam satu semester)	Guru, dan Unsur Eksternal	Dapat berupa beberapa KD atau SK	<ul style="list-style-type: none"> ◆ KHS/ Skills ◆ Passport ◆ Laporan Hasil Belajar ◆ Leger 	<ul style="list-style-type: none"> ◆ Raport ◆ Leger
Pendidik (Satuan Pendidikan)	1	Ulangan Kenaikan Kelas/ akhir semester genap	Guru dan Unsur Eksternal	SKL yang dipelajari pada tahun yang bersangkutan	<ul style="list-style-type: none"> ◆ KHS/Skill Passport ◆ Laporan Hasil Belajar ◆ Transkrip ◆ Leger 	<ul style="list-style-type: none"> ◆ Raport ◆ Leger

Penilai	No	Jenis Penilaian	Unsur yang terlibat	Ruang lingkup materi	Bentuk Administrasi Penilaian	
					Produktif	Normatif dan Adaptif
	2	Ujian Sekolah	<ul style="list-style-type: none"> ◆ Sekolah, Pemerintah ◆ (Internal/QA dan atau Eksternal/QC) 	Mata pelajaran yang tidak diujikan dalam UN untuk seluruh SKL yang sudah diajarkan	<ul style="list-style-type: none"> ◆ KHS/ Skills ◆ Passport ◆ Laporan Hasil Belajar ◆ Transkrip ◆ Ijazah ◆ Leger 	<ul style="list-style-type: none"> ◆ Ijazah ◆ Transkrip ◆ Leger
Pemerintah	1	Ujian Nasional (UN)	Pemerintah dan Du/Di	Seluruh SKL Ujian Nasional	<ul style="list-style-type: none"> ◆ Transkrip ◆ Ijazah ◆ SKHUN ◆ Sertifikat Kompetensi 	<ul style="list-style-type: none"> ◆ Ijazah ◆ SKHUN ◆ Leger

F. Kriteria Penilaian

1. Validitas

Validitas berarti menilai apa yang seharusnya dinilai dengan menggunakan alat yang sesuai untuk mengukur kompetensi. Misal, dalam pelajaran bahasa Indonesia, pendidik/guru ingin menilai kompetensi berbicara. Bentuk penilaian valid jika menggunakan tes lisan. Jika menggunakan tes tertulis penilaian tidak valid.

2. Reliabilitas

Reliabilitas berkaitan dengan konsistensi (keajegan) hasil penilaian. Penilaian yang *reliable* (ajeg) memungkinkan perbandingan yang *reliable* dan menjamin konsistensi. Misal, guru menilai dengan proyek, penilaian akan reliabel jika hasil yang diperoleh itu cenderung sama bila proyek itu dilakukan lagi dengan kondisi yang relatif sama. Untuk menjamin penilaian yang reliabel petunjuk pelaksanaan proyek dan penskorannya harus jelas.

3. Berfokus pada kompetensi

Dalam pelaksanaan kurikulum tingkat satuan pendidikan yang berbasis kompetensi, penilaian harus terfokus pada pencapaian kompetensi (rangkaiian kemampuan), bukan hanya pada penguasaan materi (pengetahuan).

4. Menyeluruh/Komprehensif

Penilaian harus menyeluruh dengan menggunakan beragam cara dan alat untuk menilai beragam kompetensi atau kemampuan peserta didik, sehingga tergambar profil kemampuan peserta didik.

5. Objektivitas

Penilaian harus dilaksanakan secara objektif. Untuk itu, penilaian harus adil, terencana, berkesinambungan, dan menerapkan kriteria yang jelas dalam pemberian skor.

6. Mendidik

Penilaian dilakukan untuk memperbaiki proses pembelajaran bagi pendidik/guru dan meningkatkan kualitas belajar bagi peserta didik.

Penilaian Hasil Belajar Kelompok Mata Pelajaran adalah sebagai berikut:

1. Penilaian hasil belajar kelompok mata pelajaran agama dan akhlak mulia serta kelompok mata pelajaran kewarganegaraan dan kepribadian dilakukan melalui:
 - a. Pengamatan terhadap perubahan perilaku dan sikap untuk menilai perkembangan afeksi dan kepribadian peserta didik.
 - b. Ujian, ulangan, dan/atau penugasan untuk mengukur aspek kognitif peserta didik.

BAB II

Karakteristik Penilaian di Sekolah Menengah Kejuruan

Karakteristik penilaian yang digunakan di SMK harus mempunyai tujuan untuk mendorong peningkatan kualitas pembelajaran yang secara langsung dapat berdampak pada peningkatan kualitas pendidikan. Pada tiap lembaga pendidikan mempunyai sistem penilaian yang mampu memberikan informasi yang akurat yang meliputi kompetensi dasar yang telah dicapai dan yang belum dicapai peserta didik. Dengan hasil penilaian yang telah diketahui maka dapat menemukan strategi pembelajaran yang tepat bagi peserta didik dan pendidik. Sehingga diharapkan akan berdampak pada kinerja lembaga meningkat dan diharapkan akan meningkatkan kualitas pendidikan.

Penilaian yang dapat meningkatkan kualitas pendidikan adalah sistem penilaian berkelanjutan yang mempunyai prinsip menilai seluruh kompetensi dasar, menganalisis hasil penilaian dan melakukan tindak lanjut berupa program pengayaan atau perbaikan. Penilaian yang dimaksud adalah penilaian berbasis kompetensi yang mencakup kemampuan kognitif, afektif dan psikomotor.

Aspek psikomotor terletak pada ketepatan gerakan yang dilakukan oleh peserta didik dilihat dari penampilan peserta didik dalam melakukan praktek dengan fokus penilaian terletak pada gerakan, waktu, hasil yang dicapai dan keselamatan kerja.

Penilaian kognitif menekankan pada kemampuan atau pengetahuan yang harus dikuasai, sedangkan penilaian aspek afektif meliputi kelakuan, kebersihan, dan kerajinan. Ketiga aspek di atas merupakan bagian dari kompetensi, oleh karena itu penilaian berbasis kompetensi menekankan pada keadaan sebenarnya yaitu kompetensi dasar yang benar-benar dimiliki oleh peserta didik.

Terdapat sepuluh prinsip penilaian berbasis kompetensi, antara lain:

- 1). **Valid dan reliabel:** apabila hasil penilaian akurat dan ajeg atau konsisten. Hal ini dapat dicapai apabila instrumen yang digunakan baik dan tepat. Cara yang paling mudah untuk mendapatkan instrumen yang valid dan reliabel apabila butir-butir soal ditulis berdasarkan indikator, dilakukan kaji ulang terhadap kalimat-kalimat yang sudah dituliskan agar tidak terjadi kesalahan serta menggunakan tata bahasa yang mudah dipelajari oleh peserta didik.
- 2). **Mendidik:** apabila penilaian guru mampu mendidik agar peserta didik dapat termotivasi untuk belajar lebih baik. Hal ini dapat dilakukan dengan cara guru

memberikan hasil pekerjaan peserta didik dan mengarsip secara lengkap hasil penilaian sehingga obyektivitas penilaian terjamin.

- 3). **Adil:** apabila ditinjau dari butir-butir soal dan cara pemberian skor. Hal ini dapat dilakukan dengan cara tidak mengujikan materi-materi yang belum diajarkan kepada peserta didik.
- 4). **Menyenangkan:** melakukan penilaian dalam suasana yang menyenangkan tidak menimbulkan peserta didik merasa takut dan stres.
- 5). **Mengacu kepada kompetensi:** dalam melakukan penilaian acuan yang dipakai adalah kesatuan kompetensi bukan waktu.
- 6). **Meyeluruh:** penilaian dilakukan secara menyeluruh, mencakup aspek kognitif, afektif dan psikomotor.
- 7). **Berkesinambungan:** hasil evaluasi harus dianalisis untuk mengetahui kompetensi yang belum dikuasai peserta didik sehingga bisa diberikan refleksi dan perbaikan proses pembelajaran di kelas.
- 8). **Mengakui kompetensi yang telah dimiliki sebelumnya:** memungkinkan peserta didik pindah kelas tanpa harus mengulang kompetensi sebelumnya.
- 9). **Menggunakan acuan kriteria:** nilai yang diberikan kepada peserta didik berdasarkan perbandingan dengan standar mutlak.
- 10). **Bisa menggunakan external assesor:** untuk menguji kompetensi yang telah dikuasai peserta didik, sehingga kredibilitas hasil pegujian dapat dipertanggungjawabkan (Djemari Mardapi: 2004).

b. Pengembangan Instrumen Penilaian Berbasis Kompetensi

Penilaian berbasis kompetensi menekankan pada kompetensi dasar yang harus dimiliki oleh peserta didik. Kompetensi dasar tersebut dibandingkan dengan standar atau kriteria yang telah ditetapkan sebelumnya. Hasil penilaian berbasis kompetensi adalah lulus dan belum lulus. Lulus berarti peserta didik telah memiliki kompetensi dasar, yaitu sama atau lebih besar dari standar. Peserta didik yang belum lulus berarti kemampuan yang dimiliki belum mencapai standar.

Alat ukur yang digunakan dalam penilaian berbasis kompetensi diusahakan memberikan informasi yang sahih dan handal. Kriteria keberhasilan dalam melaksanakan program pembelajaran adalah jumlah kompetensi dasar yang dicapai oleh peserta didik. Pengembangan sistem penilaian berbasis kompetensi hasil kegiatan belajar mengajar harus bersifat hirarki, secara berurutan yaitu: standar

kompetensi, kompetensi dasar, standar materi, indikator dan soal ujian. Hal yang penting dalam mengembangkan sistem penilaian adalah menyusun spesifikasi penilaian, meliputi: tujuan, lama penilaian dan instrumen penilaian

1). Pengembangan Instrumen Penilaian Bentuk Tes

Penguasaan peserta didik dalam aspek pengetahuan tentang suatu pekerjaan merupakan faktor yang penting. Metode yang lazim digunakan dalam mengukur aspek pengetahuan adalah melalui tes tertulis dan lisan. Scott (1993) menjelaskan berbagai variasi yang mungkin dilakukan dalam mengembangkan tes tertulis, yaitu: *multiple choice, sentence completion, listing, true-false, matching, essay, dan modified form*. Sementara itu berdasarkan Depdiknas (2001) pengembangan tes lisan terdiri dari: wawancara terstruktur, tak terstruktur dan pemecahan masalah.

Tujuan utama dalam penilaian aspek pengetahuan ini adalah mengembangkan proses belajar mengajar. Selain itu juga memberikan motivasi kepada peserta didik, mendiagnosis kemampuan peserta didik, mengembangkan pengajaran efektif, memecahkan masalah, dan memberikan informasi kepada peserta didik untuk mengembangkan diri.

Hal-hal yang perlu dipersiapkan oleh guru dalam merancang tes adalah: a). menentukan dan mengembangkan jumlah dan jenis item tes; b). menentukan dan memilih tujuan yang akan diukur; c). Menganalisa tujuan dan menentukan isi tes; d). Mengembangkan garis besar isi untuk konstruksi item tes; e). Mengkonstruksi item tes; f). Membuat tabel perencanaan untuk memilih item tes.

Tes yang akan digunakan untuk mengukur aspek pengetahuan harus memenuhi persyaratan sebagai berikut:

a). Hanya mengukur satu dimensi (*Unidimensionality*)

Suatu tes yang mengukur suatu bidang studi tertentu, setiap butir soal pada perangkat tes hanya mengukur satu bidang saja. Dengan dipenuhinya persyaratan ini, maka tes tersebut valid.

b). Keandalan (*reliabilitas*)

Keandalan tes meliputi kecermatan (*precision*) dan keajegan (*consistency*) dari hasil pengukuran. Sebuah tes dengan jumlah soal yang banyak dan seluruh soal bertaraf kesukaran sedang bagi peserta didik yang menempuh ujian, akan menghasilkan informasi yang lebih teliti mengenai orang yang diuji jika

dibandingkan dengan tes yang soalnya sedikit dan tingkat kesukaran rendah atau tinggi. Soal-soal tes tidak boleh terlalu jauh di bawah atau di atas kemampuan tingkat pencapaian hasil belajar peserta didik, dan tingkat kesukaran butir-butir soal sebaiknya homogen (Depdikbud: 1999).

Menurut Allen (1979) tahapan dalam pengembangan perangkat tes adalah:

a). *Plan the test*

Rumuskan tujuan tes yang akan dilakukan. Tujuan tes harus dirumuskan secara jelas sehingga memberikan arah dan lingkup pengembangan tes selanjutnya. Setelah tujuan dirumuskan, buat kisi-kisi tes (*table of specification*).

b). *Write item for each of the areas in the plan*

Penulisan soal adalah penjabaran indikator kompetensi yang hendak diukur menjadi pertanyaan-pertanyaan yang karakteristiknya hendak diukur sesuai dengan pertanyaan-pertanyaan yang karakteristiknya sesuai dengan kisi-kisi. Setiap butir soal yang dibuat harus jelas apa yang akan ditanyakan dan jelas pula apa yang dituntut. Pada soal yang telah dibuat, dilakukan review dan revisi oleh orang lain, karena seringkali kekurangan yang terdapat pada soal tidak terlihat oleh penulis.

c). *Administer all the item to a reasonably large sample of at least 50 examinees*

Soal-soal yang telah direview dan direvisi secara teoritis sudah merupakan soal yang baik. Guna memperoleh gambaran empirik perangkat tes perlu diujicobakan pada kelompok subyek yang memiliki karakteristik sama atau hampir sama dengan kelompok dimana tes tersebut akan digunakan.

d). *Conduct an item analysis*

Analisis butir hasil uji coba bertujuan untuk memperoleh informasi soal-soal mana yang sudah baik, perlu direvisi dan harus dibuang analisis butir dilakukan berdasarkan teori tes klasik ataupun modern.

e). *Administer the revised test the another representative sample of examinee*

Ujikan kembali tes yang sudah direvisi pada sample lain yang memiliki karakteristik peserta tes yang akan diuji, kemudian ulangi lagi langkah keempat. Langkah ini untuk selanjutnya disebut dengan *test standardization*.

Macam-macam tes untuk menilai hasil belajar antara lain:

- a). Tes Obyektif adalah bentuk soal atau tes yang telah mengandung kemungkinan jawaban yang harus dipilih atau dikerjakan oleh peserta tes. Secara umum ada tiga tipe tes obyektif:

- 1). Benar-Salah (*true-false*)
 - 2). Menjodohkan (*matching*)
 - 3). Pilihan ganda (*multiple choice*)
- b). Tes Non-Objektif adalah jenis tes yang mengandung pertanyaan atau tugas dimana jawaban soal tersebut dilakukan dengan cara mengekspresikan pikiran peserta tes. Jawaban soal tidak disediakan oleh pembuat tes. Tes Non-Objektif meliputi:
- 1). Isian/melengkapi
 - 2). Jawaban singkat
 - 3). Uraian/ essay

Hal-hal yang perlu diperhatikan dalam penyajian adalah tes adalah waktu penyajian, petunjuk pengerjaan soal yang jelas, ruangan dan tempat duduk peserta tes. Sejalan dengan hal di atas Depdiknas (2002) mengemukakan tahapan dalam mengembangkan tes prestasi hasil belajar, yaitu: (1). Menyusun spesifikasi tes; (2). Menulis butir soal; (3). Menelaah butir soal; (4). Merakit butir soal; (5). Pembakuan alat melalui uji coba; (6). Menganalisis butir soal tes; (7). Memperbaiki tes; (8). Merakit kembali butir soal; (9). Melaksanakan tes; (10). Menafsirkan hasil tes.

2). Pengembangan Instrumen Penilaian Bentuk Non Tes

Instrumen non tes digunakan untuk mengukur perubahan tingkah laku yang berhubungan dengan aspek psikomotor dan afektif terutama yang berhubungan dengan apa yang dikerjakan oleh peserta didik. Dengan kata lain instrumen ini digunakan untuk mengukur penampilan yang dapat diamati dengan menggunakan indera. Instrumen non tes merupakan bagian dari keseluruhan instrumen penilaian hasil belajar, yang umum digunakan adalah: (a) bagan partisipasi; (b) daftar cek (c) skala tujuan; dan (d) skala sikap. Dalam penilaian aspek psikomotor atau keterampilan yang diamati adalah kemampuan dan penampilannya.

Penilaian aspek psikomotor peserta didik dilakukan dengan penilaian psikomotor (*performance assesment*). Menurut Trespeces (Depdiknas: 2003) *performance assesment* adalah berbagai tugas dan situasi dimana peserta tes diminta untuk mendemonstrasikan pemahaman dan mengaplikasikan pengetahuan yang mendalam, serta psikomotor di dalam berbagai macam konteks. Dengan kata lain *performance assesment* merupakan suatu penilaian yang meminta peserta tes

untuk mendemonstrasikan dan mengaplikasikan pengetahuan ke dalam berbagai konteks sesuai dengan kriteria yang diinginkan.

Untuk mengevaluasi apakah penilaian psikomotor sudah dianggap berkualitas, maka paling tidak harus diperhatikan tujuh kriteria yang dibuat oleh Popham (1995). Kriteria-kriteria tersebut adalah:

- 1). *Generalizability*, artinya apakah psikomotor peserta tes dalam melakukan tugas yang diberikan sudah memadai untuk digeneralisasikan kepada tugas-tugas lain. Semakin dapat digeneralisasikan tugas-tugas tersebut dalam penilaian psikomotor maka semakin baik tugas tersebut.
- 2). *Authenticity*, artinya apakah tugas yang diberikan sudah serupa dengan kehidupan sehari-hari.
- 3). *Multiple foci*, apakah tugas yang diberikan sudah mengukur lebih dari satu kemampuan yang diinginkan.
- 4). *Teachability*, artinya tugas yang diberikan merupakan tugas yang hasilnya semakin baik karena adanya usaha pembelajaran. Jadi tugas-tugas yang diberikan dalam penilaian psikomotor adalah tugas yang relevan dapat diajarkan oleh dosen.
- 5). *Fairness*, artinya apakah tugas-tugas yang diberikan sudah adil untuk semua kelompok dalam peserta tes. Tugas yang diberikan harus bersifat adil untuk semua jenis kelamin, status sosial, agama dan suku bangsa.
- 6). *Feasibility*, artinya apakah tugas-tugas dalam penilaian psikomotor sudah mempertimbangkan aspek-aspek biaya, ruang, waktu atau peralatan yang digunakan.
- 6). *Scorability*, artinya apakah tugas yang diberikan dapat di skor dengan akurat dan reliabel.

Tahap-tahap pengembangan penilaian aspek psikomotor, antara lain sebagai berikut:

- 1). Identifikasi semua langkah-langkah penting yang diperlukan atau yang akan mempengaruhi hasil akhir yang terbaik.
- 2). Tuliskan semua perilaku kemampuan spesifik untuk menyelesaikan tugas dan hasil akhir yang terbaik.
- 3). Membuat kriteria kemampuan yang akan diukur tidak terlalu banyak agar dapat diobservasi dengan baik.
- 4). Definisikan dengan jelas kriteria atau produk yang akan dihasilkan.

5). Urutkan semua kriteria kemampuan berdasarkan urutan yang diamati.

Hal yang penting pada penilaian psikomotor adalah cara mengamati dan menskor kemampuan peserta didik. Penilaian psikomotor dapat dilakukan dengan dua pendekatan, yaitu: (a). Metode *holistic*, dengan menggunakan satu skor (*single rating*) terhadap keseluruhan hasil psikomotor; (b). Metode *analytic*, dengan memberikan skor pada berbagai aspek yang berbeda berhubungan dengan psikomotor yang akan dinilai.

Menurut Depdiknas (2003) sikap adalah kumpulan hasil evaluasi seseorang terhadap obyek, orang atau masalah tertentu. Kompetensi aspek sikap yang harus dicapai dalam hasil pembelajaran meliputi tingkat pemberian respon, apresiasi, penilaian dan internalisasi. Penilaian aspek sikap sebaiknya lebih ditekankan kepada sikap kerja yang terintegrasi dalam pelaksanaan penilaian aspek psikomotor, dengan tidak mengabaikan aspek sikap lain selama proses pembelajaran.

Tahap dalam mengembangkan instrumen aspek sikap (Depdiknas: 2003) adalah: (1). Menentukan definisi konseptual atau konstruk yang akan diukur; (2). Menentukan definisi operasional; (3). Menentukan indikator; (4). Menulis instrumen. Penilaian sikap dapat dilakukan dengan beberapa cara, antara lain: observasi, perilaku, pertanyaan langsung, laporan pribadi atau menggunakan skala sikap dengan model pengembangan skala Likert.

Syarat utama penilaian adalah diperolehnya data hasil pengukuran dengan tingkat akurasi yang tinggi sesuai dengan kompetensi dasar, materi pokok, dan indikator. Untuk memperoleh hasil penilaian yang akurat diperlukan alat ukur yang valid dan reliabel, sehingga harus mengikuti langkah-langkah pengembangan instrumen secara ketat. Melalui pengembangan sistem penilaian berbasis kompetensi diharapkan dapat meningkatkan kualitas hasil pembelajaran.

C. Pembakuan Instrumen Penilaian Berbasis Kompetensi

Salah satu prinsip penilaian berbasis kompetensi adalah instrumen harus valid dan reliabel. Khususnya penilaian berbentuk tes, apalagi jika tes tersebut digunakan dalam skala besar, pengambilan keputusan yang mendasar, dan berdampak luas sehingga syarat valid dan reliabel harus disertai dengan uji statistik. Dengan cara tersebut dapat dihasilkan instrumen penilaian yang baku.

1). Uji Validitas Instrumen Penilaian

Secara umum, instrumen dapat dikatakan baik apabila instrumen tersebut valid atau dengan kata lain dapat mengukur apa yang seharusnya diukur. Menurut Gronlund (1990) validitas dapat dibedakan menjadi tiga, yaitu: validitas isi, validitas konstruk dan validitas kriteria. Validitas isi mengacu pada sejauhmana materi tes tersebut dapat mengukur keseluruhan materi yang telah diajarkan. Validitas konstruk mengacu pada sejauh mana instrumen tersebut dapat mengungkapkan keseluruhan konstruk yang digunakan sebagai dasar penyusunan tes. Jika suatu tes dimaksudkan untuk memprediksi keberhasilan seseorang di masa yang akan datang atau untuk mengetahui kesesuaian antara pengetahuan dan psikomotor yang dimiliki maka digunakan validitas kriteria.

Perhitungan validitas instrumen yang berbentuk non tes menggunakan bantuan komputer program *SPSS 11,5 for Windows*. Berdasarkan hasil perhitungan apabila nilai *corrected item-total correlation* tersebut $\geq 0,30$ maka butir pernyataan dalam angket dinyatakan valid. Untuk menentukan validitas instrumen yang berupa tes, analisis validitas menggunakan **Program Iteman**, meliputi tingkat kesukaran, daya beda dan distraktor. Adapun secara terperinci dapat dijelaskan sebagai berikut:

a). Tingkat Kesukaran (*Difficulty level*)

Tingkat kesukaran butir soal dihitung berdasarkan proporsi peserta tes menjawab benar terhadap butir soal tersebut. Tingkat kesukaran butir soal dilambangkan dengan p . Makin besar nilai p berarti makin besar proporsi yang menjawab benar terhadap butir soal tersebut sehingga makin rendah tingkat kesukaran butir soal itu. Tingkat kesukaran butir menunjukkan bahwa butir soal itu sukar atau mudah untuk kelompok peserta tes tertentu. Dalam analisis ini digunakan *proportion correct* (p), untuk menilai tingkat kesukaran butir soal.

Besarnya tingkat kesukaran berkisar antara 0,00 sampai 1,00. Untuk sederhananya, tingkat kesukaran butir dapat dibagi menjadi tiga kelompok.

Adapun kriteria yang digunakan dalam penelitian ini, dijelaskan dalam tabel sebagai berikut:

Tabel 8.
Klasifikasi Tingkat Kesukaran Butir Soal

Tingkat Kesukaran	Nilai p
Sukar	0,00 – 0,25
Sedang	0,26 – 0,75
Mudah	0,76 – 1,00

b). Daya beda

Daya beda butir soal dalam penelitian ini adalah indeks yang menunjukkan tingkat kemampuan butir soal membedakan kelompok yang berprestasi tinggi (kelompok atas) dari kelompok yang berprestasi rendah (kelompok bawah) diantara para peserta uji coba instrumen. Tujuan pokok mencari daya beda adalah untuk menentukan apakah butir tersebut memiliki kemampuan membedakan kelompok dalam aspek yang diukur, sesuai dengan perbedaan yang ada pada kelompok itu.

Dalam analisis ini digunakan nilai koefisien korelasi biserial untuk menentukan daya beda butir soal. Koefisien korelasi biserial menunjukkan hubungan antara dua skor, yaitu skor butir soal dan skor keseluruhan dari peserta tes yang sama.

Koefisien daya beda berkisar antara $-1,00$ sampai dengan $+1,00$. Daya beda $+1,00$ berarti bahwa semua anggota kelompok atas menjawab benar terhadap butir soal itu, sedangkan kelompok bawah seluruhnya menjawab salah terhadap butir soal itu. Sebaliknya daya beda $-1,00$ berarti bahwa semua anggota kelompok atas menjawab salah butir soal itu, sedangkan kelompok bawah seluruhnya menjawab benar terhadap soal itu.

Kriteria besarnya koefisien daya beda dalam penelitian ini diklasifikasikan menjadi empat kategori. Secara lebih rinci dijelaskan dalam tabel di bawah ini.

Tabel 9.
Klasifikasi Daya Beda Butir Soal

Kategori Daya Beda	Koefisien Korelasi
Baik	0,40 – 1,00
Sedang (tidak perlu revisi)	0,30 – 0,39
Perlu direvisi	0,20 – 0,29
Tidak baik	-1,00 – 0,19

c). Distribusi Jawaban (*Distraktor*)

Alternatif jawaban (*distraktor*) terdiri dari dua bagian, yaitu kunci jawaban dan pengecoh. Pengecoh dikatakan berfungsi apabila semakin rendah tingkat kemampuan peserta tes semakin banyak memilih pengecoh, atau makin tinggi tingkat kemampuan peserta tes akan semakin sedikit memilih pengecoh. Proporsi alternatif jawaban masing-masing butir soal dapat dilihat pada kolom *proportion endorsing* pada hasil analisis iteman.

Untuk menilai pengecoh (*distraktor*) dari masing-masing butir soal dapat dikategorikan sebagai berikut:

Tabel 10.
Klasifikasi Distraktor Butir Soal

Kategori Distraktor	Nilai <i>Proportion Endorsing</i>
Baik	$\geq 0,025$
Revisi	$< 0,025$
Tidak Baik / Tolak	0,000

Berdasarkan uraian di atas, menurut pandangan teori tes klasik secara empiris mutu butir soal ditentukan oleh statistik butir soal yang meliputi : tingkat kesukaran, daya beda dan efektifitas *distraktor*. Untuk menentukan kualitas butir soal menggunakan kriteria sebagai berikut:

Tabel 11.

Klasifikasi Kualitas Butir Soal

Kategori	Kriteria Penilaian
Diterima	Apabila (1). Tingkat kesukaran $0,25 \leq p \leq 0,75$, (2). Korelasi biserial butir soal $\geq 0,40$ dan (3). Korelasi biserial alternatif jawaban (distraktor) bernilai negatif.
Direvisi	Apabila (1). Tingkat kesukaran $p < 0,25$ atau $p > 0,75$ tetapi korelasi biserial butir $\geq 0,40$ dan korelasi biserial distraktor bernilai negatif, (2). Tingkat kesukaran $0,25 \leq p \leq 0,75$ dan korelasi biserial butir soal $\geq 0,40$ tetapi ada korelasi biserial pada distraktor yang bernilai positif, (3). Tingkat kesukaran $0,25 \leq p \leq 0,75$ dan korelasi biserial butir soal antara 0,20 sampai 0,30 tetapi korelasi distraktor bernilai negatif selain kunci atau tidak ada yang lebih besar nilainya dari kunci jawaban.
Ditolak	Apabila (1). Tingkat kesukaran $p < 0,25$ atau $p > 0,75$ dan ada korelasi biserial pada distraktor bernilai positif, (2). Korelasi biserial butir soal $< 0,20$, (3). Korelasi biserial butir soal $< 0,30$ dan korelasi biserial distraktor bernilai positif.

2). Uji Reliabilitas Instrumen Penilaian

Instrumen dikatakan reliabel (konsisten atau andal) apabila hasil pengukuran menunjukkan sejauhmana dapat memberikan hasil yang relatif sama bila dilakukan pengukuran kembali terhadap subyek yang sama. Reliabilitas berlaku untuk semua butir yang ada dalam soal atau perangkat soal.

Perhitungan reliabilitas instrumen non tes dapat menggunakan formula Alpha dari Cronbach dengan bantuan Komputer program *SPSS 11,5 for Windows*, dengan Uji Reliabilitas (*Reliability Scale-Alpha*). Analisis statistik, untuk instrumen yang berupa tes menggunakan Program Iteaman. Suatu instrumen dikatakan reliabel jika nilai alpha $\geq 0,70$.

2. Penilaian hasil belajar kelompok mata pelajaran ilmu pengetahuan dan teknologi diukur melalui ulangan, penugasan, dan/atau bentuk lain yang sesuai dengan karakteristik materi yang dinilai,
3. Penilaian hasil belajar kelompok mata pelajaran estetika dilakukan melalui pengamatan terhadap perubahan perilaku dan sikap untuk menilai perkembangan afeksi dan ekspresi psikomotorik peserta didik.
4. Penilaian hasil belajar kelompok mata pelajaran jasmani, olahraga, dan kesehatan dilakukan melalui:
 - a. Pengamatan terhadap perubahan perilaku dan sikap untuk menilai perkembangan psikomotorik dan afeksi peserta didik, dan
 - b. Ulangan dan/atau penugasan untuk mengukur aspek kognitif peserta didik.