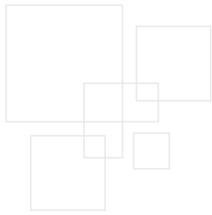


ASESMEN PEMBELAJARAN KEJURUAN

BIDANG PARIWISATA & TATA BOGA



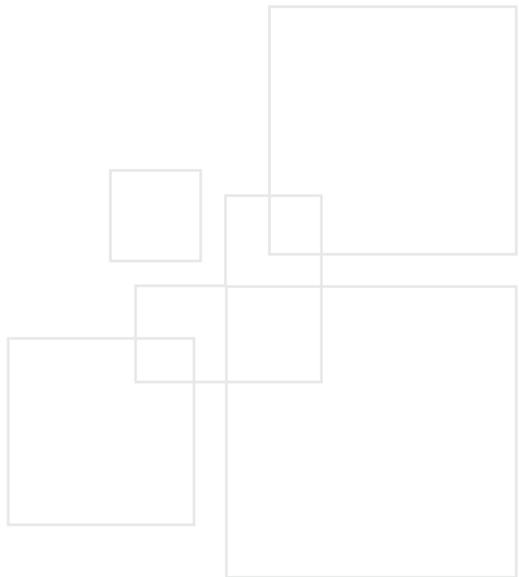
Dr. Endang Mulyatiningsih



ASESMEN PEMBELAJARAN KEJURUAN

BIDANG PARIWISATA & TATA BOGA

Dr. Endang Mulyatiningsih



**UNDANG-UNDANG REPUBLIK INDONESIA
NOMOR 28 TAHUN 2014
TENTANG HAK CIPTA**

Pasal 2

Undang-Undang ini berlaku terhadap:

- a. semua Ciptaan dan produk Hak Terkait warga negara, penduduk, dan badan hukum Indonesia;
- b. semua Ciptaan dan produk Hak Terkait bukan warga negara Indonesia, bukan penduduk Indonesia, dan bukan badan hukum Indonesia yang untuk pertama kali dilakukan Pengumuman di Indonesia;
- c. semua Ciptaan dan/atau produk Hak Terkait dan pengguna Ciptaan dan/atau produk Hak Terkait bukan warga negara Indonesia, bukan penduduk Indonesia, dan bukan badan hukum Indonesia dengan ketentuan:
 1. negaranya mempunyai perjanjian bilateral dengan negara Republik Indonesia mengenai perlindungan Hak Cipta dan Hak Terkait; atau
 2. negaranya dan negara Republik Indonesia merupakan pihak atau peserta dalam perjanjian multilateral yang sama mengenai perlindungan Hak Cipta dan Hak Terkait.

BAB XVII KETENTUAN PIDANA

Pasal 112

Setiap Orang yang dengan tanpa hak melakukan perbuatan sebagaimana dimaksud dalam Pasal 7 ayat (3) dan/atau Pasal 52 untuk Penggunaan Secara Komersial, dipidana dengan pidana penjara paling lama 2 (dua) tahun dan/atau pidana denda paling banyak Rp300.000.000,00 (tiga ratus juta rupiah).

- (1) Setiap Orang yang dengan tanpa hak melakukan pelanggaran hak ekonomi sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf i untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 1 (satu) tahun dan/atau pidana denda paling banyak Rp100.000.000 (seratus juta rupiah).
- (2) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf c, huruf d, huruf f, dan/atau huruf h untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 3 (tiga) tahun dan/atau pidana denda paling banyak Rp500.000.000,00 (lima ratus juta rupiah).
- (3) Setiap Orang yang dengan tanpa hak dan/atau tanpa izin Pencipta atau pemegang Hak Cipta melakukan pelanggaran hak ekonomi Pencipta sebagaimana dimaksud dalam Pasal 9 ayat (1) huruf a, huruf b, huruf e, dan/atau huruf g untuk Penggunaan Secara Komersial dipidana dengan pidana penjara paling lama 4 (empat) tahun dan/atau pidana denda paling banyak Rp1.000.000.000,00 (satu miliar rupiah).
- (4) Setiap Orang yang memenuhi unsur sebagaimana dimaksud pada ayat (3) yang dilakukan dalam bentuk pembajakan, dipidana dengan pidana penjara paling lama 10 (sepuluh) tahun dan/atau pidana denda paling banyak Rp4.000.000.000,00 (empat miliar rupiah).

ASESMEN PEMBELAJARAN KEJURUAN

BIDANG PARIWISATA & TATA BOGA

Dr. Endang Mulyatiningsih



**ASESMEN PEMBELAJARAN KEJURUAN
BIDANG PARIWISATA & TATA BOGA**

Oleh:

Dr. Endang Mulyatiningsih

ISBN: 978-602-5566-69-1

©2018 Endang Mulyatiningsih

Edisi Pertama

Diterbitkan dan dicetak oleh:

UNY Press

Jl. Gejayan, Gg. Alamanda, Komplek Fakultas Teknik UNY

Kampus UNY Karangmalang Yogyakarta 55281

Telp: 0274-589346

Mail: unypress.yogyakarta@gmail.com

Anggota Ikatan Penerbit Indonesia (IKAPI)

Anggota Asosiasi Penerbit Perguruan Tinggi Indonesia (APPTI)

Editor: Shendy Amalia

Desain Sampul: Ngadimin

Tata Letak: Arief Mizuary

ASESMEN PEMBELAJARAN KEJURUAN
BIDANG PARIWISATA & TATA BOGA
-Ed.1, Cet.1.- Yogyakarta: UNY Press 2018
xii + 178 hlm; 16 x 23 cm
ISBN: 978-602-5566-69-1

1. ASESMEN PEMBELAJARAN KEJURUAN
BIDANG PARIWISATA & TATA BOGA

Isi di luar tanggung jawab percetakan

PRAKATA

Buku ini ditulis untuk membantu pendidik (dosen dan guru) dalam melaksanakan tanggung jawab profesionalnya (*professional responsibility*) kepada peserta didik. Salah satu tanggung jawab profesi pendidik adalah memberi penilaian yang objektif kepada peserta didik. Jika penilaian dilakukan secara objektif diharapkan dapat memberi motivasi kepada peserta didik untuk memperoleh prestasi/nilai yang sebaik-baiknya dan selalu meningkatkan hasil belajarnya. Dalam buku ini dijelaskan tentang berbagai macam alat penilaian, cara penulisan, telaah soal secara teoritis maupun empiris.

Penulisan buku ini mendapat bantuan dana dari Universitas Negeri Yogyakarta, melalui program pemberdayaan dosen melalui penulisan buku teks. Oleh sebab itu penulis mengucapkan terima kasih atas kesempatan yang telah diberikan. Semoga buku ini bermanfaat untuk pembaca maupun lembaga.

Yogyakarta, 30 Agustus 2018

Penulis

DAFTAR ISI

PRAKATA	v
DAFTAR ISI.....	vii
DAFTAR TABEL.....	ix
DAFTAR GAMBAR.....	xi
BAB I	
KONSEP PENGUKURAN, PENILAIAN, DAN EVALUASI.....	1
A. Pengukuran (<i>measurement</i>).....	2
B. Penilaian (<i>Assessment</i>).....	5
C. Evaluasi (<i>Evaluation</i>).....	14
D. Implikasi Pengukuran, Penilaian, dan Evaluasi.....	17
BAB II	
PERAN PENILAIAN DALAM PEMBELAJARAN.....	21
A. Belajar dan Proses Pembelajaran.....	21
B. Faktor Penentu Hasil belajar.....	22
C. Peran Evaluasi dalam Proses Pembelajaran.....	26
D. Prinsip-prinsip Penilaian.....	30
E. Tanggung Jawab Profesional dan Etika Guru dalam Penilaian..	31
BAB III	
ANALISIS KOMPETENSI DAN TUJUAN PEMBELAJARAN	35
A. Analisis Kompetensi.....	35
1. <i>Cognitive Domain</i>	36
2. <i>Affective Domain</i>	43
3. <i>Psychomotor Domain</i>	50
B. Tujuan Pembelajaran.....	55
C. Pengukuran Hasil Belajar.....	63
1. Alat Ukur Tes.....	64

2. Alat Ukur non Tes	91
3. <i>Authentic assessment</i>	104
BAB IV	
PENGEMBANGAN TES BAKU	107
A. Perspektif Penggunaan Tes Baku	107
B. Prosedur Pengembangan Tes Baku	110
1. <i>Defining the test universe, audience, and purpose</i>	111
2. <i>Developing a test plan</i>	112
3. <i>Composing the test items</i>	116
4. <i>Writing the administration instructions</i>	128
5. <i>Conduct Piloting Test</i>	128
6. <i>Conduct item analysis</i>	129
7. <i>Revising the test</i>	138
8. <i>Validation the test</i>	139
9. <i>Developing norms</i>	144
10. <i>Complete test manual</i>	145
BAB V	
PENGENDALIAN KUALITAS ALAT PENILAIAN	147
A. Validitas	147
1) Face validity	149
2) <i>Content validity</i>	150
3) Validitas kriteria	153
4) <i>Construct validity</i>	155
5) Validitas Penilaian Kinerja	160
B. Reliabilitas	162
1) Stability Reliability	163
2) Representative reliability	164
3) <i>Equivalence Reliability</i>	165
4) Reliabilitas Penilaian Kinerja	167
5) Upaya Meningkatkan Reliabilitas	167
6) Hubungan antara Reliabilitas dan Validitas	169
7) Bias dalam Penilaian Pendidikan	170
DAFTAR PUSTAKA	173

DAFTAR TABEL

Tabel 1.1	Penetapan Nilai Akhir	7
Tabel 1.2	Penetapan KKM mata pelajaran.....	9
Tabel 1.3	Contoh Penetapan KKM pada KD Ilmu Gizi.....	9
Tabel 1.4	Implementasi Penilaian dan Evaluasi pada Pembelajaran.....	20
Tabel 2.1	Contoh Lembar Observasi Pelaksanaan Pembelajaran.....	29
Tabel 3.1	Tingkat Kompetensi Psikomotorik (Keterampilan)	50
Tabel 3.2	Contoh KK Operasional dan Kurang Operasional.....	55
Tabel 3.3	Contoh Ranah Kompetensi Tujuan Pembelajaran.....	56
Tabel 3.3	Contoh Rubrik Penilaian Kinerja Praktik Harian.....	60
Tabel 3.4	Contoh Soal Benar Salah dengan Koreksi.....	65
Tabel 3.5	Contoh Soal Menjodohkan	65
Tabel 3.6	Contoh Soal Pilihan Ganda.....	66
Tabel 3.7	Telaah Soal Uraian.....	72
Tabel 3.8	Lembar Penilaian Ujian Praktik Kejuruan Mata pelajaran : Melayani Makan dan Minum di Restoran	79
Tabel 3.9	Contoh Lembar Penilaian UKK Patiseri	83
Tabel 3.10	Lembar Penilaian Lomba Kompetensi Siswa.....	85
Tabel 3.11	Indikator Uji Kompetensi Keahlian Basic <i>Culinary Art</i>	87
Tabel 3.12	Contoh Penilaian Sikap dalam Proses Pembelajaran	92
Tabel 3.13	Contoh Rekap Hasil Penilaian Sikap	93
Tabel 3.14	Contoh Rubrik Penilaian Tugas Proyek.....	94
Tabel 3.15	Contoh Rubrik Penilaian Proyek Penelitian	96
Tabel 3.16	Contoh Rubrik Penilaian Portofolio	98

Tabel 3.17	Contoh Penilaian Dokumen Portofolio Project ke 1	99
Tabel 3.18	Contoh Komponen Penilaian Portofolio Guru	100
Tabel 3.19	Contoh Jurnal dari Observasi Selama Satu Semester	101
Tabel 3.0	Contoh Instrumen <i>Self Evaluation</i>	103
Tabel 4.1	Kartu Telaah Soal Pilihan Ganda	118
Tabel 4.1	Contoh Input Data Respon Butir untuk Menghitung Tingkat Kesulitan dan Korelasi Point Biserial	130
Tabel 4.2	Kriteria Indeks Tingkat Kesulitan Butir (p).....	131
Tabel 4.3	Contoh Input Data Respon Butir Untuk Menghitung Daya Pembeda	132
Tabel 4.4	Kriteria Indeks Daya Pembeda Butir (D).....	134
Tabel 4.5	Contoh Tabel Persiapan Analisis Korelasi Point Biserial ...	135
Tabel 4.6	Interpretasi Koefisien Korelasi.....	137
Tabel 4.7	Contoh Rangkuman Hasil Analisis Butir.....	138
Tabel 4.8	Persiapan Analisis Reliabilitas Belah Dua Ganjil Genap....	139
Tabel 4.9	Persiapan Analisis Reliabilitas Menggunakan Rumus Alpha Cronbach's dan KR 20	142
Tabel 5.1	Kisi-kisi Butir Soal Ilmu Gizi.....	150
Tabel 5.2	Input Data Validitas Konstruk dari Program Excel	157
Tabel 5.3	Pedoman untuk Memberikan Interpretasi Terhadap Koefisien Korelasi.....	166
Tabel 5.4	Rangkuman Cara-cara Pengukuran Reliabilitas	167
Tabel 5.4	Rangkuman Tipe-Tipe Pengukuran Reliabilitas dan Validitas	169

DAFTAR GAMBAR

Gambar 1.1	Contoh Pengukuran.....	3
Gambar 1.2	Hasil Pengukuran Kuantitatif.....	6
Gambar 1.3	Konversi Penilaian Kualitatif.....	7
Gambar 1.4	Alur Kegiatan Penilaian Pembelajaran.....	11
Gambar 1.5	Ilustrasi Rangkaian Kegiatan Pengukuran, Penilaian, dan Evaluasi.....	18
Gambar 1.6	Contoh Pengambilan Keputusan Hasil Evaluasi.....	18
Gambar 1.7	Perbedaan Asesmen dan Evaluasi.....	19
Gambar 2.1	Proses Pembelajaran.....	22
Gambar 2.2	Pengalaman Gagal.....	22
Gambar 2.3	Pengalaman Sukses.....	23
Gambar 2.4	Pengaruh Karakteristik Masuk dan Kualitas Pembelajaran pada <i>Learning Outcome</i>	24
Gambar 2.5	Dampak Penilaian terhadap Motivasi Belajar.....	27
Gambar 2.6	Peran Evaluasi bagi Guru.....	28
Gambar 3.1	Taksonomi Kompetensi Kognitif.....	36
Gambar 3.2	Pembentukan Kebiasaan Merokok.....	46
Gambar 3.3	Tahap Perkembangan Belajar Keterampilan.....	54
Gambar 3.4	Konsistensi Tujuan dan Evaluasi dalam Pembelajaran.....	55
Gambar 3.5	Indikator Perumusan Tujuan.....	58
Gambar 3.6	Skema Alat Pengukuran Kemampuan Kognitif.....	63
Gambar 3.7	Klasifikasi Potongan Daging SNI: 3932: 2008.....	75
Gambar 3.8	Skema Penilaian Sikap (Direktorat PSMK, 2017).....	91

Gambar 5.1	Ilustrasi <i>Face Validity</i>	149
Gambar 5.2	Ilustrasi validitas Isi.....	152
Gambar 5.3	Konsep Validitas Kriteria.....	154
Gambar 5.4	Model Pengukuran Validitas Konstruks.....	156
Gambar 5.5	Ilustrasi Validitas Konvergen dan Divergen	160
Gambar 5.6	Contoh Hasil Penilaian Interrater	161
Gambar 5.7	Tipe Kestabilan dalam Mengukur Reliabilitas.....	163
Gambar 5.8	Tipe Representatif dalam Mengukur Reliabilitas.....	164
Gambar 5.9	Tipe Equivalensi dalam Mengukur Reliabilitas.....	166
Gambar 5.10	Hubungan antara Validitas dan Reliabilitas.....	169

BAB I

KONSEP PENGUKURAN, PENILAIAN, DAN EVALUASI

Dalam proses evaluasi dikenal beberapa istilah yang sering digunakan yaitu pengukuran, penilaian, dan evaluasi. Pengukuran, penilaian, dan evaluasi merupakan serangkaian kegiatan yang bersifat hierarkis. Penilaian dilakukan setelah ada data hasil pengukuran. Evaluasi dilakukan setelah ada data hasil pengukuran dan penilaian. **Pengukuran** merupakan suatu kegiatan untuk mendapatkan informasi atau data kuantitatif tentang karakteristik belajar yang dimiliki oleh peserta didik menurut aturan dan formulasi yang jelas. **Penilaian** merupakan suatu kegiatan menginterpretasikan hasil pengukuran, menetapkan kategori, kualitas hasil belajar yang akan digunakan sebagai bahan pertimbangan dalam pengambilan keputusan. Evaluasi adalah suatu proses untuk mengambil keputusan dengan menggunakan informasi yang diperoleh melalui pengukuran hasil belajar baik yang menggunakan instrumen test maupun non-test.

Pengukuran, penilaian dan evaluasi merupakan tiga kegiatan yang berbeda tetapi sering dianggap sama. Pendidik sering menggunakan salah satu istilah yaitu pengukuran atau penilaian atau evaluasi saja untuk mewakili tiga kegiatan yang berbeda tersebut. Istilah pengukuran (*measurement*), penilaian (*assessment*), dan evaluasi (*evaluation*) merupakan satu rangkaian kegiatan yang berurutan. Griffin (1992: 5) memberi contoh penggunaan istilah *measurement*, *assessment*, dan *evaluation* dalam kegiatan: (1) penetapan angka terhadap objek yang diobservasi/pengambilan data termasuk kategori kegiatan pengukuran; (2) interpretasi hasil observasi dan pendeskripsian hasil pengukuran secara keseluruhan termasuk dalam kegiatan asesmen; (3) penggunaan hasil asesmen sebagai bahan pertimbangan dalam pembuatan keputusan

dan implikasinya termasuk dalam kegiatan evaluasi (Griffin & Peter, N., 1991). Keputusan diambil dengan cara membandingkan hasil pengukuran dan asesmen dengan standar yang telah ditetapkan (Johnson & Johnson, 2002). Untuk lebih memahami makna istilah pengukuran, penilaian dan evaluasi maka berikut ini dibahas lebih mendalam tentang penggunaan masing-masing istilah tersebut di dunia pendidikan.

A. Pengukuran (*measurement*)

Pengukuran adalah pemberian angka kepada suatu atribut atau karakteristik yang dimiliki oleh orang (siswa) atau objek tertentu dengan menggunakan aturan atau formulasi yang jelas. Pengukuran didefinisikan oleh beberapa tokoh sebagai berikut: (1) Kizlik mendefinisikan *measurement refers to the process by which the attributes or dimensions of some physical object are determined* (Kizlik, 2012) pengukuran mengacu pada proses dimana atribut atau dimensi dari beberapa objek fisik ditentukan); (2) Bachman (Bachman, 1990) mendefinisikan *measurement* sangat dekat dengan penguantifikasian; (3) Okpalla (Okpalla P. M., 1999) menjelaskan *measurement is therefore a process of assigning numerals to objects, quantities or events in order to give quantitative meaning to such qualities* (pengukuran adalah proses menetapkan angka atau jumlah ke objek atau memberikan arti kuantitatif pada kualitas objek tertentu; (4) GAO (2005) menjelaskan arti pengukuran dalam konteks evaluasi kinerja program sebagai kegiatan pemantauan dan pelaporan prestasi program terus menerus, khususnya untuk mengukur kemajuan program dalam mencapai tujuan yang telah ditetapkan (*Performance measurement is the on going monitoring and reporting of program accomplishments, particularly progress towards preestablished goals*) (Anonim, 2005).

Berdasarkan beberapa definisi di atas dapat diambil pengertian pengukuran (*measurement*) adalah suatu proses pengambilan data karakteristik yang melekat pada orang atau benda yang menjadi sasaran pengukuran. Objek pengukuran dapat berupa karakteristik fisik yang nampak seperti tinggi, berat, panjang, dan karakteristik non fisik yang tidak tampak seperti kecerdasan, sikap, keterampilan sosial, dsb.

Pada saat melakukan pengukuran digunakan alat pengukur. Jika objek yang diukur berupa sifat-sifat fisik yang melekat pada benda seperti panjang, berat, suhu dan volume maka telah tersedia alat ukur yang baku seperti meteran, timbangan, termometer, gelas ukur, dsb. Sebagai

contoh: jika seseorang ingin mengetahui tinggi badan, maka orang tersebut akan mengukur menggunakan alat pengukur panjang (meteran, mitlin, penggaris, dll) yang sudah baku. Hasil pengukuran berupa angka yang menunjukkan tinggi badan. Seseorang yang ingin mengetahui berat badan, maka dia bisa menimbang berat badannya dengan berbagai macam alat penimbang yang tersedia. Setelah mengukur atau menimbang berat badan, maka akan diperoleh data hasil pengukuran dalam bentuk angka/ kuantitatif. Data hasil pengukuran belum diinterpretasikan dalam bentuk apapun. Hasil pengukuran dapat/tidak dapat digunakan tergantung pada keakuratan alat pengukur dan keterampilan orang yang mengukur. Jika alat ukur telah dikalibrasi dan orang yang mengukur terampil maka hasil pengukuran tersebut akurat, atau dapat dipercaya. Hasil pengukurannya yang akurat dapat digunakan untuk berbagai kepentingan asesmen dan evaluasi pendidikan. Proses pengukuran diilustrasikan pada Gambar 1.1



Gambar 1.1 Contoh Pengukuran

Dalam dunia pendidikan, objek yang diukur berupa perilaku (pengetahuan, sikap dan keterampilan) siswa. Mengukur perilaku siswa lebih sulit daripada mengukur benda yang tidak bergerak atau objek fisik

yang tampak karena banyak perilaku manusia yang tersembunyi sehingga hanya dapat diungkap melalui pertanyaan, wawancara atau pengamatan. Untuk mendapatkan informasi yang akurat tentang perilaku siswa diperlukan alat ukur yang berkualitas. Alat ukur yang berkualitaspun tidak berarti apa-apa jika siswa yang diukur tidak memberikan respon positif.

Hasil belajar secara keseluruhan mencakup tiga ranah yaitu pengetahuan (*cognitive*), sikap (*affective*) dan keterampilan (*psichomotor*). Hasil belajar diukur dengan alat tes dan non tes. Tes berupa seperangkat pertanyaan yang harus dijawab/direspon, atau tugas yang harus dikerjakan oleh orang yang dites (diuji) untuk mengungkap pengetahuan atau keterampilan siswa dalam menyelesaikan tugas-tugas tertentu. Alat pengukuran non tes sering digunakan untuk mengukur ranah afektif.

Penggunaan alat tes dan non tes dalam pembelajaran dapat diilustrasikan dari beberapa contoh berikut ini: (1) untuk mengukur pengetahuan, kecerdasan dan olah pikir digunakan tes yang berisi seperangkat pertanyaan. Dari hasil tes tersebut akan diperoleh skor kuantitatif dari perhitungan jumlah jawaban benar/salah; (2) Untuk mengukur kompetensi atau keterampilan kerja diberikan tes kinerja yang berisi tugas-tugas yang harus diselesaikan siswa. Tugas yang diberikan merupakan bagian dari pekerjaan tertentu. Guru mengamati cara siswa menyelesaikan pekerjaan tersebut. Skor ditetapkan dengan cara membandingkan proses dan hasil penyelesaian tugas/pekerjaan terhadap standar kerja yang ada; (3) Jika objek yang diukur berupa sikap maka alat ukur yang tepat digunakan adalah lembar pengamatan perilaku dan wawancara (non tes).

Penetapan skor pada alat pengukur tes dan non tes berbeda. Pada alat ukur tes, setiap butir pertanyaan atau tugas pada tes mempunyai kriteria jawaban atau ketentuan yang dianggap benar atau salah. Skor ditentukan dari kumpulan respon/jawaban yang benar. Alat pengukur non tes tidak menuntut seseorang yang diukur memberi respon pada pertanyaan/tugas secara langsung dan tidak selalu menggunakan kriteria jawaban/respon benar atau salah tetapi lebih banyak menggunakan pertimbangan baik dan kurang baik. Bentuk alat ukur non tes lebih bervariasi yaitu dapat berupa lembar observasi, dokumen portofolio, penugasan, dll.

Kinerja seorang siswa di kelas dapat diketahui dengan cara membandingkan kinerjanya dengan kinerja siswa lainnya. Siswa mendapat skor Matematika 50 memiliki posisi lebih rendah dari siswa lain yang

mendapat skor 70. Pengukuran berhenti pada skor, nilai/kuantitas dan tidak membuat pertimbangan apapun pada kinerja siswa atau tidak ada interpretasi bahwa ia telah lulus atau gagal. Jika skor Matematika 50 diinterpretasikan untuk memberi pertimbangan lulus/tidak lulus maka hal ini sudah masuk ke tahap asesmen. Jika hasil asesmen ditindaklanjuti dengan tindakan lain (pemberian penghargaan atau perbaikan) maka kegiatan sudah menginjak pada tahap evaluasi.

Pengukuran hasil belajar disarankan tidak hanya menggunakan satu alat ukur dan satu kali pengukuran saja. Guru dapat menggunakan berbagai alat pengukuran secara komplementer (saling melengkapi) sesuai dengan kompetensi yang dinilai (kognitif, afektif, dan psikomotorik). Pengukuran hasil belajar siswa yang menggunakan alat ukur bervariasi dan dilakukan berulang-ulang selama dan sesudah proses pembelajaran akan memperoleh hasil pengukuran yang lebih komprehensif dan objektif. Pengukuran hasil belajar yang hanya dilakukan satu kali rawan mendapatkan informasi yang salah dan memiliki banyak kelemahan yang disebabkan oleh: (1) kondisi fisik dan psikis siswa tidak siap saat dilakukan pengukuran; (2) wilayah materi yang dipelajari siswa kemungkinan berbeda dengan materi yang diujikan; (3) siswa kemungkinan lemah pada ranah kognitif tetapi unggul pada ranah lainnya. Berdasarkan kelemahan ini maka guru diharapkan mampu mengembangkan dan menggunakan berbagai alat ukur.

B. Penilaian (*Assessment*)

Pembelajaran, belajar, asesmen, dan evaluasi saling berhubungan dan tidak dapat dipisahkan. Guru bertanggung jawab memberi pembelajaran agar siswa dapat belajar. Secara periodik, asesmen dilakukan untuk memberi pertimbangan secara kuantitatif maupun kualitatif tentang hasil belajar atau kompetensi untuk menentukan kelulusan siswa. Asesmen berfungsi untuk memperbaiki kegiatan pembelajaran supaya lebih efektif. Asesmen dapat dilakukan tanpa evaluasi tetapi evaluasi tidak dapat dilakukan tanpa asesmen. Asesmen selalu diawali dengan pengukuran melalui tes/ujian tetapi tidak semua tes/ujian digunakan dalam asesmen. Tes bukan satu-satunya sumber data asesmen tetapi hanya digunakan untuk mengambil data kuantitatif. Asesmen yang baik seyogianya menggunakan berbagai sumber data hasil pengukuran misalnya inventories, kuesionair, observasi, dsb. Semua sumber data kemudian diorganisasikan untuk menilai kemampuan siswa.

Ada beberapa definisi yang memperjelas pengertian asesmen. Kizlik (2012) mengemukakan bahwa *assessment is a process by which information is obtained relative to some known objective or goal* (Kizlik, 2012). Definisi tersebut mengandung maksud penilaian adalah suatu proses dimana informasi diperoleh untuk mengetahui ketercapaian tujuan. Terry Overton (2011) menyatakan *assessment: the process of gathering information to monitor progress and make educational decisions if necessary. an assessment may include a test, but also includes methods such as observations, interviews, behavior monitoring, etc.* (Overton, 2011). Definisi tersebut dalam terjemahan bebas dapat diartikan sebagai berikut: penilaian adalah proses pengumpulan informasi untuk memantau kemajuan dan membuat keputusan pendidikan jika diperlukan. penilaian tidak hanya menggunakan hasil ujian/tes, tetapi juga mencakup hasil observasi, wawancara, pemantauan perilaku, dll. Johnson (2002) mendefinisikan asesmen adalah proses pengumpulan informasi kualitatif maupun kuantitatif atas perubahan yang terjadi pada diri siswa, kelompok, guru atau penyelenggara pendidikan (sekolah) selama proses belajar (Johnson & Johnson, 2002).

Asesmen dilakukan dengan cara mengorganisasikan berbagai macam data hasil pengukuran agar dapat diinterpretasikan. Pada dunia pendidikan asesmen digunakan untuk menjelaskan kemajuan siswa dalam mencapai tujuan pendidikan pada titik waktu tertentu. Di dalam kelas, asesmen digunakan untuk menjelaskan proses dan produk yang dicapai siswa secara alami. Ahli pendidikan sering membandingkan asesmen dengan evaluasi. Secara umum, evaluasi lebih ditekankan pada efektivitas program pembelajaran atau kelembagaan sedangkan asesmen digunakan untuk mendiskripsikan kemampuan atau prestasi siswa secara perseorangan atau kelompok. Proses asesmen yang dapat diilustrasikan pada Gambar 1.2 dan 1.3.



Gambar 1.2 Hasil Pengukuran Kuantitatif



Gambar 1.3 Konversi Penilaian Kualitatif

Asesmen dalam arti yang sempit sering disamakan dengan penilaian yaitu pemberian nilai terhadap data hasil pengukuran. Menilai adalah mengambil suatu keputusan secara kualitatif yaitu baik atau buruk terhadap karakteristik objek yang telah diukur. Data kuantitatif yang diperoleh pada saat melakukan pengukuran kemudian diinterpretasikan menurut kriteria yang telah ditetapkan. Penilaian hasil belajar sering menggunakan beberapa tingkatan kriteria yaitu mulai dari sangat baik, baik, cukup baik, kurang baik, dan sangat kurang baik; kompeten atau tidak kompeten dan tuntas atau tidak tuntas. Predikat kelulusan menggunakan 3 kriteria yaitu memuaskan, sangat memuaskan dan dengan pujian. Penilaian pada jenjang pendidikan tinggi menggunakan nilai huruf (A, B, C, D, E) yang merupakan konversi dari nilai angka seperti pada Tabel 1.1.

Tabel 1.1 Penetapan Nilai Akhir

NILAI ANGKA	NILAI HURUF	INDEX NILAI	NILAI AKHIR	KONVERSI	
				Huruf	Bobot
80 – 100	A	4	86 – 100	A	4.00
75 – 79	AB	3,5	81 – 85	A-	3.67
70 – 74	B	3	76 – 80	B+	3.33
65 – 69	BC	2,5	71 – 75	B	3.00
60 – 64	C	2	66 – 70	B-	2.67
55 – 59	CD	1,5	61 – 65	C+	2.33
40 – 54	D	1	56 – 60	C	2.00
< 40	E	0	41 – 55	D	1.00
			0 – 40	E	0.00

Sumber: Peraturan Akademik Politeknik Negeri Bandung

Peraturan Akademik UNY

Setiap lembaga dapat menentukan kriteria sendiri sehingga nilai yang dimunculkan berbeda-beda. Panjang interval kelas pada nilai paling atas (nilai A) dan panjang interval kelas paling bawah (E) memiliki jarak yang tidak sama dibanding panjang interval kelas pada nilai lainnya. POLBAN menetapkan panjang interval kelas nilai A sebesar 20 sedangkan UNY menetapkan panjang interval kelas sebanyak 15. Batas kelulusan untuk setiap mata kuliah berbeda. Kelulusan mahasiswa pada akhir studi ditentukan dengan IPK (*Index Prestasi Kumulatif*) yang merupakan akumulasi nilai semua mata kuliah.

Hasil pembelajaran kejuruan ditetapkan menggunakan standar kompetensi sehingga keputusan hasil penilaian dinyatakan dengan kata kompeten dan tidak kompeten atau tuntas dan tidak tuntas. Penilaian kompetensi menggunakan pendekatan penilaian acuan kriteria (PAK). PAK merupakan penilaian pencapaian kompetensi yang didasarkan pada kriteria ketuntasan minimal (KKM). KKM merupakan kriteria ketuntasan belajar minimal yang ditentukan oleh satuan pendidikan dengan mempertimbangkan karakteristik kompleksitas Kompetensi Dasar yang akan dicapai, daya dukung, dan karakteristik intake peserta didik (Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2013 Tentang Standar Penilaian Pendidikan).

Tingkat kompleksitas yaitu tingkat (kesulitan dan kerumitan) setiap KD atau indikator yang harus dicapai oleh peserta didik dalam belajar. Tingkat kompleksitas dinyatakan tinggi jika kompetensi yang harus dicapai peserta didik memerlukan tingkat pemahaman yang tinggi, guru harus kreatif dan inovatif dalam melaksanakan pembelajaran, waktu pencapaian kompetensi cukup lama karena perlu pengulangan dan untuk mencapai kompetensi tersebut siswa harus memiliki daya nalar dan kecermatan yang tinggi. KKM ditetapkan **rendah** jika tingkat **kompleksitas tinggi**. Daya dukung meliputi ketersediaan tenaga, sarana dan prasarana dan biaya operasional dari sekolah. KKM ditetapkan tinggi jika daya dukung yang tersedia di sekolah memadai. *Intake* meliputi tingkat kemampuan rata-rata siswa. KKM ditetapkan tinggi jika *intake* kemampuan rata-rata siswa tinggi. Kriteria ideal ketuntasan untuk masing-masing indikator program normatif dan adaptif adalah 75%.

KKM untuk program Normatif dan Adaptif:

Contoh cara penetapan KKM kompetensi dasar diambil dari materi bimtek Evaluasi: Penetapan KKM pada KTSP oleh Departemen Pendidikan Nasional (Endang Mulyani: 2009).

Tabel 1.2 Penetapan KKM mata pelajaran

	Indikator	Kriteria nilai		
		Tinggi	Sedang	Rendah
A	Kompleksitas materi (KM)	1	2	3
B	Daya dukung sekolah (DD)	3	2	1
C	Tingkat kemampuan rata-rata siswa	3	2	1
	Nilai rerata kemampuan siswa	80 - 100	60 - 79	< 60

Jika indikator memiliki kriteria: A-kompleksitas rendah (3), B-daya dukung tinggi (3) dan C-tingkat kemampuan rata-rata peserta didik sedang (2) maka nilai KKM-nya adalah

$$\text{KKM} = \frac{A+B+C}{9} \times 100 \text{ atau } \frac{3+3+2}{9} \times 100 = 88,9$$

Contoh penerapan penetapan KKM pada beberapa kompetensi dasar (KD) mata pelajaran ilmu gizi dapat disimak pada tabel 3. Kompleksitas materi (A), daya dukung (B) dan tingkat kemampuan rata-rata peserta didik (C) ditafsirkan sendiri oleh guru karena gurulah yang paling tahu tentang kondisi lingkungan sekolah, kemampuan siswa dan tingkat kesulitan materi.

Tabel 1.3 Contoh Penetapan KKM pada KD Ilmu Gizi

Kompetensi Dasar	Faktor			KKM	
	A	B	C		
4.1. Menganalisis angka kecukupan gizi	2	3	2	7/9	77,8
4.2. Menjelaskan menu seimbang	3	3	2	8/9	88,9
4.3. Menyusun menu untuk bayi dan balita	1	3	2	6/9	66,7
Rerata seluruh KD	2	3	2	7;9	77,8

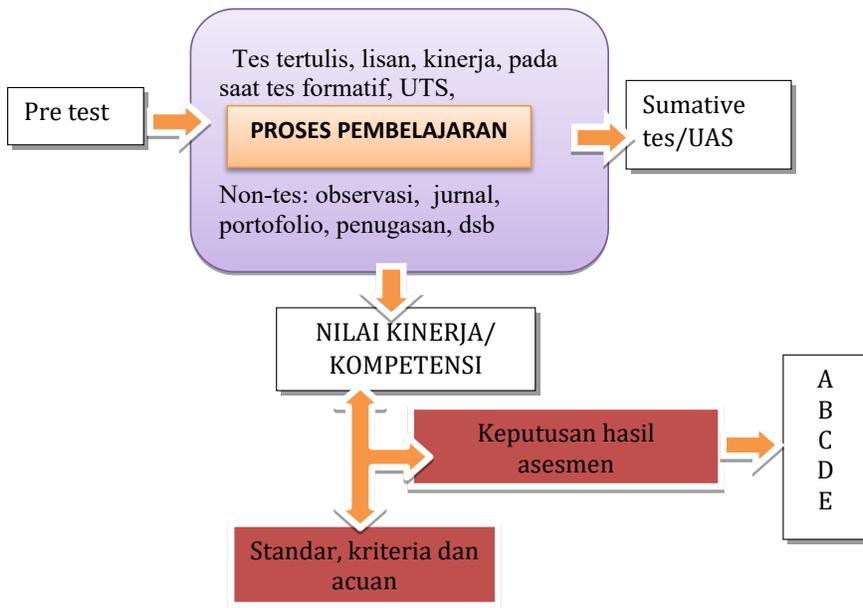
KKM pada ilmu sosial sangat normatif karena dipengaruhi kemampuan pribadi dan tingkat kesulitan materi pelajaran. Jika materi pelajaran sulit dan kemampuan rata-rata siswa rendah maka KKM juga menjadi rendah. Hal ini tentu saja dibuat dengan pertimbangan kemanusiaan supaya banyak siswa yang lulus ujian apalagi untuk siswa dari daerah-daerah terpencil yang daya dukung sekolahnya rendah. Jika KKM ditetapkan tinggi di daerah terpencil maka akan berakibat banyak siswa yang tidak lulus. Batas kelulusan ini bertolak belakang dengan pendidikan profesi dan kejuruan yang menggunakan standar kompetensi.

KKM Program Produktif

KKM program produktif mengacu kepada standar minimal penguasaan kompetensi yang berlaku di dunia kerja yang bersangkutan. Kriteria ketuntasan untuk masing-masing kompetensi dasar (KD) adalah terpenuhinya indikator yang dipersyaratkan dunia kerja yaitu kompeten atau belum kompeten dan diberi lambang/skor 7,00 bila memenuhi persyaratan minimal. Nilai <7 dinyatakan belum kompeten. Pencapaian kompetensi setiap indikator kemudian diberi tingkatan dari cukup kompeten (7,0-7,9); kompeten (8,0-8,9); dan sangat kompeten (9,0-10). Hasil akhir nilai kompetensi merupakan nilai rerata skor setiap indikator pada komponen: persiapan, proses, hasil kerja, sikap kerja, dan penggunaan waktu.

Urgensi atau tingkat kepentingan mata pelajaran bagi lulusan dan dunia kerja berpengaruh terhadap batas nilai (*cut score*) kelulusan yang harus dicapai siswa. Beberapa pekerjaan di bidang sains, teknologi dan kesehatan pada umumnya menetapkan standar ketuntasan/kompetensi yang tinggi karena jika masih terdapat kesalahan pekerjaan akan berbahaya bagi keselamatan manusia. Jika siswa tidak mampu mencapai standar kerja yang ditentukan maka dia tidak berhak lulus. Contoh: KD meracik obat pada SMK kesehatan termasuk materi sulit dan kompleks. Meskipun sulit dan kompleks, KKM tetap menggunakan standar tinggi karena jika terjadi kesalahan dalam meracik obat maka berakibat penyakit tidak sembuh atau bahkan meninggal. Pelajaran mengoperasikan CNC (*Computer Numerical Control*) juga sulit dan kompleks tetapi jika siswa tidak bisa mencapai standar kompetensi yang tinggi maka dapat menyebabkan kerusakan pada alat maupun produk yang dihasilkan, padahal alat dan bahan pembuat produk sangat mahal.

Penilaian hasil belajar seharusnya dilakukan secara komprehensif yaitu dengan menggunakan berbagai macam alat pengukur dan beberapa kali pengukuran. Guru dapat memanfaatkan momen-momen penting untuk melaksanakan pengukuran sebelum, selama dan sesudah proses pembelajaran dengan menggunakan berbagai jenis alat pengukur berupa tes dan non tes. Pengukuran dan penilaian sebelum, selama, dan setelah pembelajaran diilustrasikan pada Gambar 1.4.



Gambar 1.4 Alur Kegiatan Penilaian Pembelajaran

Penjelasan:

- 1) mengukur hasil belajar dengan menggunakan berbagai macam alat pengukur seperti tes (lisan/tertulis), dan non tes seperti observasi, wawancara, presentasi, pengumpulan dokumen portofolio, dsb.
- 2) Memberi skor/nilai kuantitatif pada semua dokumen pengukuran dari alat tes maupun non tes.
- 3) Membandingkan skor/nilai yang diperoleh dengan standar kompetensi atau kriteria ketuntasan minimal (KKM).
- 4) Menginterpretasikan nilai sesuai dengan standar kompetensi atau kriteria ketuntasan minimal menggunakan pernyataan kualitatif

seperti **kompeten** >< **tidak kompeten**; menggunakan huruf alphabet (A, B, C, D, E) sangat baik, baik, cukup baik, kurang baik, dan tidak layak.

Penilaian merupakan tahap yang penting dalam pembelajaran. Suryabrata (1981) menjelaskan pengukuran dan penilaian dalam pendidikan mempunyai tiga fungsi yaitu:

- 1) melayani kebutuhan psikologis seseorang dalam memperoleh informasi tentang potensi yang dimiliki untuk dijadikan acuan dalam menentukan arah ke masa depan;
- 2) melayani kebutuhan instruksional, untuk mengetahui kesesuaian materi pelajaran, metode pembelajaran dengan kemampuan dan potensi peserta didik dan mengetahui posisi masing-masing siswa di dalam kelas;
- 3) melayani kebutuhan administrasi untuk mengetahui peringkat atau indeks prestasi siswa (Suryabrata, 1981).

Johnson (2002) menjelaskan hal yang serupa yaitu hasil penilaian dapat digunakan oleh siswa, guru dan sekolah yaitu:

- 1) Siswa mendapat informasi kelulusan, penghargaan dan pengakuan terhadap pengetahuan dan keterampilan yang telah dipelajari.
- 2) Guru memanfaatkan hasil asesmen untuk menemukan informasi apakah proses pembelajaran yang dijalankan sudah atau belum efektif.
- 3) Sekolah dapat menentukan keefektifan sekolah dibanding dengan sekolah lainnya (Johnson & Johnson, 2002).

Belajar pada hakikatnya adalah mengubah perilaku siswa dari yang semula tidak bisa menjadi bisa. Untuk mengetahui perubahan perilaku tersebut dilakukan pengukuran dan penilaian. Siswa yang telah sukses dalam belajar perlu mendapat penghargaan dengan memberi nilai yang bagus. Penghargaan yang sepadan dengan usahanya dapat memberikan rasa puas sehingga mendorong siswa untuk belajar lebih giat belajar lagi. Sebaliknya, penilaian hasil belajar yang tidak objektif yaitu tidak mampu membedakan antara siswa yang pandai, tekun, ulet, dan rajin belajar dengan siswa yang kurang pandai dan malas belajar justru akan berdampak buruk pada motivasi belajar siswa. Jika hal ini terjadi, siswa menjadi malas berusaha karena hasil belajar yang akan diperoleh sama meskipun

intensitas belajarnya tidak sama. Dalam Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 66 Tahun 2013 Tentang Standar Penilaian Pendidikan dinyatakan bahwa Penilaian hasil belajar didasarkan pada prinsip-prinsip sebagai berikut.

- 1) *Objektif*, berarti penilaian berbasis pada standardan tidak dipengaruhi faktor subjektivitas penilai.
- 2) *Terpadu*, berarti penilaian oleh pendidik dilakukan secara terencana, menyatu dengan kegiatan pembelajaran, dan berkesinambungan.
- 3) *Ekonomis*, berarti penilaian yang efisien dan efektif dalam perencanaan, pelaksanaan, dan pelaporannya.
- 4) *Transparan*, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diakses oleh semua pihak.
- 5) *Akuntabel*, berarti penilaian dapat dipertanggungjawabkan kepada pihak internal sekolah maupun eksternal untuk aspek teknik, prosedur, dan hasilnya.
- 6) *Edukatif*, berarti mendidik dan memotivasi peserta didik dan guru.

Hasil penilaian digunakan untuk mengambil keputusan penting harus memenuhi persyaratan di atas. Tes yang digunakan untuk mengambil keputusan penting dinamakan *high stake testing* dan keputusannya disebut *high stake decisions*. Menurut Nitko (2011: 87) keputusan hasil tes dianggap penting jika memberi konsekuensi yang serius, memberi manfaat besar dan jika gagal dapat menyebabkan kehilangan beberapa hal yang sangat berharga yang tidak dapat dikembalikan setelah keputusan dibuat (Nitko & Brookhart, 2011). Beberapa contoh *high stake testing/decisions* antara lain tes yang digunakan untuk penentuan kelulusan, seleksi sekolah atau pekerjaan dan penempatan. Hasil penilaian tersebut akan menentukan nasib seseorang pada perjalanan hidupnya dimasa yang akan datang. Hasil penilaian yang tidak objektif akan merugikan lembaga pengguna karena orang-orang yang mendapatkan nilai tinggi ternyata kemampuannya rendah. Hasil penilaian yang tidak objektif juga tidak akan akuntabel atau tidak dapat dipertanggungjawabkan kepada pihak lain.

Keputusan hasil penilaian akan memberi konsekuensi positif atau negatif. Guru memiliki tanggung jawab profesional dalam memberi keputusan. Apabila keputusan yang diberikan tidak berkualitas, informasi yang diperoleh salah dan menyebabkan keputusan yang diambil salah maka dapat menyebabkan guru tidak profesional. tidak etis dan tidak legal. Salah

satu cara untuk memperbaiki keputusan pendidikan adalah menggunakan *high-quality assessment information* yaitu dengan menggunakan informasi yang memiliki validitas dan reliabilitas tinggi.

C. Evaluasi (*Evaluation*)

Evaluasi mempunyai arti yang sangat luas dan tidak hanya digunakan pada kegiatan pembelajaran saja. Evaluasi juga digunakan oleh semua institusi untuk mengetahui apakah program-program yang dirancang telah sukses atau gagal dilaksanakan. Untuk memahami makna evaluasi secara lebih mendalam, berikut ini dikutip beberapa definisi evaluasi yang mengarah pada evaluasi program dan evaluasi hasil belajar. Dalam konteks evaluasi program, (1) Alkin (Alkin & Woolley, 1969) menyebutkan bahwa evaluasi adalah: *The process of ascertaining the decision of concern, selecting appropriate information and collecting and analysing information in order to report summary data useful to decision makers in selecting among alternatives*. Pengertian tersebut mengandung makna bahwa evaluasi adalah proses menentukan keputusan, memilih informasi yang tepat, mengumpulkan dan menganalisis informasi dan melaporkan data secara ringkas untuk para pengambil kebijakan dalam menyeleksi beberapa alternatif keputusan; (2) Madaus, G. F., & Stufflebeam, D. L. menyatakan evaluasi adalah: *“the process of delineating, obtaining and providing useful information for judging decision alternatives”* (Madaus & Stufflebeam, D L, 1985). Definisi evaluasi tersebut dapat diartikan sebagai proses yang menggambarkan, memperoleh dan memberikan informasi yang berguna untuk menetapkan beberapa alternatif keputusan; (3) Bachman (Bachman, 1990), yang mengutip pendapat Weiss (1972) mendefinisikan evaluasi sebagai *“the systematic gathering of information for the purpose of making decisions”* atau “pengumpulan informasi secara sistematis untuk tujuan membuat keputusan;

Dalam konteks evaluasi hasil belajar, Terry Overton menyatakan *evaluation is procedures used to determine whether the subject (i.e. student) meets a preset criteria, such as qualifying for special education services. This uses assessment to make a determination of qualification in accordance with a predetermined criteria.* (Overton, 2011). Dalam terjemahan bebas definisi evaluasi tersebut mengandung pengertian sebagai prosedur yang digunakan untuk menentukan apakah subjek (yaitu siswa) telah memenuhi kriteria yang telah ditetapkan, seperti kualifikasi pendidikan

khusus. Evaluasi menggunakan penilaian untuk membuat penentuan kualifikasi sesuai dengan kriteria yang telah ditentukan. Burden & Byrd menambahkan evaluasi diperlukan untuk membantu guru mengetahui tingkat pencapaian tujuan pembelajaran dan membantu guru mengetahui posisi siswa secara individual (Burden & Byrd, 1998).

Berdasarkan beberapa definisi yang telah dipaparkan di atas, evaluasi mengandung pengertian sebagai suatu proses yang sistematis untuk menentukan atau membuat keputusan tentang ketercapaian tujuan program atau pembelajaran. Sistematis karena evaluasi hanya dapat dilakukan setelah ada data hasil pengukuran dan penilaian. Substansi yang diukur dan dinilai harus konsisten dengan tujuan pembelajaran. Hasil penilaian kemudian dibandingkan dengan kriteria, acuan atau standar yang telah ditetapkan sebelumnya. Dengan evaluasi akan diketahui apakah proses dan hasil belajar sudah atau belum sesuai dengan kriteria, acuan atau standar yang telah ditetapkan untuk mengambil keputusan. Asesmen dapat berdiri sendiri tanpa evaluasi tetapi evaluasi tidak dapat dilakukan tanpa pengukuran dan asesmen.

Setiap kegiatan pembelajaran memerlukan evaluasi proses dan hasil belajar. Evaluasi proses pembelajaran merupakan evaluasi yang dilaksanakan oleh pendidik untuk mengetahui efektivitas pemilihan strategi, metode, media dan teknik pembelajaran terhadap peningkatan motivasi, aktivitas dan kemajuan belajar siswa. Evaluasi proses pembelajaran memberi manfaat penting untuk mendiagnosis kesulitan belajar dan meningkatkan kualitas proses pembelajaran. Evaluasi hasil belajar dilakukan untuk mengetahui pencapaian tujuan pembelajaran oleh setiap peserta didik. Evaluasi hasil belajar bermanfaat untuk mengetahui tingkat keberhasilan peserta didik, tingkat pencapaian kurikulum dan memotivasi belajar siswa.

Tujuan menjadi pengendali semua rangkaian kegiatan proses dan hasil pembelajaran. Tujuan pembelajaran telah dirumuskan pendidik (guru/dosen) pada saat menyusun rencana pelaksanaan pembelajaran (RPP). Tujuan pembelajaran berisi kompetensi-kompetensi yang harus dicapai setelah belajar. Evaluasi hasil belajar dilakukan untuk mengetahui pencapaian kompetensi yang telah dirumuskan pada tujuan. Dalam RPP ditetapkan indikator pencapaian sebagai acuan, kriteria untuk menentukan apakah hasil belajar telah sukses atau gagal.

Sasaran hasil belajar adalah penguasaan kompetensi. Kompetensi adalah seperangkat tindakan cerdas penuh tanggung jawab yang dimiliki

seseorang sebagai syarat untuk dianggap mampu oleh masyarakat dalam melaksanakan tugas-tugas di bidang pekerjaan tertentu (SK. Mendiknas No. 045/U/2002). Peserta didik (mahasiswa) dinyatakan berkompeteren dalam pekerjaan tertentu manakala ia memiliki pengetahuan, keterampilan dan sikap kerja minimum yang digunakan untuk menyelesaikan pekerjaan tersebut dalam bentuk unjuk kerja/kinerja/perilaku. Dalam pembelajaran, unjuk kerja merupakan penampilan peserta didik dalam mengerjakan sesuatu tugas yang terkait dengan pembelajaran yang dilakukan.

Evaluasi tidak hanya diperlukan pada kegiatan pembelajaran. Semua program/kegiatan juga menggunakan evaluasi untuk melihat tingkat keberhasilan program yang telah dicapai, mengetahui efektivitas dan efisiensi program yang sedang berjalan dan memperoleh informasi untuk penetapan kegiatan berikutnya. Dalam proses evaluasi, informasi dari data pengukuran diinterpretasikan menurut standar/kriteria yang ditetapkan sebagai dasar untuk membuat keputusan. Sukses evaluasi tergantung pada keakuratan data hasil pengukuran yang dikumpulkan. Apabila hasil pengukuran tidak konsisten (atau reliabel) dan tidak valid (*truthful*) maka keakuratan hasil evaluasi tidak mungkin dapat dicapai. Data yang kurang akurat dapat menyebabkan keputusan yang diambil tidak adil dan merugikan orang lain sehingga berdampak pada program kurang efisien dan efektif.

Evaluasi pendidikan dilakukan dari waktu ke waktu. Oguniyi merangkum tujuan evaluasi program pembelajaran dan program lain yang relevan. Secara umum, evaluasi bertujuan untuk:

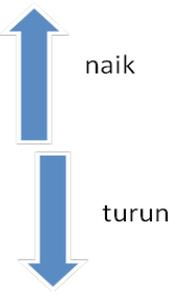
- 1) menentukan efektivitas relatif program terhadap output perilaku siswa;
- 2) membuat keputusan yang dapat diandalkan tentang perencanaan pendidikan;
- 3) memastikan kelayakan waktu, energi dan sumber daya yang diinvestasikan dalam program;
- 4) mengidentifikasi perkembangan siswa dalam memperoleh pengetahuan, keterampilan, sikap dan nilai-nilai sosial yang diinginkan;
- 5) membantu guru menentukan efektivitas teknik mengajar dan materi pembelajaran;
- 6) membantu memotivasi siswa untuk belajar lebih lanjut karena mereka menemukan kemajuan atau kurangnya kemajuan dalam mengerjakan tugas yang diberikan;

- 7) mendorong siswa untuk mengembangkan rasa disiplin dan kebiasaan belajar yang sistematis;
- 8) memberikan informasi yang memadai tentang efektivitas guru dan sekolah kepada penyelenggara pendidikan;
- 9) melaporkan kemajuan anak-anak kepada orang tua atau wali;
- 10) mengidentifikasi masalah yang mungkin menghambat pencapaian tujuan yang ditetapkan;
- 11) memprediksi kecenderungan umum dalam pengembangan proses belajar mengajar;
- 12) memastikan efisiensi manajemen sumber daya yang langka;
- 13) dasar objektif untuk memberi penghargaan, sertifikat, kenaikan kelas;
- 14) menjadi dasar untuk membuat ijazah yang berguna untuk melamar pekerjaan (Ogunniyi, 1984).

Pada evaluasi program, hasil evaluasi pada umumnya digunakan untuk mengambil keputusan yang berdampak pada: (1) melanjutkan program jika hasil yang dicapai memenuhi kriteria sangat baik; (2) memperbaiki program jika hasil yang dicapai kurang memuaskan karena terjadi beberapa kesalahan selama pelaksanaan program; (3) menghentikan program jika program tidak efektif dan hanya menimbulkan kerugian. Pada evaluasi hasil belajar, hasil evaluasi digunakan untuk mengambil keputusan program remedial bagi peserta didik yang belum mencapai KKM (kriteria ketuntasan minimum) dan program pengayaan bagi peserta didik yang sudah mencapai KKM lebih cepat. Program remedial individu dapat dilakukan melalui penugasan, ujian ulang, tutor sebaya, dan bimbingan belajar. Program remedial kelas dilakukan melalui pendalaman materi, bimbingan belajar klasikal. Program pengayaan dapat dilakukan melalui kegiatan penugasan belajar materi baru yang lebih kompleks.

D. Implikasi Pengukuran, Penilaian, dan Evaluasi

Setelah mempelajari konsep pengukuran, penilaian dan evaluasi secara terpisah, berikut ini diilustrasikan contoh penerapan kegiatan tersebut sehari-hari. Berikut ini contoh gambar langkah-langkah pengukuran, penilaian dan evaluasi pada program diet penurunan berat badan.

 <p>Sumber: https://fragilecekal.wordpress.com/2012/02/page/4/</p>		 <p>Sumber: http://www.alodokter.com/pilihan-makanan-untuk-diet-yang-mengenyangkan</p>
<p>Pengukuran</p>	<p>Penilaian</p>	<p>Keputusan evaluasi</p>

Gambar 1.5 Ilustrasi Rangkaian Kegiatan Pengukuran, Penilaian, dan Evaluasi

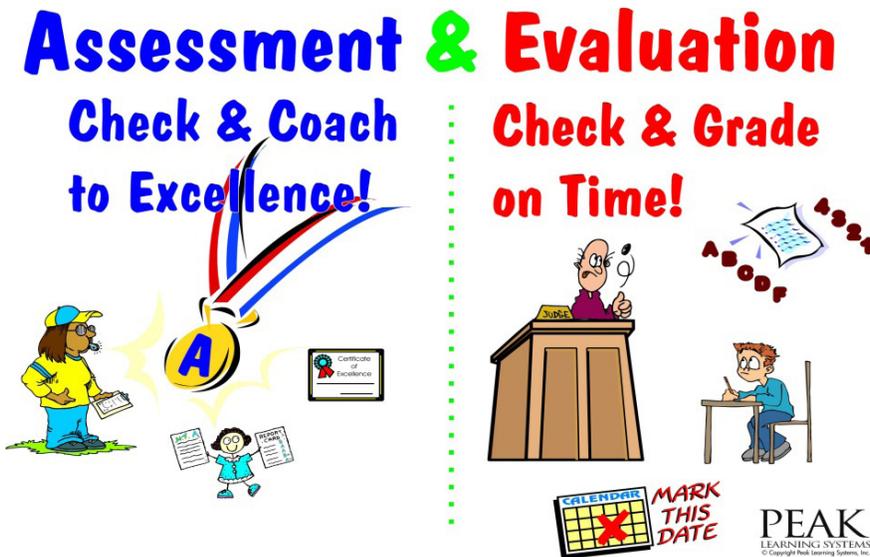
Contoh ketika seorang ingin mengetahui efektivitas program diet yang dilakukan untuk menurunkan berat badan maka dia menimbang (mengukur) berat badannya. Setelah ditimbang, dia mendapatkan informasi tentang berat badannya telah turun atau belum. Jika berat badannya turun maka dapat diputuskan program dietnya telah berhasil dan dapat dilanjutkan. Tetapi jika berat badan tidak bisa turun maka diputuskan untuk menghentikan program diet dan menggantinya dengan program lain misalnya dengan olah raga dan mengatur pola makan seperti Gambar 1.6 berikut ini.

 <p>Sumber: http://siti042915.blogspot.co.id/2011/12/cara-menghitung-berat-badan-ideal.html</p>	<p>Berat badan sama tapi Lingkar perut beda?</p>	 <p>Sumber: http://kawulo-nyantrik.blogspot.co.id/2011/04/melatih-otot-bahu-1.html#.WpWL9XerfU</p>
<p>Pengukuran</p>	<p>Penilaian</p>	<p>Keputusan berolah raga</p>

Gambar 1.6 Contoh Pengambilan Keputusan Hasil Evaluasi

Ilustrasi di atas menunjukkan dua orang yang memiliki pola makan sama tetapi aktivitas berbeda. Setelah diukur, berat badan mereka sama tetapi lingkaran perut berbeda. Orang yang rajin berolahraga memiliki otot yang lebih padat sedangkan orang yang tidak pernah berolahraga memiliki otot yang kendur. Dari sisi penampilan dan kesehatan, orang yang rajin berolahraga memiliki status yang lebih baik. Evaluasi menyarankan agar orang mengikuti olahraga rutin supaya bentuk badannya bagus dan sehat.

Perbedaan antara asesmen dan evaluasi diilustrasikan pada gambar 7. Asesmen berfungsi untuk mengecek kemajuan belajar dan membimbing siswa agar menjadi lebih baik (*excellence*). Evaluasi digunakan untuk mengecek dan mengambil keputusan kelulusan pada saat itu. Dalam proses pembelajaran, evaluasi merupakan tahap terakhir untuk menetapkan siswa dapat melanjutkan belajar tahap berikutnya atau mengulang kembali pelajaran yang belum tuntas.



Gambar 1.7 Perbedaan Asesmen dan Evaluasi

Sumber: <https://ronatoralde.wordpress.com/2016/03/08/module-1-what-is-the-difference-between-assessment-and-evaluation/>

Implikasi penilaian dan evaluasi dalam kegiatan pembelajaran dirangkum pada Tabel 1.4 berikut:

Tabel 1.4 Implementasi Penilaian dan Evaluasi pada Pembelajaran

Dimensi	Penilaian	Evaluasi
Waktu	Formatif	Sumatif
Fokus pengukuran	Berorientasi proses	Berorientasi hasil
Hubungan antara penyelenggara dengan peserta	Reflektif (melihat kembali hal-hal yang masih kurang)	Preskriptif (bersifat menentukan)
Temuan dan penggunaan	Diagnostik	Keputusan
Kemungkinan perubahan kriteria pengukuran	Fleksibel, luwes masih dapat diubah	Fix, keputusan terakhir yang tidak dapat diubah
Standar pengukuran	Mutlak untuk individu siswa	Dibandingkan dengan siswa lain
Hubungan antar siswa	Kerjasama (kooperatif)	Kompetisi (bersaing)

Pengukuran, penilaian, dan evaluasi merupakan serangkaian kegiatan yang dilakukan secara berurutan. Pengukuran dilakukan pertama kali ketika guru ingin mengetahui kemampuan siswa dengan menggunakan alat ukur tes atau non tes. Penilaian dilakukan untuk memberikan informasi tentang kemampuan yang diperoleh siswa setelah diukur. Evaluasi menggunakan data hasil pengukuran dan penilaian untuk membuat keputusan lulus/tidak lulus dan kompeten/tidak kompeten.

BAB II

PERAN PENILAIAN DALAM PEMBELAJARAN

A. Belajar dan Proses Pembelajaran

Belajar adalah proses untuk mengubah perilaku siswa dari yang tadinya tidak bisa menjadi bisa dalam bidang yang dipelajari yang diukur dari aspek pengetahuan (kognitif), sikap (afektif) maupun keterampilan (psikomotor). Definisi belajar tersebut diperkuat oleh beberapa tokoh pendidikan yang dikutip oleh Ady Ferdian Noor yaitu: (1) Travers: belajar adalah proses menghasilkan penyesuaian tingkah laku; (2) Cronbach dan Morgan: Belajar adalah perubahan perilaku yang bersifat permanen sebagai hasil dari pengalaman; (3) Geoch: Belajar adalah perubahan *performance*/penampilan sebagai hasil latihan (Noor, 2013).

Proses belajar mengajar (PBM) merupakan suatu bentuk interaksi antara peserta didik dengan pendidik. Kegiatan belajar dilakukan oleh peserta didik sedangkan kegiatan mengajar dilakukan oleh pendidik. PBM bertujuan men-transfer ilmu pengetahuan, keterampilan, sikap ataupun nilai-nilai dari pendidik kepada peserta didik yang menjadi objek belajar. Dalam PBM digunakan berbagai macam metode dan media agar ilmu pengetahuan, keterampilan, sikap yang ditransfer lebih mudah diterima peserta didik. Untuk mengetahui hasil belajar dilakukan evaluasi hasil belajar. Berdasarkan ilustrasi tersebut dapat diidentifikasi komponen pembelajaran yaitu terdiri atas: (1) peserta didik dan pendidik; (2) metoda dan media; (3) tujuan pembelajaran, materi yang dipelajari dan evaluasi hasil belajar. Contoh pembelajaran tertera pada Gambar 2.1. Tugas saudara; identifikasilah komponen apa saja yang terdapat pada gambar tersebut.

<p>TUJUAN belajar agar siswa dapat:</p> <ul style="list-style-type: none"> • Tahu • Paham • Bisa • Terampil.... 		<p>Sumber: https://iris.peabody.vanderbilt.edu/module/cnm/cresource/q2/p07/</p>	<p>EVALUASI hasil belajar yang sudah dicapai:</p> <ul style="list-style-type: none"> • Tahu • Paham • Bisa • Terampil •Yes or No
<p>Input</p>	<p>Proses ada interaksi antar input guru, siswa, materi, media dan metode</p>	<p>Output</p>	

Gambar 2.1 Proses Pembelajaran

PBM dinyatakan efektif jika tujuan pembelajaran dapat tercapai. Untuk mengetahui ketercapaian tujuan pembelajaran dilakukan pengukuran, penilaian dan evaluasi. Pengetahuan dan keterampilan guru dalam melakukan evaluasi sangat membantu untuk mengetahui ketercapaian hasil belajar.

B. Faktor Penentu Hasil belajar

Hasil belajar diperoleh secara bertahap melalui proses belajar. Hasil belajar perlu dicek kemajuannya tahap demi tahap. Hal ini penting karena keberhasilan belajar tahap pertama sangat menentukan keberhasilan belajar tahap berikutnya. Jika kesulitan belajar tahap pertama tidak diatasi maka belajar tahap kedua akan semakin sulit dan dapat diprediksi hasil belajar akan gagal. Jika kesulitan belajar pada tahap pertama tidak diatasi maka akan berakibat seperti yang tertera pada ilustrasi di Gambar 2.1. Namun, jika kesulitan belajar pada tahap pertama dapat diatasi, maka pada tahap berikutnya siswa mampu belajar mandiri sehingga motivasi belajar meningkat dan hasil belajar sukses seperti terdapat pada contoh di Gambar 2.2.



Gambar 2.2 Pengalaman Gagal



Gambar 2.3
Pengalaman Sukses

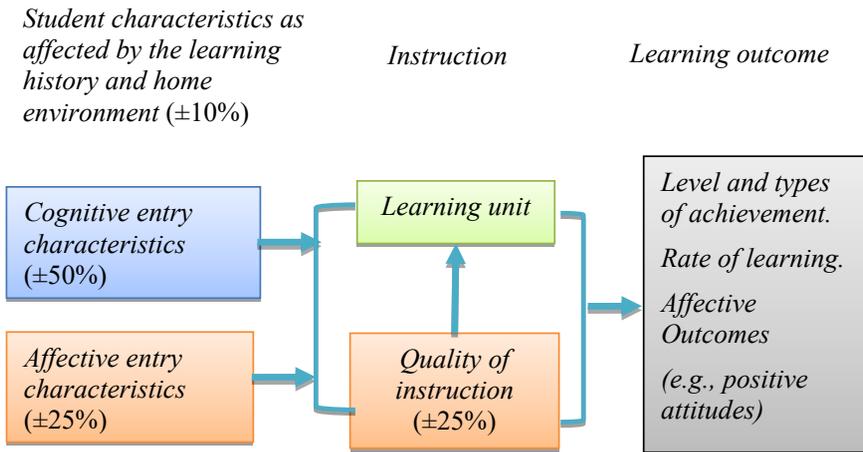
Motivasi menentukan persepsi individu tentang prestasi sekarang dan prestasi berikutnya. Motivasi dapat membangkitkan usaha yang diperlukan untuk belajar atau usaha ketika siswa menghadapi kesulitan dan frustrasi. Pengalaman sukses yang dihargai atau diberi nilai dapat memberi pengaruh positif bagi siswa. Penilaian mempunyai pengaruh terhadap motivasi dan sikap belajar pada waktu masuk dan prestasi belajar pada waktu keluar dari sekolah. Apabila siswa masuk dengan sikap belajar negatif, maka mereka membutuhkan pembelajaran yang berkualitas tinggi untuk mencapai prestasi yang bagus. Siswa yang masuk dengan sikap belajar sangat positif dapat mengalami kegagalan apabila mereka kurang mempunyai potensi kognitif. Sikap belajar positif yang dikombinasi dengan potensi kognitif yang memadai dapat membuat siswa lebih mudah belajar jika diberi pembelajaran yang berkualitas.

Bloom dalam Roid & Haladyna menggambarkan model regresi linier untuk menjelaskan tentang pengaruh potensi kognitif dan afektif terhadap prestasi belajar sebagai berikut:

$$\hat{Y} = F(X_1 + X_2 + X_3) \text{ Sumber (Roid \& Haladyna, 1982)}$$

Dimana \hat{Y} adalah prestasi yang diukur dengan test acuan kriteria, X_1 adalah *cognitive entry characteristics*, X_2 adalah *affective entry characteristics* seperti sikap dan kepribadian, dan X_3 adalah kualitas pembelajaran. Bloom's menganalisis tiga konstruk yang berhubungan dengan belajar yaitu *cognitive entry characteristics* menyumbang sekitar 50%, *affective entry characteristic* menyumbang sekitar 25%, sedangkan sekitar 25% sisanya dipengaruhi oleh kualitas pembelajaran. Apabila ketiga variabel

(kognitif, afektif dan kualitas pembelajaran) tersebut dijumlahkan maka dapat diprediksi menyumbang sekitar 90% prestasi belajar dan 10% sisanya dipengaruhi oleh sejarah keluarga, lingkungan rumah, dan latar belakang lain. Bloom menjelaskan pengaruh kognitif, afektif dan kualitas pembelajaran terhadap prestasi belajar seperti pada Gambar 2.4.



Sumber: (Bloom, 1976)

Gambar 2.4 Pengaruh Karakteristik Masuk dan Kualitas Pembelajaran pada *Learning Outcome*.

Karakteristik pertama yang dipercaya dapat menentukan prestasi belajar adalah *cognitive entry behaviors* siswa dan karakteristik yang kedua adalah *the affective entry characteristics* yang memotivasi siswa untuk belajar tugas-tugas baru. Kualitas pembelajaran yang sesuai dengan kebutuhan siswa merupakan variabel yang menentukan hasil belajar siswa secara bersama-sama. Apabila siswa memiliki *prerequisite* (kognitif dan afektif) yang baik kemudian dilayani dengan pembelajaran yang berkualitas sesuai dengan kebutuhan maka dapat diprediksi semua siswa akan mampu belajar tugas-tugas baru dengan tingkat pencapaian prestasi atau waktu yang diperlukan untuk penyelesaian tugas yang hampir sama. Bloom (Bloom, 1976) mengutip definisi motivasi dari Carroll (1963) yang dianalogikan dengan istilah *perseverance* (ketekunan) individu dalam mengerjakan tugas atau kemauan berusaha untuk menyelesaikan tugas. Karakteristik afektif merupakan campuran kompleks antara minat, sikap, dan pandangan terhadap diri sendiri (*self-views*).

Prestasi belajar dapat diprediksi dari prestasi pada mata pelajaran yang menjadi prasyarat belajar (*prerequisite*). Skor komposit (gabungan) tes bakat dan prestasi (aritmatik, matematik, membaca, bahasa) berkorelasi dengan prestasi belajar dalam rentangan antara 0,5 s/d 0,7. *Prerequisite* keterampilan dasar kuantitatif (matematika) diperlukan tidak hanya dalam pelajaran matematika dan sains tetapi juga dalam tugas-tugas belajar sekolah (Bloom, 1976). Skor *reading comprehension* mempunyai korelasi tinggi terhadap prestasi membaca literatur. Jika materi (*subject matter*) pelajaran banyak diberikan dalam bentuk literatur atau buku teks maka siswa harus memiliki kemampuan membaca tinggi. Korelasi *reading comprehension* terhadap hasil belajar sekitar 0,7 tetapi masih tergantung pada berat dan ringannya bacaan.

Hamilton melaporkan hasil penelitian ada hubungan antarprestasi siswa dipengaruhi oleh praktik perubahan yang dilakukan guru, latar belakang siswa seperti status sosio ekonomi dan etnik (Hamilton, 2003). Willingham, Pollack, Lewis menemukan enam kelompok variabel yang berpengaruh terhadap prestasi siswa, yaitu: karakteristik kemampuan siswa, inisiatif, kegiatan kompetisi, latar belakang keluarga, sikap dan penilaian guru (Willingham, Pollack, & Lewis, C., 2002). Faktor penilaian guru berpengaruh paling tinggi (0,9) sedangkan karakteristik kemampuan siswa menunjukkan pengaruh yang lebih kecil yaitu 0,86. Faktor non tes yang sering mempengaruhi penilaian guru seperti: kehadiran, perilaku di kelas, kelengkapan tugas, motivasi, dan bimbingan guru. Hasil penelitian ini mengindikasikan bahwa penilaian yang dilakukan guru kurang objektif karena ada beberapa hal di luar kemampuan kognitif yang menjadi pertimbangan guru.

Skor tes tunggal yang diperoleh siswa tidak dapat menunjukkan kemampuan siswa secara mutlak. Beberapa siswa mempunyai perbedaan skor karena perbedaan latar belakang keluarga, budaya, kurikulum, jenis kelamin, dan penilaian subjektif guru. Perbedaan kemampuan antara pria dan wanita dalam matematika diteliti oleh (Gierl, Bisanz & Boughton, 2003) dan Bielinski (Bielinski & Davison, M. L, 2001) menggunakan metode analisis DIF (*differential item functioning*). Hasil penelitian mengungkapkan wanita mempunyai kinerja yang lebih baik dalam item keterampilan verbal dan pada item yang menentukan kematangan isi matematika. Pria mempunyai kinerja yang lebih baik dari wanita pada item yang mempunyai karakteristik kontekstual seperti keterampilan spasial, dan item yang mempunyai jalur pemecahan masalah ganda.

Penelitian yang dilakukan oleh Mehrens (Mehrens & Lehman, 1973) mempelajari perbedaan kurikulum terhadap skor tes menunjukkan hasil ‘tidak ada buku teks yang tidak sepadan dengan kurikulum yang diujikan’ dan tidak ada perbedaan skor tes membaca buku teks namun guru sering membuat klaim bahwa perbedaan skor yang diperoleh karena perbedaan kurikulum. Jika soal yang disusun sudah sesuai dengan materi yang terdapat dalam kurikulum, maka hasil tes dipengaruhi oleh kemampuan peserta didik dan kemampuan guru mengajarkan materi sampai peserta didiknya paham dan mampu.

C. Peran Evaluasi dalam Proses Pembelajaran

Salah satu rangkaian kegiatan pembelajaran adalah melaksanakan evaluasi selama dan sesudah proses pembelajaran. Selama proses pembelajaran, evaluasi dilakukan untuk mengetahui kesesuaian antara strategi/metode, media pembelajaran dengan materi pelajaran dan kemajuan belajar peserta didik. Sesudah proses pembelajaran, evaluasi dilakukan untuk mengetahui tingkat pencapaian hasil belajar oleh setiap peserta didik. Burden & Byrd (Burden & Byrd, 1998) menegaskan evaluasi diperlukan untuk membantu guru mengetahui tingkat ketercapaian tujuan pendidikan dan membantu guru mengetahui posisi siswa secara individual.

Evaluasi dilakukan menggunakan data pengukuran dan asesmen. Pengukuran dan asesmen dalam pendidikan mempunyai tiga fungsi, yaitu:

- (1) melayani kebutuhan psikologis seseorang dalam memperoleh informasi tentang potensi yang dimiliki untuk dijadikan acuan dalam menentukan arah ke masa depan;
- (2) melayani kebutuhan instruksional, untuk mengetahui kesesuaian materi pelajaran, metode pembelajaran dengan kemampuan dan potensi peserta didik dan mengetahui posisi masing-masing siswa di dalam kelas;
- (3) melayani kebutuhan administrasi untuk mengetahui peringkat atau indeks prestasi siswa (Suryabrata, 1981).

Evaluasi memiliki peran yang penting dalam kegiatan pembelajaran. Informasi hasil evaluasi belajar memberi dampak pada peningkatan proses dan hasil belajar. Sebagai contoh, ketika evaluasi hasil belajar digunakan untuk menetapkan kelulusan maka peserta didik yang berada dalam keadaan normal akan berusaha untuk belajar supaya dapat lulus. Ketika

guru memberi umpan balik hasil evaluasi belajar secara rutin, maka peserta didik akan termotivasi untuk meningkatkan kinerja belajar secara terus menerus. Dengan motivasi belajar yang tinggi diharapkan suasana pembelajaran di kelas menjadi lebih kondusif, peserta didik lebih aktif sehingga hasil belajar lebih mudah dicapai. Ilustrasi dampak penilaian terhadap motivasi dan dampak motivasi terhadap prestasi diilustrasikan pada Gambar 2.5 berikut.



Gambar 2.5 Dampak Penilaian terhadap Motivasi Belajar

Evaluasi proses pembelajaran dilakukan apabila pendidik ingin mengetahui efektivitas proses pembelajaran yang telah dilaksanakan. Evaluasi proses pembelajaran dilaksanakan pada saat pembelajaran berlangsung, setiap akhir proses pembelajaran, tengah semester atau akhir semester. Selama proses pembelajaran minimal ada beberapa komponen yang terlibat antara lain: pendidik, peserta didik, materi/bahan ajar, strategi penyampaian materi, dan media/perangkat pembelajaran lainnya. Proses pembelajaran dinyatakan efektif apabila telah mampu memberdayakan semua komponen pembelajaran dalam mencapai tujuan/hasil yang diinginkan. Untuk mengetahui efektivitas proses pembelajaran tersebut dapat dilakukan dengan cara melihat:

- (1) Apakah **strategi** pembelajaran yang dipilih dan dipergunakan oleh pendidik (guru/dosen) sudah tepat, menarik dan menyenangkan?
- (2) Apakah **media/perangkat** pembelajaran yang digunakan oleh pendidik sudah dapat membantu pemahaman peserta didik?
- (3) Apakah **gaya mengajar** pendidik sudah menarik dan menyenangkan sehingga peserta didik (mahasiswa) antusias untuk mengikutinya?
- (4) Apakah **materi/bahan ajar** yang dibahas pendidik sudah sesuai dengan tujuan pembelajaran dan mendukung pencapaian kompetensi?
- (5) Bagaimanakah **interaksi** pendidik dan peserta didik berlangsung cukup harmonis?
- (6) Apakah **peserta didik** sudah terlibat aktif dalam proses pembelajaran?



<https://gifer.com/en/LYcL>

Peran evaluasi bagi guru dapat memperoleh umpan balik untuk perbaikan proses pembelajaran berikutnya

Gambar 2.6 Peran Evaluasi bagi Guru

Untuk mendapat informasi yang telah ditulis di atas, diperlukan beberapa instrumen/alat evaluasi, antara lain lembar observasi atau angket. Instrumen berisi butir-butir evaluasi kegiatan pendidik (guru/dosen) dalam melaksanakan pembelajaran dan kegiatan peserta didik dalam mengikuti pembelajaran. Pelaksanaan pembelajaran di tingkat pendidikan dasar dan menengah dievaluasi dengan lembar observasi yang diisi oleh guru sejawat, kepala sekolah dan pengawas sekolah. Pelaksanaan pembelajaran di perguruan tinggi biasanya dilakukan dengan kuesioner yang diisi oleh mahasiswa. Contoh instrumen (lembar observasi) pelaksanaan pembelajaran yang dikutip dari buku pedoman penyusunan portofolio guru (Dikti, 2010) adalah:

Tabel 2.1 Contoh Lembar Observasi Pelaksanaan Pembelajaran

NO	INDIKATOR/ASPEK YANG DIAMATI
I	PRA PEMBELAJARAN
1	Mempersiapkan siswa untuk belajar
2	Melakukan kegiatan apersepsi
II	KEGIATAN INTI PEMBELAJARAN
A	Penguasaan materi pelajaran
3	Menunjukkan penguasaan materi pelajaran
4	Mengaitkan materi dengan pengetahuan lain yang relevan
5	Menyampaikan materi dengan jelas, sesuai dengan hierarki belajar dan karakteristik siswa
6	Mengaitkan materi dengan realitas kehidupan
B	Pendekatan/strategi pembelajaran
7	Melaksanakan pembelajaran sesuai dengan kompetensi (tujuan) yang akan dicapai dan karakteristik siswa
8	Melaksanakan pembelajaran secara runtut
9	Menguasai kelas
10	Melaksanakan pembelajaran yang bersifat kontekstual
11	Melaksanakan pembelajaran yang memungkinkan timbulnya kebiasaan positif
12	Melaksanakan pembelajaran yang sesuai dengan alokasi waktu yang direncanakan

C	Pemanfaatan sumber belajar/media pembelajaran
13	Menggunakan media secara efektif dan efisien
14	Menghasilkan pesan yang menarik
15	Melibatkan siswa dalam pemanfaatan media
D	Pembelajaran yang memicu dan memelihara keterlibatan siswa
16	Menumbuhkan partisipasi aktif siswa dalam pembelajaran
17	Menunjukkan sikap terbuka terhadap respon siswa
18	Menumbuhkan keceriaan dan antusiasme siswa dalam belajar
E	Penilaian proses dan hasil belajar
19	Memantau kemajuan belajar selama proses
20	Melakukan penilaian akhir sesuai dengan kompetensi (tujuan)
F	Penggunaan bahasa
21	Menggunakan bahasa lisan dan tulis secara jelas, baik dan benar
22	Menyampaikan pesan dengan gaya yang sesuai
III	PENUTUP
23	Melakukan refleksi atau membuat rangkuman dengan melibatkan siswa
24	Melaksanakan tindak lanjut dengan memberi arahan, atau kegiatan atau tugas sebagai bagian remidi/pengayaan
	Total Skor

Pendidik dapat mengembangkan alat evaluasi pelaksanaan pembelajaran yang lebih sederhana dari contoh lembar observasi di atas. Data hasil evaluasi pelaksanaan pembelajaran kemudian dianalisis secara deskriptif kuantitatif dan kualitatif. Hasil evaluasi ini dapat menjadi umpan balik bagi pendidik untuk memperbaiki proses pembelajaran.

D. Prinsip-prinsip Penilaian

Berdasarkan Permendikbud Nomor 23 Tahun 2016 tentang Standar Penilaian Pendidikan, dalam setiap aktivitas penilaian pendidikan tidak dapat dilepaskan dari prinsip-prinsip penilaian sebagai berikut:

- 1) **sahih**, berarti penilaian didasarkan pada data yang mencerminkan kemampuan yang diukur; materi yang diujikan sesuai dengan kurikulum, silabus, kisi-kisi materi dan indikator pencapaian kompetensi.

- 2) **objektif**, berarti penilaian didasarkan pada prosedur dan kriteria yang jelas, tidak dipengaruhi subjektivitas penilai;
- 3) **adil**, berarti penilaian tidak menguntungkan atau merugikan peserta didik karena berkebutuhan khusus serta perbedaan latar belakang agama, suku, budaya, adat istiadat, status sosial ekonomi, dan gender;
- 4) **terpadu**, berarti penilaian merupakan salah satu komponen yang tak terpisahkan dari kegiatan pembelajaran;
- 5) **terbuka**, berarti prosedur penilaian, kriteria penilaian, dan dasar pengambilan keputusan dapat diketahui oleh pihak yang berkepentingan;
- 6) **menyeluruh dan berkesinambungan**, berarti penilaian mencakup semua aspek kompetensi dengan menggunakan berbagai teknik penilaian yang sesuai, untuk memantau dan menilai perkembangan kemampuan peserta didik;
- 7) **sistematis**, berarti penilaian dilakukan secara berencana dan bertahap dengan mengikuti langkah-langkah baku;
- 8) **beracuan kriteria**, berarti penilaian didasarkan pada ukuran pencapaian kompetensi yang ditetapkan;
- 9) **akuntabel**, berarti penilaian dapat dipertanggungjawabkan, baik dari segi mekanisme, prosedur, teknik, maupun hasilnya; dan
- 10) **andal**, berarti dapat dipercaya dan memberikan hasil yang konsisten pada ujian atau pengukuran yang berulang.

E. Tanggung Jawab Profesional dan Etika Guru dalam Penilaian

Pendidik (guru dan dosen) memiliki tugas pokok dan fungsi yang berbeda. Menurut Undang-undang Guru dan Dosen, guru adalah pendidik profesional dengan tugas utama mendidik, mengajar, membimbing, mengarahkan, melatih, menilai, dan mengevaluasi peserta didik pada pendidikan anak usia dini jalur pendidikan formal, pendidikan dasar, dan pendidikan menengah. Dosen adalah pendidik profesional dan ilmuwan dengan tugas utama mentransformasikan, mengembangkan, dan menyebarkan ilmu pengetahuan, teknologi, dan seni melalui pendidikan, penelitian, dan pengabdian kepada masyarakat. Profesional adalah pekerjaan atau kegiatan yang dilakukan oleh seseorang dan menjadi sumber penghasilan kehidupan yang memerlukan keahlian, kemahiran,

atau kecakapan yang memenuhi standar mutu atau norma tertentu serta memerlukan pendidikan profesi. Guru dan dosen dapat bekerja profesional jika dia telah kompeten dalam bidangnya. Kompetensi adalah seperangkat pengetahuan, keterampilan, dan perilaku yang harus dimiliki, dihayati, dan dikuasai oleh guru atau dosen dalam melaksanakan tugas keprofesionalan. Bukti formal sebagai pengakuan yang diberikan kepada guru dan dosen sebagai tenaga profesional diberikan dalam bentuk sertifikat pendidik.

Guru dan dosen sering mengalami kendala untuk bekerja profesional. Guru dan dosen terpaksa meninggalkan beberapa tugas pokok dan bekerja kurang profesional karena sibuk atau beban kerja berlebihan. Dalam penilaian pendidikan, guru dan dosen harus membuat banyak keputusan pembelajaran. Dalam kegiatan ini ternyata ada banyak hal yang dapat membuat keputusan kurang adil bagi siswa. Alat-alat penilaian dapat membantu guru untuk membuat keputusan. Masing-masing keputusan akan memberi konsekuensi positif atau negatif kepada siswa. Menurut Nitko (Nitko & Brookhart, 2011) terdapat 6 tanggung jawab profesional guru dalam penilaian yaitu:

- 1) membuat prosedur asesmen yang berkualitas.
- 2) memilih prosedur asesmen yang sesuai untuk digunakan.
- 3) mengadministrasikan semua prosedur asesmen yang adil untuk semua siswa.
- 4) mensekor hasil asesmen, mengevaluasi jawaban dengan akurat dan melaporkan hasilnya kepada siswa.
- 5) menginterpretasikan dan menggunakan hasil asesmen yang valid, mempromosikan dampak positif dan meminimalkan dampak negatif terhadap siswa.
- 6) mengomunikasikan hasil asesmen.

Nitko (Nitko & Brookhart, 2011) memberi penjelasan bawa dalam membuat prosedur asesmen yang berkualitas, tanggung jawab umum pada saat guru adalah membuat prosedur asesmen yang valid dan reliabel. Tanggung jawab khusus pada saat guru membuat prosedur asesmen: (a) menerapkan sejumlah prinsip perencanaan asesmen, perancangan asesmen, pengembangan tugas, penulisan butir, pengembangan rubrik, memberi nilai pada masing-masing asesmen formal yang digunakan; (b) membuat prosedur asesmen yang relevan untuk mengukur target belajar.

Prosedur asesmen tidak menguntungkan atau merugikan status gender, etnik, ras, kelas sosial dan agama tertentu; (c) mengakomodasi cara yang sesuai untuk siswa berkesulitan belajar atau berkebutuhan khusus; (d) meminta ijin kepada yang berhak untuk menggunakan salinan materi yang akan dibuat soal asesmen; (e) menjamin bahwa materi atau isi sudah akurat, perintah/pertanyaan dan kunci jawaban sudah benar sehingga dapat dijamin tidak ada kesalahan (*error*) sedikitpun. Apabila terjadi kesalahan setelah dilakukan asesmen, segera betulkan dan lakukan penyekoran ulang.

Asesmen dapat dilakukan dalam berbagai bentuk dan cara. Tanggung jawab profesional pendidik ketika memilih alat asesmen adalah mengecek kesesuaian materi tes dengan materi yang diajarkan dan kualitas isi asesmen. Fakta yang ditemukan adalah asesmen yang dipublikasikan tidak menjamin isi atau kualitas sebagai sebuah alat asesmen. Banyak tugas-tugas asesmen tidak disiapkan dan direview secara profesional sehingga isi, kalimat pertanyaan, jawaban benar masih mengandung kesalahan. Mengevaluasi alat asesmen sangat penting sebelum digunakan.

Guru dan dosen memiliki tanggung jawab profesional dalam **mengadministrasikan semua prosedur asesmen yang adil untuk semua siswa**. Etika profesi yang harus dipatuhi dalam konteks ini adalah guru berkewajiban memberi informasi yang mencukupi untuk siswa yaitu: (a) ketika asesmen akan dimulai, (b) kondisi kinerja yang harus ditunjukkan, (c) isi dan kemampuan yang akan dinilai, (d) komponen penilaian yang akan ditekankan; (e) standar atau tingkat kinerja yang diharapkan, bagaimana penyekoran kinerja, dan pengaruh hasil untuk beberapa keputusan. Jika di sekolah terdapat siswa berkesulitan belajar, guru wajib mengakomodasi siswa berkesulitan belajar. Asesmen harus dimodifikasi agar siswa yang berkesulitan belajar dapat mengikuti. Nitko (Nitko & Brookhart, 2011) menjelaskan 5 cara untuk mengakomodasi siswa yang berkesulitan belajar yaitu: (a) *presentation* dengan cara mengulangi perintah, membacakan pelan, memperbesar huruf atau menggunakan huruf braille; (b) *equipment and material*, yaitu perlengkapan dipenuhi seperti kalkulator, audio dan alat manipulasi lain; (c) *response* yaitu membantu siswa menghitami jawaban atau merekam jawaban; (d) *timing/scheduling* yaitu dengan memperpanjang waktu ujian atau sering beristirahat jika kelelahan; (e) *setting* ruang yang terpisah.

Jika sekolah wajib mengikuti tes standar nasional yaitu tes yang diselenggarakan oleh negara, tanggung jawab profesional guru selama

ujian standar nasional adalah: (a) memberi informasi kepada siswa dan orangtua siswa tentang ujian standar dan bagaimana hasil ujian tersebut akan digunakan; (b) menjaga kondisi ujian agar siswa dapat mengerjakan ujian secara maksimum; (c) skor siswa diinterpretasikan menggunakan norma yang berlaku secara nasional. Praktik-praktik yang tidak etis pada saat ujian standar nasional berlangsung adalah jika guru memberitahu siswa mengisi jawaban yang benar, atau mengubah jawaban siswa untuk memperbaiki skor. Apabila guru melihat ada guru yang melakukan hal tidak etis tersebut maka wajib segera melaporkannya.

Setelah selesai penyelenggaraan ujian, guru memiliki tanggung jawab profesional untuk **menilai hasil ujian**. Etika yang harus dipenuhi oleh guru adalah: (a) memberi skor jawaban siswa secara akurat dengan menggunakan kunci jawaban, skor, rubrik, checklist, atau skala penilaian yang sesuai; (b) supaya hasil penilaian objektif maka dilakukan dengan cara memberi skor secara terbuka; memberi skor dengan tidak melihat nama siswa, memberi skor jawaban esai pada satu nomor soal yang sama untuk semua siswa sebelum beralih pada nomor yang lain; (c) memberi umpan balik kepada siswa agar mereka memiliki peluang untuk memperbaiki belajarnya; (d) memberi penjelasan tentang jawaban yang benar secara rasional, dan cara penilaian menggunakan rubrik; (e) memberi kesempatan kepada siswa untuk mengevaluasi jawabannya sendiri; (d) mengoreksi kesalahan penilaian dan membuat keputusan secepat mungkin; (e) memberi nilai dan mengumumkan hasil penilaian dalam waktu yang telah ditetapkan.

Setelah jawaban ujian diberi skor, guru memiliki tanggung jawab profesi untuk **menginterpretasikan dan menggunakan hasil asesmen** untuk: (a) mengetahui apakah target belajar yang ditetapkan sudah tercapai; (b) membandingkan kinerja siswa dengan hasil asesmen lain; (c) membantu siswa dan orangtua siswa menginterpretasikan hasil penilaian yang diperoleh; (d) membantu siswa dan orangtua siswa memahami konsekuensi hasil penilaian kelas untuk mengetahui posisi siswa diantara siswa lainnya; (e) menginterpretasikan kinerja siswa untuk mengevaluasi pencapaian target belajar mereka sebagai senjata untuk menghukum/memotivasi dan mengendalikan perilaku siswa; (f) menjaga hak pribadi siswa jika hasil penilaian kelasnya tidak ingin diketahui orang lain. Tahap terakhir yang menjadi tanggung jawab profesional guru adalah **mengomunikasikan hasil asesmen** kepada siswa, orangtua siswa dan sekolah untuk berbagai kepentingan, antara lain melaporkan kemajuan hasil belajar siswa (Nitko & Brookhart, 2011).

BAB III

ANALISIS KOMPETENSI DAN TUJUAN PEMBELAJARAN

A. Analisis Kompetensi

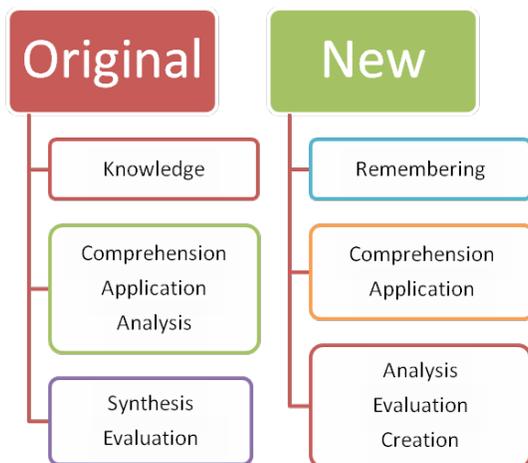
Hasil belajar pendidikan kejuruan dinyatakan dalam rumusan kompetensi. Menurut definisi dari dictionary.com (<http://dictionary.reference.com>) kompetensi adalah kualitas yang menjadikan kompeten, keadekuatan, kepemilikan keterampilan, pengetahuan, kualifikasi atau kapasitas yang dituntut (*the quality of being competent; adequacy; possession, of required skill, knowledge qualification, or capacity*). Dari sumber yang sama, British dictionary mendefinisikan kompetensi sebagai kondisi yang menjadikan mampu atau kemampuan (*the condition of being capable; ability*). Jika dikaitkan dengan pekerjaan maka kompetensi berarti kemampuan/keterampilan untuk menyelesaikan pekerjaan tertentu. Orang yang mampu bekerja cepat, tepat dan berkualitas menunjukkan dia sudah kompeten.

Hasil belajar pendidikan kejuruan ditetapkan dengan standar kompetensi lulusan yaitu kualifikasi kemampuan lulusan yang mencakup sikap, pengetahuan, dan keterampilan (Permendiknas nomor 19 tahun 2005 tentang Standar Nasional Pendidikan). Kompetensi hasil belajar sering disingkat KSAO (*cognitive, affective, psycomotor, and other*) atau 3H (*Head, Heart, Hand*). Spektrum SMK terdiri atas banyak keahlian, oleh sebab itu selain kompetensi umum yang terdiri atas pengetahuan, sikap dan keterampilan juga terdapat kompetensi khusus yang membedakannya yaitu kompetensi lainnya (*other*) sesuai dengan tuntutan keahlian.

Kompetensi hasil belajar ranah pengetahuan, sikap dan keterampilan memiliki tingkatan dari rendah sampai mahir. Pengelompokan tingkatan kompetensi dikenal dengan istilah taksonomi (*taxonomy*). Ada beberapa tokoh yang menyusun tingkatan kompetensi hasil belajar antara lain Bloom; Krathwohl, Dave, Simpson dan Gagne. Berikut diuraikan tingkatan kompetensi tiap-tiap ranah.

1. Cognitive Domain

Bloom (Krathwohl, Bloom, B.S, & Masia, 1964) mengklasifikasikan kompetensi *cognitive* menjadi enam tingkat dari rendah atau LOTS (*lower order thinking skills*), sedang atau MOTS (*middle order thinking skills*) sampai tingkat tinggi atau HOTS *higher order thinking skills*) seperti tertera pada Gambar 3.1. Kompetensi kognitif tingkat paling rendah/LOTS adalah mengingat sedangkan tingkat yang paling tinggi (HOTS) adalah analisis, evaluasi, dan kreasi. Domain kognitif dari Bloom dinyatakan pada Gambar 3.1.



Gambar 3.1 Taksonomi Kompetensi Kognitif

Sub-Domain Cognitive

a. Ingatan (*remembering*)

“Mengingat” merupakan level domain kognitif yang paling rendah. Evaluasi hasil belajar sebaiknya tidak terlalu banyak mengukur aspek “mengingat” karena dapat merangsang siswa menyontek dan kurang mampu

melatih siswa untuk berpikir kreatif memecahkan masalah. Kompetensi kognitif level “mengingat” biasanya diukur dari *kemampuan mengenali dan mengingat peristilahan, definisi, fakta-fakta, gagasan, pola, urutan, metodologi, prinsip dasar, peristiwa, kejadian, sejarah, dan informasi yang telah diterima sebelumnya*. Dalam perumusan tujuan pembelajaran, kata-kerja operasional yang sering digunakan untuk mengukur kemampuan “ingatan” adalah sebagai berikut.

KK operasional	
mengidentifikasi (<i>identify</i>)	mendeskripsikan (<i>describe</i>)
mendefinisikan (<i>define</i>)	menyebutkan kembali (<i>recall</i>)
menamai (<i>name</i>)	melengkapi (<i>complete</i>)
mencocokkan	mendaftar (<i>list</i>)
memasangkan	menceritakan (<i>recite</i>)
menghapal	menirukan/mengulangi

Kata kerja operasional digunakan untuk menyusun tujuan pembelajaran dan evaluasi hasil belajar. Tujuan pembelajaran dan evaluasi hasil belajar harus konsisten karena tujuan pembelajaran harus dapat dicapai saat akhir pembelajaran. Ada cara yang dapat digunakan untuk merumuskan kata kerja operasional yaitu dengan mengubah kata kerja pada tujuan pembelajaran menjadi kata tanya atau perintah pada saat mengevaluasi hasil belajar.

Contoh:

Tujuan Pembelajaran

- Siswa dapat mendefinisikan pengertian masakan tradisional.
- Siswa dapat menyebutkan 10 nama masakan khas daerah Jawa Timur.

Soal evaluasi hasil belajar

- Definisikan arti masakan tradisional?
- Sebutkan 10 nama masakan khas daerah Jawa Timur!

b. Pemahaman (*comprehension*)

Pemahaman merupakan kompetensi kognitif level kedua setelah ingatan. Seseorang yang telah mendapat informasi atau pengetahuan, akan menyimpan informasi tersebut menjadi ingatan kemudian memanggil kembali (*retrieval*) informasi yang telah disimpan dalam “ingatan” dengan menggunakan kalimatnya sendiri. Jika kalimat yang diucapkan masih sama persis dengan kalimat yang diterima, berarti siswa belum mencapai level pemahaman tetapi masih mengingat-ingat atau menghafal.

Kompetensi kognitif level “pemahaman” biasanya diukur dari *kemampuan seseorang memaparkan kembali suatu prosedur kerja, proses, kejadian, materi ilmu pengetahuan dengan menggunakan bahasa atau kalimatnya sendiri*. Kata kerja yang sering digunakan untuk mengukur level pemahaman hampir sama dengan level ingatan tetapi respon (jawaban) yang dikehendaki lebih luas atau mendalam dan menggunakan bahasanya sendiri. Kata kerja operasional yang dapat digunakan dalam rumusan tujuan pembelajaran antara lain:

KK operasional	
menjelaskan	merinci
memaparkan	menguraikan
membedakan	membandingkan
menceritakan kembali	menerangkan

Contoh:

Tujuan pembelajaran

- Siswa dapat membedakan ciri-ciri wortel yang sudah tua dan masih muda.
- Siswa dapat menjelaskan prosedur pembuatan *nata de coco*.

Soal evaluasi hasil belajar

- Bedakan ciri-ciri wortel yang sudah tua dan masih muda.
- Jelaskan prosedur pembuatan *nata de coco*.

c. Penerapan (*application*)

Penerapan (*application*) merupakan kompetensikognitif level 3 setelah “ingatan dan pemahaman”. Tingkat kesulitan sub domain *application* berada satu peringkat di atas “ingatan dan pemahaman”. Kompetensi kognitif level penerapan mengukur “kemampuan untuk menggunakan dan menerapkan gagasan, prosedur, metode, rumus, teori dan informasi yang telah dipelajari ke dalam kondisi kerja atau konteks lain yang baru”. Evaluasi hasil belajar dilakukan dengan memberi masalah (problem) yang serupa dengan materi yang dipelajari, kemudian siswa dituntut memecahkan masalah dengan menggunakan ilmu yang dimiliki, cara atau prosedur yang sama. Kata kerja operasional yang sering digunakan dalam tujuan pembelajaran level aplikasi adalah:

KK operasional	
menghitung	menemukan
mengubah/memodifikasi	menunjukkan
memilih	meramal/memprediksi
menggunakan	menyiapkan
mengoperasikan	memecahkan masalah

Contoh:

Tujuan pembelajaran

- Siswa dapat menghitung kebutuhan bahan makanan untuk memasak 100 porsi bakmie goreng.
- Siswa dapat memodifikasi/mengubah resep kroket kentang menjadi kroket talas.
- Siswa dapat memilih teknik membuat adonan yang tepat untuk mengolah cake pisang.
- Siswa dapat menggunakan teori 7P dalam penyusunan strategi pemasaran.

Soal evaluasi hasil belajar

- Hitunglah kebutuhan bahan makanan untuk memasak 100 porsi bakmi goreng!
- Ubahlah resep kroket kentang menjadi kroket talas!
- Pilihlah teknik membuat adonan yang tepat untuk mengolah cake pisang!
- Susun strategi pemasaran menggunakan teori 7P!

d. Analisis (*analysis*)

Kompetensi kognitif level analisis berada di atas level *application*. Tingkat kesulitan level analisis hampir sama dengan level evaluasi dan sintesis atau kreasi. Kompetensi kognitif level analisis mengukur *kemampuan memisahkan materi atau konsep ke dalam bagian-bagian untuk diorganisasikan kembali menjadi struktur yang mudah dipahami*. Kata kerja operasional yang digunakan untuk menyusun tujuan pembelajaran level analisis adalah:

KK operasional	
menganalisis	membuat kerangka berpikir
mendiagnosis	mengilustrasikan
mem-breakdown	memberi alasan
mengidentifikasi masalah	memilah

Contoh:

Tujuan pembelajaran:

- Siswa dapat mendiagnosis sebab-sebab tekstur mayonaise pecah atau tidak homogen.
- Siswa dapat menganalisis sebab-sebab bolu yang dipanggang tidak mengembang.
- Siswa dapat memberi alasan mengapa terjadi *case hardening* pada daging beku yang direbus.

Soal evaluasi hasil belajar:

Untuk menilai kemampuan siswa berpikir analisis, ada beberapa kata tanya yang menjadi kunci untuk mengawali pertanyaan yaitu: Apa sebab?; Mengapa? ; Bagaimana? Contoh pertanyaan yang menuntut siswa berpikir pada level analisis:

- Mengapa tekstur mayonaise pecah atau tidak homogen?
- Apa penyebab bolu tidak bisa mengembang?
- Mengapa daging beku yang direbus mengalami *case hardening*?

e. Evaluasi (*evaluation*)

Evaluasi merupakan kompetensi kognitif level tinggi karena orang yang dapat mengevaluasi harus sudah tahu kriteria hasil yang benar, memberi contoh cara yang benar, dan memberi solusi cara memperbaiki pada hasil yang salah. Ranah kognitif yang diukur dalam level evaluasi adalah: *kemampuan membuat penilaian dan keputusan tentang nilai suatu gagasan, metode, produk atau benda dengan menggunakan kriteria tertentu*. Evaluasi hasil belajar ranah kognitif level “evaluasi” lebih banyak diberikan dalam bentuk tugas. Kata kerja operasional yang digunakan untuk menyusun tujuan pembelajaran level evaluasi adalah:

KK operasional	
mereview	mengevaluasi
menilai	mengoreksi
mengapresiasi	membetulkan
memperbaiki	menginterpretasikan
menghubungkan	mengkritik
menyimpulkan	mempertimbangkan

Contoh:

Tujuan pembelajaran

- Siswa dapat membetulkan komposisi bahan yang kurang tepat pada resep kue nastar.

- Siswa dapat mengoreksi kesalahan prosedur dalam membuat kue putu mayang.
- Siswa dapat mengoreksi kata kerja yang belum operasional dalam perumusan tujuan pembelajaran.
- Siswa dapat menilai kreativitas desain penyajian tumpeng nasi kuning.

Soal evaluasi hasil belajar

Untuk menilai kemampuan siswa melakukan evaluasi, guru perlu menyiapkan standar/kriteria terlebih dahulu. Evaluasi hasil belajar dilakukan dengan memberi tugas-tugas yang konsisten dengan tujuan pembelajaran, yaitu:

- Berikut terdapat resep nastar yang standar, carilah komposisi bahan yang kurang tepat pada resep kue nastar.
- Berikut terdapat resep putu mayang yang standar, betulkan prosedur pembuatan kue putu mayang yang belum runtut.
- Berikut ini terdapat kriteria dan daftar kata kerja operasional untuk menyusun tujuan pembelajaran. Perbaikilah susunan tujuan pembelajaran yang belum menggunakan kata kerja operasional.
- Nilailah kreativitas design penyajian tumpeng nasi kuning dengan menggunakan rubrik penilaian yang telah tersedia.

f. Sintesis/Kreasi

Kompetensi kognitif yang paling tinggi adalah menyusun (sintesis) atau membuat (kreasi). Dalam kompetensi ini, siswa sudah memiliki berbagai macam ilmu pengetahuan, kemudian menggunakan ilmunya untuk menyusun: rencana/program, proposal, karya ilmiah, laporan, persiapan praktik, design, dll. Konsep pengukuran kompetensi kognitif level sintesis atau kreasi adalah siswa mampu: *“membangun sebuah struktur atau pola dari berbagai elemen atau menggabungkan bagian-bagian untuk membentuk sebuah kesatuan yang utuh dengan penekanan pada hasil berupa sebuah pengertian atau struktur baru”*. Kata kerja operasional yang digunakan dalam rumusan pembelajaran adalah:

KK operasional	
menyusun	mengategorikan
merancang/mendesain	mengombinasikan
membuat konsep	menata
mengorganisasikan	membuat diagram

Contoh:

Tujuan pembelajaran

- Siswa dapat menyusun menu makan siang keluarga yang memiliki variasi bahan, teknik olah dan rasa.
- Siswa dapat membuat rencana usaha (*bussines plan*) catering Panji.
- Siswa dapat merancang strategi pemasaran dodol Carica.
- Siswa dapat mendesain penataan meja untuk penyajian menu prasmanan.

Soal evaluasi hasil belajar

Evaluasi hasil belajar level sintesis dapat berbentuk tugas atau proyek. Guru bertugas menyiapkan rubrik untuk menilai hasil belajar siswa. Contoh pernyataan tugas berdasarkan tujuan pembelajaran di atas adalah sebagai berikut.

- Susunlah menu makan siang keluarga yang memiliki variasi bahan, teknik olah dan rasa!
- Susunlah rencana usaha (*bussines plan*) catering Panji!
- Rancanglah strategi pemasaran dodol Carica!
- Desainlah penataan meja untuk penyajian menu prasmanan!

2. *Affective Domain*

Krathwohl, Bloom, Masia (Krathwohl, Bloom, B.S, & Masia, 1964) mengembangkan taksonomi tujuan yang berorientasi pada emosi seperti perasaan, nilai-nilai, apresiasi, antusiasme, motivasi, dan sikap. Ada lima kategori tujuan afektif yang dikemukakan, yaitu: penerimaan (*receiving*), pemberian respon (*responding*), menilai (*valuing*), pengorganisasian (*organization*), dan *Characterization* atau internalisasi (*internalizing*).

Penjelasan yang lebih mendalam dari lima kategori tujuan afektif dalam proses perubahan perilaku, membentuk perilaku baru yang menjadi kebiasaan (habit) hidup sehari-hari dapat diceritakan sebagai berikut:

- (1) Penerimaan (*receiving*) terjadi ketika siswa melihat atau menerima rangsang (stimuli) dari luar. Dalam tahap penerimaan, siswa menyadari atau merasakan keberadaan ide-ide, materi, atau fenomena dan bersedia untuk menerimanya. (*Receiving describes the stage of being aware of or sensitive to the existence of certain ideas, material, or phenomena and being willing to tolerate them*). Contoh perilaku: siswa pertama melihat orang membuang sampah sembarangan. Siswa masih bersikap pasif yaitu hanya melihat atau mendengar cerita tanpa memberi respon apapun.
- (2) *Responding* terjadi jika siswa terlibat aktif dalam menanggapi apa yang dilihat atau dirangsang dari luar. *Responding refers to a commitment in some small measure to the ideas, materials, or phenomena involved by actively responding to them. Responding* mengacu pada komitmen terhadap ide-ide, materi pelajaran, atau fenomena yang melibatkan mereka untuk aktif menanggapi. Respon berbentuk positif/negatif. Respon positif terjadi jika siswa mengikuti apa yang diharapkan dari pemberi rangsang (pendidik). Contoh sikap yang menunjukkan respon positif adalah: mematuhi, mengikuti, menanggapi, dll. Dalam konteks pembentukan sikap sehari-hari misalnya: ketika guru menyuruh siswa untuk mengambil dan membuang sampah pada tempatnya, siswa yang memberi respon positif akan mengambil dan membuang sampah sesuai dengan perintah guru;
- (3) Krathwohl menjelaskan *Valuing means being willing to be perceived by others as valuing certain ideas, materials, or phenomena*. Menilai (*valuing*) berarti kesediaan menghargai hal-hal yang dilakukan atau dibuat oleh orang lain seperti penyampaian ide-ide tertentu, materi, atau fenomena. Penilaian terjadi ketika seseorang sudah memberi respon positif (*affective level 2*), kemudian dilanjutkan dengan mulai menilai kelebihan dan keburukan. *Valuing* positif dilakukan dengan memberi dukungan, menghargai pendapat/ide. Dalam kasus pembiasaan perilaku membuang sampah, siswa mulai menilai baik jika membuang sampah pada tempatnya dan menilai buruk jika ada orang yang tidak mau membuang sampah pada tempatnya.

- (4) Pengorganisasian terjadi ketika siswa dihadapkan pada beberapa pilihan yang sama beratnya maka akan muncul sikap baru yang konsisten. Krathwohl menyatakan *organization involves relating the new value to those one already holds and bringing it into a harmonious and internally consistent philosophy*. Organisasi berkaitan dengan nilai baru yang sudah dipegang. Nilai yang membawanya ke sebuah filsafat yang harmonis dan konsisten secara internal. Pada kasus ini, jika siswa yang sudah memberi penilaian positif pada tugas membuang sampah pada tempatnya tiba-tiba ada pengaruh dari luar yang mengajaknya untuk membuang sampah sembarangan, maka dalam diri siswa tersebut muncul perang batin antara ya dan tidak sehingga perlu pengorganisasian sikap/perilaku. Tingkat pengorganisasian membutuhkan banyak pertimbangan positif dan negatif sebelum sikap/perilaku membuang sampah pada tempatnya menjadi pola/kebiasaan. Jika siswa lolos pada tingkat ini, maka tingkat terakhir dari perilaku afektif adalah internalisasi.
- (5) Level terakhir dari kompetensi afektif adalah karakterisasi atau internalisasi Krathwohl menyatakan *characterization by value or value set means acting consistently in accordance with the values the individual has internalized*. Karakterisasi adalah nilai sikap telah menjadi tindakan yang konsisten atau telah diinternalisasi menjadi nilai-nilai individu yang permanen. *Internalization refers to the process whereby a person's affect toward an object passes where the affect is 'internalized' and consistently guides or controls the person's behavior* (Seels & Glasgow, 1990). Internalisasi mengacu pada proses dimana seseorang bersikap terhadap suatu objekdimana sikap tersebut 'diinternalisasikan' dan secara konsisten memandu atau mengendalikan perilaku seseorang.

Internalisasi terjadi jika sikap/perilaku positif/negatif menang pada tingkat pengorganisasian. Perilaku yang menjadi kebiasaan akan melekat dan sulit diubah meskipun mendapat berbagai pengaruh dari luar. Guru harus sering menanamkan kebiasaan positif karena jika perilaku afektif negatif yang menang dalam pengorganisasian maka perilaku itupun akan sulit diubah. Contoh tindakan yang mencerminkan sikap positif sudah diinternalisasi antara lain: memperbaiki sikap, menuntut sikap positif, menghindari sikap negatif, melawan sikap negatif, dan menyelesaikan konflik nilai. Contoh internalisasi sikap dalam kehidupan sehari-hari yaitu jika perilaku membuang sampah

pada tempatnya sudah diinternalisasi menjadi kebiasaan maka mereka akan melawan pada orang-orang yang tidak mau membuang sampah pada tempatnya. Contoh pembentukan sikap/kebiasaan merokok/tidak merokok melalui Gambar 3.2.



Gambar 3.2 Pembentukan Kebiasaan Merokok

Keterangan gambar sebagai berikut

- Penerimaan (*receiving*) terjadi ketika siswa melihat ada orang yang sedang merokok, siswa tidak memberi respon apapun tentang orang yang sedang merokok.
- Pemberian respon (*responding*) terjadi ketika siswa memberi respon positif/negatif. Respon positif dapat dilihat jika siswa menanyakan keingintahuannya tentang rokok. Respon negatif

dapat dilihat jika siswa terus menghindar dari orang yang sedang merokok.

- c) Penghargaan (*valuing*) terjadi ketika siswa mendapat informasi positif dan negatif tentang rokok. Siswa memberi penilaian dalam dirinya untuk ikut atau tidak ikut merokok.
- d) Pengorganisasian (*organization*) terjadi ketika siswa sudah memiliki penilaian negatif terhadap rokok, maka ketika ada orang yang mengajaknya merokok dia mulai mengoordinasikan antara bujukan, keinginan, resiko dan nilai-nilai yang telah dimiliki sebelumnya. Jika siswa memberi penilaian positif terhadap rokok, maka dalam pengorganisasian tidak terlalu banyak pertimbangan.
- e) *Characterization* atau internalisasi (*internalizing*) dalam kondisi ini, perilaku merokok/tidak merokok telah menjadi kebiasaan. Jika dalam pengorganisasian, siswa memilih tidak merokok, maka selamanya dia tidak akan merokok meskipun banyak stimulus yang menggodanya untuk merokok.

Penilaian domain afektif dalam proses pembelajaran

Domain afektif merupakan domain yang paling sulit diukur dan dinilai dalam proses pembelajaran. Guru harus melakukan pengamatan secara individu untuk dapat memberikan nilai. Sejalan dengan kebijakan pendidikan karakter, penanaman sikap/ afektif positif perlu dilakukan oleh semua guru karena ranah afektif ini dapat mendukung hasil belajar siswa menjadi lebih baik.

a. Penerimaan (*Receiving Phenomena*)

Sikap (afektif) pertama yang ditunjukkan seorang siswa pada saat belajar adalah menerima stimulus dari guru. Apa saja yang dirasakan siswa dari panca inderanya (dilihat, didengar, dicium, diraba, dan dicecap) akan menimbulkan kesan yang melekat pada perasaannya. Kesan ini menjadi dasar untuk menentukan sikap berikutnya. Sikap-sikap “penerimaan” ini terlihat dari: *kesadaran, kesiediaan untuk mendengar, menerima, memberi perhatian terhadap berbagai stimulus*. Dalam tahap *receiving* ini, siswa masih bersikap pasif.

Contoh perilaku yang menunjukkan domain afektif level penerimaan yaitu :

- Siswa mendengarkan penjelasan guru.
- Siswa melihat tayangan media dengan saksama.
- Siswa memperhatikan proses pembuatan bakso.

b. Pemberian respon/menanggapi (*Responding*)

Sikap afektif berikutnya setelah siswa menerima stimulus belajar adalah memberi respon. Perilaku yang nampak pada siswa adalah: *berpartisipasi aktif pada proses pembelajaran dengan cara memberikan perhatian dan reaksi pada kejadian khusus selama proses pembelajaran.* Dampak yang diperoleh adalah kepatuhan dalam menanggapi, kemauan untuk merespon, menjawab pertanyaan atau bertanya, secara sukarela memberi kepuasan kepada pemberi stimulus.

Contoh perilaku yang menunjukkan domain afektif level *responding* adalah:

- Siswa menjawab pertanyaan guru.
- Siswa mengajukan pertanyaan.
- Siswa berpartisipasi pada kegiatan yang didiskusikan.
- Siswa mau mengerjakan tugas dengan sukarela.

c. Penghargaan (*valuing*)

Setelah menerima dan memberi respon terhadap stimulus belajar, siswa mulai memberi penilaian terhadap hal baru yang dipelajarinya. Penilaian bisa bersifat positif maupun negatif. Perilaku ini bisa tidak tampak sehingga perlu ditanyakan kepada siswa. Pada proses penilaian ini terjadi: *perasaan, keyakinan, atau anggapan terhadap objek tertentu, benda, fenomena, atau perilaku yang melekat itu memiliki nilai.* Siswa secara konsisten berperilaku sesuai dengan nilai meskipun tidak ada pihak lain yang mengharuskan.

Contoh pengukuran *level valuing*:

- Siswa menyetujui hasil diskusi kelompok.
- Siswa mengemukakan argumen kelebihan dan kekurangan jika tidak mematuhi tata tertib di laboratorium.
- Siswa memuji pendapat siswa lain yang lebih pintar.

- Siswa mengakui lebih baik mengenakan seragam.

d. Pengorganisasian (*organization*)

Pengorganisasian menunjukkan saling ketergantungan antara nilai-nilai tertentu dalam suatu sistem nilai sehingga mendorong siswa untuk memandang penting semua bagian untuk dilibatkan dalam proses menilai dan mengambil keputusan bertindak. Pengorganisasian membutuhkan banyak pertimbangan untuk menetapkan tindakan. Pengorganisasian dapat membantu:

- (1) Menentukan nilai-nilai yang menjadi prioritas untuk dipilih dari berbagai nilai yang berbeda;
- (2) Upaya menyelesaikan konflik batin;
- (3) Menciptakan sistem nilai yang unik;
- (4) Mengakui perlunya keseimbangan antara perilaku yang bebas dengan perilaku yang bertanggung jawab.

Contoh:

- Siswa memilih untuk kembali mengerjakan tugas setelah ada siswa lain yang mengajaknya istirahat/bermain.
- Siswa mempertahankan pendapatnya pada saat berdebat.
- Siswa selalu berusaha untuk menepati janji dan disiplin.

e. *Internalizing values*

Pengorganisasian dan pengintegrasian nilai ke dalam sistem nilai pribadi untuk mengendalikan perilaku. Perilaku merasuk ke dalam diri menjadi konsisten dan dapat diprediksi. Siswa mengintegrasikan nilai ke dalam filsafat hidup yang lengkap dan meyakinkan. Perilaku siswa akan selalu konsisten dengan filsafat hidup yang dipilih. Perilaku yang selalu konsisten tersebut akan membentuk karakter yang melekat permanen pada masing-masing individu.

Kriteria perilaku:

- Menunjukkan perilaku positif/negatif tanpa tekanan dari pihak lain.
- Bersikap objektif dalam memberi saran pemecahan masalah.
- Menunjukkan komitmen profesional dalam pekerjaan sehari-hari.

- Memperbaiki nilai dan mengubah perilaku karena ada kejadian-kejadian baru yang lebih baik.
- Perilakunya selalu konsisten, sudah melekat.

Contoh:

- Siswa selalu datang tepat waktu (sudah konsisten dengan perilaku disiplin).
- Siswa selalu bekerja dengan rapi.

3. *Psychomotor Domain*

The psychomotor domain (Simpson, 1996) adalah domain gerakan fisik, koordinasi yang menggunakan keterampilan gerak. Pengembangan keterampilan ini memerlukan latihan dan *diukur dalam hal kecepatan, ketepatan, jarak, prosedur, atau teknik dalam pelaksanaan*. Simpson mengklasifikasikan tujuh kategori perilaku dari yang paling sederhana sampai perilaku yang paling rumit yaitu: persepsi, kesiapan, respon terpinpin, mekanistik, respon yang kompleks, penyesuaian dan penciptaan. Dave’s dan Atkinson’s (Atkinson, 2015) membagi domain psikomotor menjadi lima sub domain yang lebih nyata untuk diukur. Taksonomi sub domain psikomotor dijelaskan pada Tabel 3.1.

Tabel 3.1 Tingkat Kompetensi Psikomotorik (Keterampilan)

Simpson’s	Dave’s original	Atkinson’s adaptation	Standar Proses Kurikulum 13
<i>Perception</i>			Mengamati
<i>Set/Readiness</i>			Menanya
<i>Guided response</i>	<i>Imitation</i>	<i>Imitate</i>	Mencoba
<i>Mechanism</i>	<i>Manipulation</i>	<i>Manipulate</i>	Menalar
<i>Complex</i>	<i>Precision</i>	<i>Perfect</i>	Menyaji
<i>Adaptation</i>	<i>Articulation</i>	<i>Articulate</i>	Mencipta
<i>Origination</i>	<i>Naturalisation</i>	<i>Embody</i>	

Taksonomi kompetensi psikomotor yang dikemukakan oleh tiga ahli yaitu Simpson’s, Dave’s dan Atkinson’s pada Tabel 3.1 terdapat 5 level kompetensi yang menunjukkan kesetaraan karakteristik. Deskripsi dua

sub domain psikomotor dari Simpson's adalah sebagai berikut: (1) Persepsi berkaitan dengan kemampuan untuk menggunakan isyarat-isyarat sensoris yang memandu aktivitas motorik. Hal ini melibatkan rangsangan panca indra misalnya penglihatan dan pendengaran. Ketika seseorang pertama kali melihat atau mendengar, syaraf otak kemudian menerima isyarat tersebut untuk diterjemahkan menjadi gerakan (merefleksi). Contoh: ketika seseorang melihat orang membuat minuman, maka di dalam otak timbul berbagai macam persepsi terhadap objek minuman tersebut seperti cara mencampur, cara mengaduk, cara menyajikan, dsb; (2) Kesiapan mencakup satu kesatuan mental, fisik, dan emosional untuk bertindak. Kesiapan menentukan pola berpikir orang dalam menggunakan cara tertentu pada saat melakukan tindakan. Kesiapan dapat ditunjukkan oleh kemampuan menyesuaikan diri terhadap hal-hal baru. Jika dalam proses belajar, siswa diminta membuat minuman seperti contoh, siswa yang telah siap minimal memiliki pengetahuan tentang minuman, kondisi fisiknya kuat untuk membuat minuman, dan emosional mendukung untuk tidak malas membuat minuman. Sub domain "persepsi dan kesiapan" dari Simpson sulit diukur karena belum ada keterampilan yang nampak untuk diobservasi.

Deskripsi lebih lanjut dari lima level kompetensi psikomotor berikutnya dan kata kerja yang digunakan untuk mengukur ketercapaian kompetensi dijelaskan pada paparan berikut ini.

a. Imitation (meniru)

Imitasi atau meniru yaitu menirukan pola perilaku yang telah diamati dari orang lain. Menurut Atkinson's (Atkinson, 2015) imitation adalah *ability to copy, replicate the actions of others following observations* (kemampuan untuk menyalin, meniru tindakan, mengikuti tindakan orang lain yang diamatinya). Kegiatan yang dilakukan guru dalam melatih keterampilan pada level pertama ini adalah mendemonstrasikan keterampilan agar dapat ditiru oleh siswa. Hasil belajar keterampilan level imitasi masih berkualitas rendah dan gerakan motoriknya kasar.

Contoh tujuan pembelajaran keterampilan level imitasi:

- Siswa dapat memegang pisau sesuai dengan contoh guru.
- Siswa dapat menirukan gerakan telapak tangan pada saat menguli adonan roti.

- Siswa dapat memotong wortel berbentuk jardiner sesuai contoh dari guru.

b. Manipulation (manipulasi):

Atkinson's (Atkinson, 2015) menjelaskan *manipulate* adalah *ability to repeat or reproduce actions to prescribed standard from memory or instructions* (kemampuan untuk mengulang atau mereproduksi tindakan sesuai standar yang ditentukan dari memori atau instruksi). Belajar keterampilan pada level kedua, siswa sudah mampu melakukan tindakan tertentu karena latihan berulang-ulang (mekanis) tanpa bantuan visual dari orang lain. Siswa menggunakan ingatannya pada saat belajar keterampilan sebelumnya atau membaca ulang petunjuk kerja (*job sheet*). Kata kerja operasional masih sama seperti pada tingkat imitasi, namun pada tingkat ini, keterampilan sudah mandiri tanpa bantuan contoh dari guru.

Contoh tujuan pembelajaran keterampilan level manipulasi:

- Siswa dapat memotong karkas ayam menjadi 16 bagian, setelah mempelajari *job sheet*.
- Siswa dapat memegang pisau dengan benar sesuai dengan contoh dari guru minggu yang lalu.
- Siswa dapat memotong sayur bentuk jardiner tanpa contoh dari guru.

c. Precision (presisi)

Atkinson memberi nama *perfect* pada kompetensi keterampilan level ke-3 yang setara dengan presisi. *Perfect* menunjukkan pekerjaan yang sempurna tanpa cacat. Atkinson's (2015) menjelaskan *ability to perform actions with expertise and without interventions and the ability to demonstrate and explain actions to others*. (kemampuan untuk melakukan tindakan dengan keahlian tanpa intervensi orang lain dan kemampuan untuk menunjukkan atau menjelaskan tindakan kepada orang lain). Keterampilan level ini ditunjukkan oleh kemampuan bekerja cepat dan tepat dengan sedikit kesalahan tanpa menggunakan petunjuk visual maupun tertulis. Keterangan yang menunjukkan tingkat presisi antara lain: "dengan tepat, dengan lancar, tanpa kesalahan".

Contoh tujuan pembelajaran keterampilan level presisi:

- Siswa dapat mencetak kue donat dengan ukuran dan bentuk yang seragam.
- Siswa dapat memotong wortel dengan bentuk juliene seragam.
- Mahasiswa dapat menata meja makan (*table set up*) dengan rapi dan tanpa kesalahan.

d. Articulation (artikulasi):

Setelah mencapai tingkat presisi, keterampilan (*skill*) dapat di tingkatkan lagi satu tahap yang lebih sulit yaitu tingkat artikulasi. Atkinson's menjelaskan artikulasi adalah "*ability to adapt existing psychomotor skills in a non-standard way, in different contexts, using alternative tools and instruments to satisfy need*". Istilah artikulasi mengandung maksud kemampuan untuk beradaptasi dalam mempelajari keterampilan psikomotorik baru dengan cara non-standar, dalam konteks yang berbeda, menggunakan alat dan instrumen alternatif untuk memenuhi kebutuhan.

Pada tingkat artikulasi, siswa dapat menunjukkan serangkaian gerakan yang akurat, sesuai prosedur, cepat dan tepat. Pada level artikulasi, terdapat keleluasaan untuk berkreasi menggunakan cara baru yang lebih efektif untuk mencapai tujuan. Di sisi lain, artikulasi dapat berarti juga serangkaian gerakan yang memerlukan koordinasi menjadi tindakan yang selaras untuk menunjukkan tingkat keterampilan yang lebih tinggi.

Contoh tujuan pembelajaran keterampilan level artikulasi:

- Mahasiswa dapat memotong kentang bentuk turning yang seragam (ada koordinasi gerak tangan menekan dan mendorong).
- Siswa dapat mengukir buah labu kuning berbentuk bunga matahari (penggabungan keterampilan dan kreativitas).
- Siswa dapat bermain *juggling* (keterampilan yang kompleks).
- Siswa dapat menghias kue tart dengan bentuk hiasan kerang dan pita yang menggantung.

e. Naturalization (naturalisasi)

Atkinson menganalogikan naturalisasi dengan *embody*. Beliau menjelaskan bahwa keterampilan tingkat paling mahir diberi nama *embody* yang berarti *ability to perform actions in an automatic, intuitive or unconscious way appropriate to the context*. Dalam definisi tadi, naturalisasi/*embody* dapat berarti kemampuan untuk melakukan tindakan dengan

cara otomatis, intuitif atau spontan. Siswa yang memiliki kecerdasan kinestetik (gerak tubuh) dapat cepat berlatih keterampilan, mereka memiliki performa tingkat tinggi secara alami, mempunyai bakat alam sehingga tanpa perlu bersusah payah belajar mereka cepat bisa. Kata keterangan yang sering digunakan untuk menyertai objek pengamatan adalah otomatis, dengan lancar, dengan sempurna, dan sebagainya.

Contoh tujuan pembelajaran keterampilan level naturalisasi adalah sebagai berikut.

- Siswa dapat mencetak 100 kue semprit dengan bentuk yang seragam dalam waktu 10 menit.
- Siswa dapat memotong 10 wortel berbentuk jardiniere sesuai standardalam waktu 10 menit.
- Siswa dapat membungkus lemper dengan seragam, rapi dan lancar.
- Siswa dapat memorsi kentang goreng dengan berat yang sama.

Contoh tahap perkembangan belajar keterampilan mem-*fillet* ikan diilustrasikan dengan tahap-tahap yang terdapat pada Gambar 3.4.



Gambar 3.3 Tahap Perkembangan Belajar Keterampilan

B. Tujuan Pembelajaran

Evaluasi dilakukan untuk mengetahui ketercapaian kompetensi tertentu yang telah dinyatakan pada tujuan pembelajaran. Tujuan dan evaluasi hasil belajar harus konsisten agar dapat memberi informasi apakah proses belajar mengajar telah mencapai tujuan yang diharapkan. Oleh sebab itu, evaluasi harus mengacu pada tujuan pembelajaran yang telah ditetapkan. Konsistensi antara tujuan dan evaluasi diilustrasikan pada Gambar 3.4.



Gambar 3.4 Konsistensi Tujuan dan Evaluasi dalam Pembelajaran

Dalam tujuan pembelajaran terdapat kompetensi atau kemampuan yang diharapkan dapat ditunjukkan oleh siswa pada akhir pembelajaran. Rumusan kompetensi ditulis dalam bentuk kata kerja operasional (KKO). Ciri-ciri KKO yaitu kata kerja yang dapat diubah menjadi “kata tanya atau perintah” Respon jawaban “kata tanya atau perintah” tersebut berupa perilaku yang dapat diukur atau diamati. Kata kerja kurang operasional (KKKO) jika diubah menjadi “kata tanya atau perintah”, respon jawaban tidak dapat diketahui orang lain secara langsung. Contoh kata kerja operasional:

Tabel 3.2 Contoh KK Operasional dan Kurang Operasional

KK	Kata tanya atau perintah	Keterangan
Menyebutkan	Sebutkan	Operasional
Menjelaskan	Jelaskan	Operasional
Membuat	Buatlah	Operasional
Membedakan	Bedakan	Operasional
Membedingkan	Bandingkan	Operasional
Memahami	Pahamilah?	Kurang operasional
Mengetahui	Ketahuilah?	Kurang operasional

Kata memahami dan mengetahui tidak langsung dapat diamati hasilnya, melainkan harus diukur dengan tes yang komprehensif sehingga tidak layak menjadi tujuan pembelajaran. Kata memahami dan mengetahui dapat digunakan untuk pernyataan kompetensi dasar, yang penilaiannya dilakukan dari akumulasi data tes pemahaman/pengetahuan. Kompetensi pada tujuan pembelajaran hanya dicapai setelah akhir pembelajaran. Contoh perumusan tujuan pembelajaran terdapat pada Tabel 3.3.

Tabel 3.3 Contoh Ranah Kompetensi Tujuan Pembelajaran

No	Tujuan Pembelajaran	Domain	Sub domain
1.	Pada akhir pembelajaran, siswa dapat menjelaskan prosedur pembuatan roti menggunakan <i>straight dough method</i>	<i>Cognitive</i>	<i>Comprehensif</i>
2.	Siswa dapat menjelaskan perbedaan tekstur kue bolu yang menggunakan adonan <i>creaming</i> dan <i>sponge</i>	<i>Cognitive</i>	<i>Analysis</i>
4.	Siswa bersedia mendiskusikan materi yang diajarkan guru	<i>Affective</i>	<i>Responding</i>
5.	Siswa dapat memotong wortel bentuk <i>julienne</i> sesuai contoh	Psikomotor	Imitasi
6.	Siswa dapat menata meja makan gaya Rusia	Psikomotor	Manipulasi

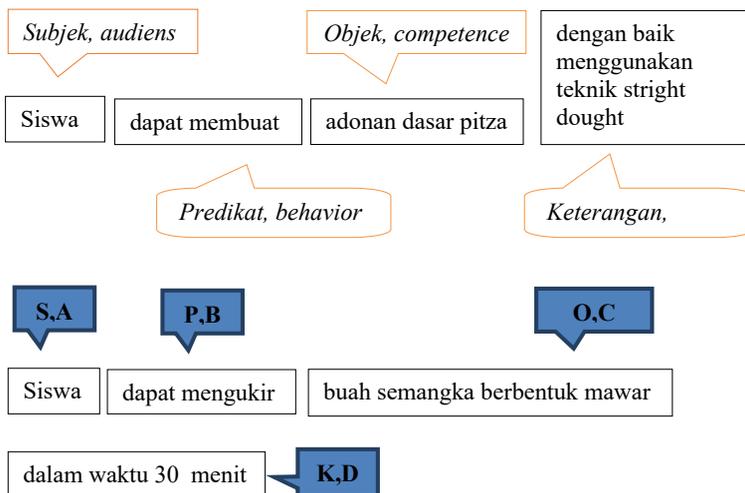
Untuk mengukur kompetensi sudah/belum tercapai maka ditetapkan indikator pencapaian kompetensi (IPK). Tujuan pembelajaran dan indikator pencapaian kompetensi telah dirumuskan pendidik (guru/dosen) pada saat menyusun rencana pelaksanaan pembelajaran (RPP). Indikator pencapaian kompetensi adalah kriteria yang menunjukkan bahwa kompetensi dasar tertentu telah tercapai. Kata kerja yang digunakan untuk merumuskan indikator pencapaian kompetensi adalah kata kerja operasional, artinya kata tersebut mencerminkan perilaku (pengetahuan, sikap, dan keterampilan) yang dapat diamati dan diukur dengan alat ukur yang jelas. Penulisan indikator menggunakan persyaratan kondisi dan ada kriteria yang menunjukkan siswa telah kompeten atau belum kompeten secara individual. Oleh sebab itu, indikator pencapaian kompetensi dalam sebuah RPP akan menjadi acuan pada saat guru menilai hasil belajar.

Penulisan tujuan pembelajaran mengikuti aturan penulisan kalimat baku yang terdiri atas unsur **SPOK** (subjek, predikat, objek, dan

keterangan). Dalam bahasa Inggris, tujuan pembelajaran disusun dalam kalimat yang mengandung unsur **ABCD** (*Audience*: Siswa, *Behavior*: Perilaku, *Competency*: Kompetensi dan *Degree*: tingkat/ukuran). Susunan kalimat pada tujuan pembelajaran diawali dengan menyebutkan pelaku (*audience*) yaitu “siswa” kemudian diikuti dengan menyebutkan perilaku dan “kompetensi” akhir yang diharapkan dapat dilakukan siswa dengan menggunakan kata kerja operasional. Contoh perumusan tujuan menggunakan sistematika SPOK/ABCD:

- (1) Subjek/*audience* dengan menyebutkan kata “siswa”,
- (2) Predikat berupa kata kerja operasional (KKO) atau *behavior* yang dapat diukur dan diamati. Contoh: membuat, memotong, menjelaskan, menganalisis, dsb .
- (3) Objek berupa inti materi yang mengiringi KKO atau kompetensi hasil belajar yang diharapkan dicapai setelah pembelajaran, Contoh: Predikat atau *behavior* dengan kata “membuat” maka objek atau kompetensi dapat berupa materi yang dibuat seperti: soto, sate, napkin, sauce, salad, mayonnaise, dsb.
- (4) Keterangan, *degree* diperlukan jika ada persyaratan kondisi atau kriteria khusus dari perilaku dikehendaki seperti waktu penyelesaian, tempat pengerjaan, jumlah yang harus diselesaikan, alat yang digunakan, dll.

Contoh perumusan tujuan pembelajaran yang menggunakan rumus SPOK/ABCD:



Prinsip yang digunakan dalam perumusan tujuan adalah **SMART** yaitu: *Specific, Measurable, Achievable/Attainable, Realistic, Timely*. Prinsip ini tidak hanya digunakan untuk merumuskan tujuan pembelajaran saja, namun dalam perumusan tujuan lain seperti tujuan program, pendirian lembaga, visi misi juga menggunakan prinsip yang sama. Contoh tujuan pembelajaran: “*Siswa dapat membuat 3 porsi soto Bandung dengan kuah yang jernih dan rasa gurih dalam waktu 120 menit*”. Indikator SMART pada tujuan pembelajaran tersebut tercermin dari: materi *spesifik* yaitu hanya membuat satu macam soto meskipun banyak jenis soto yang ada di Indonesia. *Measurable*, yaitu kemampuan mahasiswa setelah proses pembelajaran dapat diukur (benar/salah dan baik/kurang) dan ada alat ukur yang jelas yaitu membuat 3 porsi soto Bandung dengan kriteria berkuah jernih dan rasa gurih. *Achievable*, yaitu kompetensi yang diharapkan pada tujuan pembelajaran dapat dicapai pada akhir pembelajaran. *Rational*, yaitu kompetensi yang diharapkan pada tujuan pembelajaran rasional untuk dicapai, terdapat daya dukung proses pembelajaran seperti alat dan bahan tersedia, harga terjangkau, peserta didik mampu membuatnya. *Timely*, yaitu ada batasan waktu untuk mencapai tujuan pembelajaran. Dalam contoh tujuan pembelajaran di atas sudah disebutkan 120 menit, sehingga cukup dengan satu kali tatap muka.



Gambar 3.5 Indikator Perumusan Tujuan

Contoh soal evaluasi hasil belajar, kunci jawaban dan rubrik penilaian hasil belajar yang menggunakan indikator SMART yaitu:

Tujuan pembelajaran

- 1) Teori: Setelah selesai pelajaran, siswa dapat menganalisis 5 penyebab kerusakan mikrobiologis, kimia, mekanis, biologis, dan fisik pada bahan pangan.
- 2) Praktik: Siswa dapat membuat 15 potongan wortel berbentuk jardinière dengan ukuran yang tepat selama waktu 10 menit.

Soal evaluasi hasil belajar

- 1) Berilah lima contoh kerusakan bahan pangan yang disebabkan oleh faktor mikrobiologis, kimia, mekanis, biologis, dan fisik dan analisislah penyebabnya!
- 2) Buatlah 15 potongan wortel berbentuk jardinière dengan tepat dalam waktu 10 menit!

Jawaban soal:

Penyebab kerusakan bahan makanan dapat dikelompokkan menjadi beberapa jenis, yaitu:

- 1) **Kerusakan mikrobiologis.** Kerusakan ini disebabkan oleh mikroorganisme seperti bakteri, kapang, dan khamir. Jenis kerusakan ini ditandai dengan timbulnya kapang, kebusukan, lendir, dan adanya perubahan warna. Kerusakan mikrobiologis dapat berbahaya bagi kesehatan manusia karena racun yang diproduksi oleh mikroorganisme.
- 2) **Kerusakan biologis.** Kerusakan biologis disebabkan oleh kerusakan fisiologis, serangga, maupun binatang pengerat. Kerusakan ini meliputi reaksi metabolisme pada bahan atau enzim-enzim yang terdapat didalamnya sehingga terjadi proses autolysis yang menyebabkan terjadinya kerusakan.
- 3) **Kerusakan Fisik.** Kerusakan ini disebabkan oleh perlakuan-perlakuan fisik seperti pemanasan, pendinginan, dan tekanan udara. Contoh dari kerusakan fisik adalah Kerusakan warna dan tekstur pada daging yang dibekukan, tepung mengeras atau membatu karena penyimpanan pada tempat yang lembab dsb.
- 4) **Kerusakan mekanis** adalah kerusakan yang disebabkan karena bahan mengalami benturan-benturan mekanis yang terjadi selama pemanenan, transportasi ataupun penyimpanan. Contohnya: Pada waktu dipanen buah yang jatuh atau membentur permukaan keras menjadi memar.
- 5) **Kerusakan kimiawi** adalah kerusakan yang terjadi karena reaksi kimia yang berlangsung di dalam bahan makanan seperti penurunan pH, proses rigor, dan reaksi reduksi dan oksidasi. Contoh dari kerusakan kimiawi misalnya Reaksi pencoklatan pada beberapa jenis buah dan sayur, reaksi ketengikan minyak, dsb. (Muchtadi, 2011)

Rubrik Penilaian Praktik Potongan Sayur

Untuk menilai potongan sayur diperlukan rubrik penilaian. Setiap komponen yang dinilai memiliki bobot berbeda tergantung pada tingkat kepentingan komponen tersebut dalam membentuk kompetensi. Contoh rubrik penilaian terdapat pada Tabel 3.3.

Tabel 3.3 Contoh Rubrik Penilaian Kinerja Praktik Harian

No	Komponen/Sub komponen Penilaian	Skor maksimal
I	Persiapan Kerja	15
	Bahan	
	Alat	
	Penampilan/ <i>performance</i>	
II	Proses (Sistematika & Cara Kerja)	30
	Cara memegang pisau	
	Cara mengupas bahan	
	Cara memotong bahan	
	Efisiensi bahan	
III	Hasil Kerja	30
	Ketepatan ukuran	
	Kerapihan bentuk	
IV	Sikap Kerja	15
	K3	
	Tanggung jawab	
	Disiplin	
V	Waktu	10
	Waktu penyelesaian semua tahapan praktik	
	Total skor	100

Prinsip-prinsip penyusunan tujuan pembelajaran yang menggunakan indikator SMART dapat dijelaskan sebagai berikut

a. *Specific*

Tujuan pembelajaran hanya mengukur satu jenis kemampuan khusus (spesifik), tidak boleh mengukur beberapa kemampuan sekaligus atau kemampuan yang terlalu umum, misalnya:

Boleh	: • Siswa dapat menyebutkan empat macam nasi berbumbu. • Siswa dapat menjelaskan prosedur pembuatan nasi kuning. • Siswa dapat menghias tumpeng nasi kuning selama 25 menit.
Tidak boleh	: • Siswa dapat menjelaskan prosedur pembuatan nasi udak dan nasi kuning (dua objek yang berbeda). • Siswa memiliki kemampuan mengolah nasi berbumbu (kurang spesifik karena nasi berbumbu jumlahnya banyak).

b. *Measurable*

Kompetensi yang tertulis dalam tujuan pembelajaran dapat *diukur* dan memiliki kriteria yang jelas/pasti untuk mengukur keberhasilannya. Kompetensi diukur dari tiga ranah yaitu pengetahuan, sikap, dan keterampilan. Hasil belajar ranah pengetahuan diukur dengan tes. Hasil belajar ranah sikap dan keterampilan diukur dengan observasi. Alat ukur hasil belajar memiliki kualitas yaitu memenuhi unsur validitas dan reliabilitas.

Contoh:

Boleh	: • Siswa dapat menjelaskan 2 fungsi protein dalam pertumbuhan manusia. • Siswa dapat membuat 1 buah garnish tomat berbentuk bunga mawar dalam waktu 5 menit.
Tidak boleh	: • Siswa memahami sejauh mana protein berfungsi untuk pertumbuhan manusia (kata memahami tidak menunjukkan perilaku yang dapat diobservasi atau diubah menjadi kata tanya pada evaluasi sehingga sebaiknya diganti dengan kata menjelaskan). • Siswa dapat membuat garnish dari buah tomat (tidak disebutkan jenis <i>garnish</i> , batasan jumlah dan waktunya).

c. *Achievable/attainable*

Kompetensi yang tertulis dalam tujuan pembelajaran harus dapat dicapai dan terdapat cara untuk mencapainya.

-
- | | |
|--------------------|--|
| Boleh | : • Siswa dapat merangkum materi kuliah bahan pangan dalam bentuk <i>mind mapping</i> dalam waktu 5 menit.
• Siswa dapat mencetak 100 kue semprit dengan bentuk yang seragam dalam waktu 1 jam. |
| <hr/> | |
| Tidak boleh | : • Siswa dapat merangkum materi kuliah bahan pangan menjadi 100 halaman (tidak dapat dicapai pada akhir pembelajaran karena jumlah rangkuman masih terlalu banyak).
• Siswa dapat membentuk 100 kue nastar dalam waktu 30 menit (sulit dicapai karena kue dibentuk dengan tangan). |
-

d. *Realistic*

Kompetensi yang tertulis dalam tujuan pembelajaran berwujud perilaku nyata (*tangible*) yang realistis untuk dicapai, dapat diukur atau diobservasi.

-
- | | |
|--------------------|---|
| Boleh | : • Siswa mampu memasak 2 porsi nasi goreng seafood untuk menu makan pagi keluarga.
• Siswa dapat mengidentifikasi 5 nama minuman panas dari daerah Jawa Tengah. |
| <hr/> | |
| Tidak boleh | : • Siswa mampu mengolah dan menghidangkan sukiyaki sesuai dengan standar masakan Jepang (sulit dicapai jika alat, bahan dan bumbu tidak tersedia).
• Siswa dapat mengidentifikasi 115 nama minuman panas dari daerah Jawa Tengah (jumlah terlalu banyak sulit dicapai, 10 saja sudah cukup karena nama minuman panas asli daerah Jawa Tengah tidak sebanyak itu). |
-

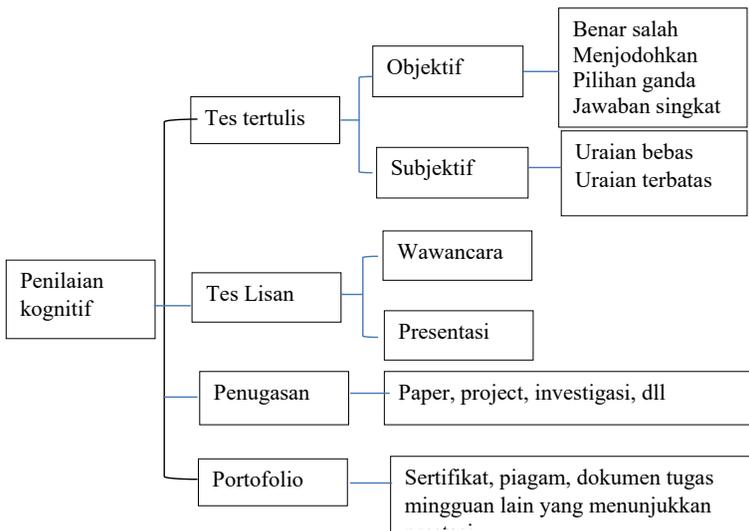
e. *Timely*

Waktu penyelesaian tugas atau soal-soal selama ujian menjadi penentu kompetensi seseorang. Orang yang kompeten dapat menyelesaikan tugas dengan benar dalam waktu singkat dan jika waktu penyelesaian lebih lama berarti belum kompeten atau masih dalam taraf latihan. Jumlah waktu yang diminta ditegaskan dalam tujuan pembelajaran dan pengukuran kompetensi. Perhitungan jumlah waktu penyelesaian tugas cukup memadai dan jelas batasnya

Boleh	: • Siswa dapat mencetak 10 butir kue semprit dalam waktu satu menit. • Siswa dapat mengukir buah semangka berbentuk bunga dalam waktu 30 menit.
Tidak boleh	: • Siswa dapat membuat kue donat dalam waktu 30 menit (tidak rasional sebab mengembangkan adonannya saja sudah membutuhkan waktu minimal 30 menit).

C. Pengukuran Hasil Belajar

Alat pengukuran hasil belajar di bedakan menjadi dua yaitu berupa tes dan non tes. Tes adalah seperangkat pertanyaan yang harus dijawab/ direspon, atau tugas yang harus dikerjakan oleh orang yang dites (diuji) dengan tujuan untuk mengukur kemampuan atau mengungkap aspek tertentu dari orang yang diuji tersebut. Setiap butir pertanyaan atau tugas mempunyai kriteria jawaban atau ketentuan yang dianggap benar atau salah. Alat pengukur non tes tidak menuntut seseorang yang diukur memberi respon pada pertanyaan/tugas secara langsung. Bentuk alat ukur non test lebih bervariasi dan tidak selalu menggunakan kriteria jawaban/respon benar atau salah tetapi lebih banyak menggunakan pertimbangan baik dan kurang baik. Alat pengukuran non tes ini dapat berupa lembar observasi, dokumen portofolio, penugasan, dll.



Gambar 3.6 Skema Alat Pengukuran Kemampuan Kognitif

Penentuan kelulusan siswa dengan menggunakan satu alat dan satu kali pengukuran melalui tes saja dianggap tidak adil terutama bagi siswa yang tidak dapat mengikuti tes secara maksimal. Penilaian tunggal tidak cukup untuk memberikan gambaran/informasi tentang kemampuan, keterampilan, pengetahuan, dan sikap seseorang. Penggunaan pengukuran ganda (*multiple measurement*) dapat mengumpulkan informasi hasil dan kemajuan belajar peserta didik secara lengkap. Menurut Montero (Montero, Yulian, & Yen, 2003) pendekatan pengukuran ganda berusaha untuk meningkatkan reliabilitas dengan cara menurunkan sejumlah kesalahan yang berhubungan dengan evaluasi pada siswa, guru, sekolah dan *local educational agency* (LEAs).

1. Alat Ukur Tes

Alat ukur tes banyak digunakan untuk mengukur kemampuan kognitif dan keterampilan siswa. Dalam bidang evaluasi dikenal tiga macam tes hasil belajar yaitu tes tertulis, tes lisan, dan tes kinerja. Tes tertulis (*paper and pencil test*) dan lisan digunakan untuk mengukur kemampuan kognitif. Tes kinerja digunakan untuk mengukur keterampilan atau kompetensi keahlian dalam bidang pekerjaan tertentu. Seiring dengan berkembangnya teknologi informasi, maka tes kognitif sudah banyak dilakukan secara online menggunakan media komputer atau *smartphone*. Penjelasan masing-masing tes tersebut adalah sebagai berikut.

Tes Tertulis

Tes tertulis terdiri atas beberapa macam bentuk. Menurut karakteristik jawaban terdapat tes yang jawabannya sudah tersedia dan tes yang jawabannya ditulis sendiri oleh peserta tes. Tes yang jawabannya sudah tersedia dikategorikan pada tipe *objective test* karena peserta tes tinggal memilih jawaban dan langsung mendapat skor benar atau salah, tanpa ada pertimbangan lain dari penilai. Tes tertulis yang jawabannya ditulis sendiri oleh peserta tes dikenal dengan *subjective test* karena pada saat penskoran atau pemberian nilai, penguji sangat dipengaruhi oleh suasana hati dan pertimbangan lain yang bersifat subjektif.

Tes Objektif

Tes objektif terdiri atas tes benar-salah (*true-false items*), menjodohkan atau mencocokkan (*matching exercises*), pilihan ganda atau MC (*multiple choice*), dan isian singkat (*short answer, fill-in items*).

Soal tes benar salah

Soal tes objektif benar-salah dapat dikembangkan menjadi soal benar salah dengan koreksi yaitu ketika peserta tes memilih jawaban benar/salah mereka dituntut untuk memberi perbaikan dan alasan terhadap pilihan jawabannya tersebut.

Tabel 3.4 Contoh Soal Benar Salah dengan Koreksi

No	Jawaban	Soal
1.	(B) – (S)	Perlakuan yang kasar seperti melempar, menjatuhkan, dan memencet buah mangga dapat menyebabkan kerusakan mekanis.
	Alasan	Kerusakan disebabkan karena bahan mengalami benturan-benturan mekanis yang terjadi selama pemanenan, transportasi ataupun penyimpanan.
2.	(B) – (S)	Buah apel yang telah dikupas/dipotong mudah mengalami pencoklatan non enzimatis.
	Alasan	Reaksi pencoklatan pada buah dan sayur merupakan reaksi enzimatis yang terjadi karena proses fisiologis antara oksigen dan enzim fenolase (Cheng & Crisosto, 1995).

Soal tes menjodohkan

Soal tes objektif menjodohkan atau mencocokkan (*matching exercises*) terdiri atas dua kolom yang berisi kolom pertanyaan dan jawaban, gambar dan nama gambar atau sebab dan akibat. Pada umumnya, soal menjodohkan hanya mengukur ranah kognitif pada level pemahaman (C2)

Tabel 3.5 Contoh Soal Menjodohkan

PERTANYAAN	JAWABAN
<p>Sauce yang paling tepat untuk hidangan berikut ini:</p> <p>1. Chicken steak (B)</p> <p>2. Beef steak (A)</p> <p>3. Fruit salad (D)</p>	<p>A. Demi-glace sauce</p> <p>B. Veloute sauce</p> <p>C. Hollandaise sauce</p> <p>D. Mayonaise</p> <p>E. Bechamel</p>

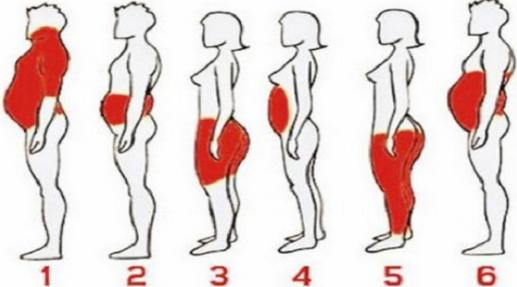
<p>BAHAN DASAR SAUS/SEBAB</p> <ol style="list-style-type: none"> 1. Milk (E) 2. White stock (B) 3. Brown stock (A) 4. Butter (C) 5. Salad oil (D) 	<p>SAUS DASAR/AKIBAT</p> <ol style="list-style-type: none"> A. Demi-glace sauce B. Veloute sauce C. Hollandaise sauce D. Mayonaise E. Bechamel
<p>GAMBAR</p> <p>1. Jawaban (B)</p> 	<p>NAMA GAMBAR</p> <p>2. Jawaban (C)</p> 
<p>3. Jawaban (A)</p> 	<p>Jawaban:</p> <ol style="list-style-type: none"> A. Demi-glace sauce B. Veloute sauce C. Hollandaise sauce D. Mayonaise E. Bechamel

Soal tes pilihan ganda:

Soal tes objektif pilihan ganda dapat dibuat menjadi hubungan antar beberapa hal, analisis kasus kompleks, dan membaca diagram. Berikut ini contoh soal pilihan ganda yang bervariasi bentuknya

Tabel 3.6 Contoh Soal Pilihan Ganda

<p>1</p>	<p>C3</p>	<p>Angka Kecukupan Energi wanita usia 19 -39, dengan BB ideal 60 dan TB 168 adalah 2725 kkal (Permenkes RI, no 75, thn 2013). Kebutuhan protein yang dianjurkan sebanyak 15% dari kebutuhan energi, 1 gram protein mengandung 4 kkal. Berapa keputusan protein wanita tersebut setiap hari?</p>
		<ol style="list-style-type: none"> A. 100, 2 gram protein B. 102,2 gram protein C. 104,2 gram protein D. 106,2 gram protein E. 108,2 gram protein

2	C4	<p>Konsumsi lemak jenuh pada penderita hipertensi dibatasi yaitu maksimal 10%energi total/hari. Dampak konsumsi lemak jenuh berlebihan pada penderita hipertensi adalah:</p> <p>(1) Meningkatkan serum kolesterol darah. (2) Menyebabkan tubuh gemuk dan perut buncit (3) Membentuk plak atheroma dalam pembuluh darah (4) pembuluh darah ke jantung mengalami penyempitan (5) mengganggu pernapasan dan sistem metabolisme tubuh</p>
		<p>Jawaban:</p> <p>A. (1), (2) dan (3) benar B. (1), (2) dan (4) benar C. (1), (3) dan (4) benar D. (1), (3) dan (5) benar E. (3), (4) dan (5) benar</p>
3	C4	 <p>The image shows six human figures from the waist up, illustrating different types of obesity. Each figure is labeled with a number from 1 to 6. Figure 1 shows a large belly. Figure 2 shows a large belly and large thighs. Figure 3 shows a large belly and large thighs. Figure 4 shows a large belly and large thighs. Figure 5 shows a large belly and large thighs. Figure 6 shows a large belly and large thighs.</p>
		<p>Pada gambar nomor berapa tipe obesitas yang disebabkan karena makan berlebihan dan kurang aktivitas?</p> <p>A. (1) Makan berlebihan dan (2) kurang aktivitas B. (2) Makan berlebihan dan (5) kurang aktivitas C. (4) Makan berlebihan dan (3) kurang aktivitas D. (6) Makan berlebihan dan (1) kurang aktivitas E. (1) Makan berlebihan dan (6) kurang aktivitas</p>

Constructed-response

Tes objektif yang jawabannya diisi sendiri oleh peserta tes (*constructed-response*: CR) dapat berbentuk jawaban singkat, melengkapi kalimat, menghitung dan lain-lain. Jawaban CR hanya satu kata atau padanan kata. Tes model ini dapat diselenggarakan dalam suatu permainan tebak kata, teka-teki silang atau *wordsquare*. Berikut ini ada sebuah contoh pertanyaan dan jawaban CR dalam bentuk *wordsquare*. Peserta tes mencari jawaban dengan cara memberi blok warna, jika dalam kotak masih kosong, peserta tes dapat mengisi kotak seperti teka-teki silang.

Soal:

1. Mineral yang diperlukan untuk pembentukan sel darah merah adalah
2. Penurunan kadar haemoglobin darah dapat menyebabkan penyakit
3. Kekurangan vitamin C dapat menyebabkan penyakit ... dengan gejala pendarahan di sekeliling gusi.
4. Vitamin A terdapat dalam 2 bentuk yaitu *preformed* vitamin A dan provitamin A atau sering disebut
5. Vitamin E dapat menjadi zat yang berfungsi melindungi asam lemak tak jenuh agar tidak mengalami kerusakan membrane fosfolipid.
6. Mineral ... berfungsi dalam pembentukan tulang.

Pilihan jawaban:

X	K	V	B	N	M	A	S	D	Z	X	A	V
C	A	N	T	I	O	K	S	I	D	A	N	T
I	L	B	E	T	A	A	R	A	G	I	T	O
O	S	V	E	D	R	R	U	I	O	P	I	K
D	I	E	N	S	K	O	R	B	U	T	G	O
I	U	H	A	K	I	T	O	I	U	Y	I	F
U	M	B	E	S	I	L	E	M	A	Z	E	
M	Q	H	E	R	A	N	E	M	I	A	I	R
A	S	A	M	F	O	L	A	L	A	T	S	O
F	O	S	F	O	R	I	S	A	S	I	V	L

Tes Subjektif

Tes subjektif banyak ditemukan dalam bentuk tes uraian yaitu tes yang memerlukan jawaban terurai. Pokok pertanyaan tes uraian diberikan dalam bentuk permasalahan kompleks sehingga menuntut jawaban tertulis berupa paparan, ungkapan pemikiran atau solusi terhadap permasalahan tersebut. Jawaban soal ditulis dengan menggunakan kalimatnya sendiri, tidak mencontoh atau sama persis dengan kalimat yang tertulis di buku atau literatur. Pada saat penskoran atau pemberian nilai, korektor sangat

dipengaruhi oleh suasana hati dan pertimbangan lain yang bersifat subjektif sehingga tes uraian dinamakan *tes subjektif*.

Tes uraian sangat tepat untuk mengukur hasil belajar/kemampuan siswa dalam hal:

- 1) Mengaplikasikan prinsip-prinsip dan teori pada mata pelajaran.
- 2) Membandingkan beberapa fenomena, menganalisis dan mengevaluasi hubungan antar sebab akibat (C4) antara fenomena/kejadian.
- 3) Mengaitkan relevansi informasi dengan mata pelajaran yang lebih kontekstual.
- 4) Merangkum, mengklasifikasikan, merumuskan hipotesis, dan mengambil keputusan, dan menarik kesimpulan.
- 5) Mengidentifikasi asumsi yang mendasari suatu kesimpulan.
- 6) Mengajukan alternatif pemecahan masalah, memberi ilustrasi atau contoh pemecahan masalah (C3).
- 7) Merancang prosedur eksperimen, pelaksanaan tugas belajar berbasis proyek atau kegiatan lain.

Ada beberapa bentuk tes uraian yaitu: (1) **Tes uraian terbatas:** tes jawaban singkat, tes melengkapi dan tes uraian terbatas sederhana; (2) **Tes uraian bebas:** tes uraian bebas sederhana dan tes uraian ekspresif. Untuk menjaga objektivitas pada saat koreksi hasil tes uraian, guru dituntut membuat kunci jawaban dan cara penskorannya. Pada tes uraian terbatas, kunci jawaban lebih mudah dibuat karena indikator merujuk pada materi pelajaran. Pada tes uraian bebas, penilaian hasil tes dibantu menggunakan rubrik penilaian yang memberi bobot lebih tinggi pada jawaban yang membutuhkan kemampuan berpikir tingkat tinggi.

Tes uraian sangat tepat untuk mengukur kemampuan metakognitif atau kemampuan berpikir tingkat tinggi (*higher order thinking skills*). Metakognitif merupakan keterampilan seseorang dalam mengatur dan mengontrol proses berpikirnya (Moshman & Schraw., 1995). Menurut Anderson pemerolehan pengetahuan memiliki empat tingkatan yaitu: *factual knowlwdge*, *conceptual knowledge*, *procedural knowledge* and *metacognitive knowledge*.

- 1) Pengetahuan faktual (*factual knowledge*) merupakan elemen dasar yang harus diketahui untuk memecahkan masalah atau mempelajari

disiplin ilmu tertentu. Dimensi ini mengacu pada fakta, terminologi, rincian, atau elemen penting yang harus diketahui atau dipahami siswa untuk memecahkan masalah di dalamnya. Jika ingin sukses dalam belajar membuat roti, maka siswa harus mengumpulkan fakta atau data resep-resep roti yang berisi komposisi, spesifikasi bahan yang digunakan dan teknik pembuatan;

- 2) Pengetahuan konseptual (*conceptual knowledge*) mengaitkan beberapa elemen dasar untuk membangun struktur yang lebih besar yang memungkinkan mereka berfungsi bersama. Pengetahuan konseptual dapat berisi tentang klasifikasi, prinsip, generalisasi, teori, model, atau struktur yang berkaitan dengan area disiplin tertentu. Jika ingin sukses belajar membuat roti, maka siswa harus memiliki banyak wawasan konseptual tentang prinsip-prinsip pembuatan roti untuk mengurangi resiko kegagalan.
- 3) Pengetahuan prosedural (*procedural knowledge*) berkaitan dengan bagaimana melakukan sesuatu, proses percobaan, kriteria keterampilan, algoritma, teknik, dan metode yang tepat. Jika ingin sukses belajar membuat roti, maka siswa harus prosedur pembuatan roti, mulai dari pembuatan adonan, fermentasi sampai ke pengemasan. Selama proses pembuatan, siswa mengamati perubahan yang terjadi untuk mengonstruksi pengetahuan berdasarkan pengalaman.
- 4) Pengetahuan metakognitif (*metacognitive knowledge*) –Kemampuan merespon situasi atau permasalahan menggunakan pengetahuan yang telah dimilikinya. Pengetahuan metakognitif menuntut orang memiliki kesadaran diri untuk mengendalikan atau mengatur kemampuan berpikirnya sendiri. Pengetahuan metakognitif meliputi pengetahuan strategis atau reflektif untuk mengatasi masalah, tugas-tugas berpikir kontekstual dan kondisional dengan menggunakan pengetahuan sendiri (Anderson, Krathwohl, & et all, 2001).

The Australian Council for Educational Research(ACER) menyatakan bahwa kemampuan berpikir tingkat tinggi (*higher order thinking skills*) merupakan proses: menganalisis, merefleksi, memberikan argumen/ alasan), menerapkan konsep pada situasi berbeda, dan kemampuan menyusun atau mencipta. Dalam taksonomi Bloom's, kemampuan berpikir tingkat tinggi dimulai dari analisis (C4), evaluasi (C5) dan kreasi (C6). Kemampuan berpikir tingkat tinggi bukan kemampuan untuk mengingat, mengetahui, atau mengulang. Kemampuan berpikir tingkat

tinggi termasuk kemampuan untuk memecahkan masalah (*problem solving*), keterampilan berpikir kritis (*critical thinking*), berpikir kreatif (*creative thinking*), kemampuan penalaran (*reasoning*), dan kemampuan mengambil keputusan (*decision making*) (Widana, 2017). Indikator yang digunakan untuk mengukur kemampuan berpikir tingkat tinggi adalah sebagai berikut:

- 1) Keterampilan pemecahan masalah, yakni suatu keterampilan seorang siswa untuk memecahkan masalah melalui pengumpulan fakta, analisis informasi, menyusun berbagai alternatif pemecahan, dan memilih pemecahan masalah yang paling efektif.
- 2) Keterampilan berpikir kritis, yakni keterampilan seseorang untuk menganalisis argumen dan memberikan interpretasi berdasarkan persepsi yang sah melalui interpretasi logis, analisis asumsi, dan argumen.
- 3) Keterampilan berpikir kreatif, yakni keterampilan seseorang untuk menghasilkan suatu ide baru, konstruktif, dan baik berdasarkan konsep-konsep, prinsip-prinsip yang rasional, maupun persepsi dan intuisi.
- 4) Kemampuan penalaran (*reasoning*) adalah proses berpikir koheren dan logis dalam penyelesaian suatu masalah atau mengambil kesimpulan.
- 5) Keterampilan pengambilan keputusan, yakni keterampilan seseorang untuk menentukan pilihan melalui pengumpulan informasi, perbandingan kebaikan dan kekurangan dari setiap alternatif sehingga memperoleh keputusan terbaik berdasarkan alasan rasional (Iskandar, 2014).

Soal tes uraian diharapkan mengukur *higher order thinking skills* yang menuntut siswa menggunakan pengetahuan metakognitifnya untuk menjawab pertanyaan. Namun demikian, bukan berarti pengetahuan faktual dan konseptual tidak penting karena kedua pengetahuan tersebut merupakan landasan untuk memperoleh pengetahuan metakognitif. Dalam penyusunan soal esai, tingkat kesulitan soal dibuat bervariasi untuk memberi kesempatan yang adil kepada semua siswa khususnya siswa yang *slow learner* agar mencapai nilai KKM. Dalam proses pembuatan soal esai, pengendalian dilakukan pada aspek konstruksinya agar butir soal uraian memiliki kualitas yang tinggi. Contoh soal-soal uraian beserta kunci jawaban dan rubrik penilaian terdapat pada lampiran 1. Berikut ini

terdapat kartu telaah soal uraian yang disusun oleh Pusat Pengembangan dan Sistem Pengujian (Umar, 1996).

Tabel 3.7 Telaah Soal Uraian

NO	KRITERIA PENILAIAN
A	MATERI
1	Soal sesuai dengan indikator
2	Batasan pertanyaan dan jawaban yang diharapkan jelas
3	Isi materi sesuai dengan tujuan pengukuran
4	Isi materi yang ditanyakan sudah sesuai dengan jenjang, jenis sekolah atau tingkatan kelas
B	KONSTRUKSI
5	Rumusan kalimat soal atau pertanyaan menggunakan kata tanya atau perintah yang menuntut jawaban terurai
6	Ada petunjuk yang jelas tentang cara pengerjaan soal
7	Ada pedoman penyekoran
8	Grafik, gambar, tabel, peta, diagram dan yang sejenisnya disajikan dengan jelas dan terbaca
C	BAHASA
9	Perumusan kalimat soal komunikatif
10	Butir soal menggunakan bahasa Indonesia yang baik dan benar
11	Rumusan soal tidak menggunakan kata-kata/kalimat yang menimbulkan penafsiran ganda atau salah pengertian
12	Tidak menggunakan bahasa yang berlaku setempat
13	Rumusan soal tidak mengandung kata-kata yang menyinggung perasaan siswa

Penjelasan:

- 1) Soal sesuai dengan indikator

Indikator soal dikembangkan dari materi yang telah diajarkan. Ruang lingkup materi sesuai dengan kurikulum, program semester, dan rencana pelaksanaan pembelajaran (RPP).

Kompetensi Dasar: Menganalisis bahan makanan dari daging dan hasil olahannya.

Baik : **Bagaimana karakteristik fisiologis daging sapi yang masih segar?**

Kurang baik : Bagaimana cara merawat sapi agar cepat gemuk dan menghasilkan daging yang empuk?

Keterangan: Soal di luar materi yang dipelajari dan tidak sesuai dengan mata pelajaran *Pengetahuan Bahan Pangan*.

- 2) Batasan pertanyaan dan jawaban yang diharapkan jelas.

Untuk mengukur kejelasan ini, maka pembuat pertanyaan harus mencoba menjawab pertanyaan yang diajukan, membuat kunci jawaban dan cara penyekoran jawaban (skor penilaian).

Kompetensi Dasar: Mengevaluasi perubahan sifat sayuran dan buah-buahan akibat perlakuan saat penyiapan bahan dan proses pengolahan.

Baik : **Mengapa buah apel setelah dikupas akan berwarna kecoklatan? Bagaimana cara menjaga agar buah apel tidak berwarna coklat setelah dikupas dan dipotong-potong?**

Kurang baik : Bagaimana cara menyimpan bahan makanan dari sayuran dan buah-buahan agar tidak mengalami kerusakan?

Keterangan: Soal menuntut jawaban yang terlalu luas karena tidak ada batasan materi sayur dan buah-buahan.

- 3) Isi materi sesuai dengan tujuan pengukuran

Tujuan pengukuran hasil belajar ranah kognitif memiliki 6 (enam) tingkatan yaitu ingatan, pemahaman, penerapan, analisis, evaluasi, dan kreasi. Soal uraian sebaiknya tidak mengukur level ingatan karena merangsang siswa untuk menyontek. Berilah kesempatan kepada siswa untuk memahami, menerapkan pengetahuan dan menganalisis sehingga soal dapat diberikan dalam bentuk permasalahan yang menuntut solusi.

Tujuan pengukuran: Siswa dapat menjelaskan kembali prosedur pembuatan choux pastry. Mengukur ranah kognitif level C2 atau pemahaman

Baik : **Buatlah diagram alir proses pembuatan choux pastry**

Kurang baik : Tulislah resep choux pastry dengan ukuran bahan yang tepat

Keterangan: soal hanya mengukur level ingatan yang mendorong siswa untuk menyontek karena ukuran bahan harus tepat. Dalam belajar pastry, ukuran bahan memang harus tepat tetapi bahan bisa disiapkan, ditimbang, dan diukur sebelum pelaksanaan praktik.

- 4) Isi materi yang ditanyakan sudah sesuai dengan jenjang, jenis sekolah atau tingkatan kelas.

Baik : Faktor apa saja yang perlu diperhatikan dalam pelaksanaan sanitasi penyelenggaraan makanan di usaha jasa boga?

Kurang baik : Berilah contoh bagaimana mengendalikan bahaya keracunan makanan melalui 7 prinsip HACCP (Hazard Analysis Critical Control Point) di UMKM.

Keterangan: Soal untuk level SMK sebaiknya menggunakan istilah yang lebih sederhana sedangkan untuk mahasiswa bisa diberikan soal yang memerlukan keterampilan berpikir tingkat tinggi (higher order thinking skills/HOTS).

- 5) Rumusan kalimat soal atau pertanyaan menggunakan kata tanya atau perintah yang menuntut jawaban terurai

Baik : Mengapa panas yang terlalu tinggi pada saat memanggang cake dapat menyebabkan case hardening?

Kurang baik : Istilah apa yang paling tepat digunakan untuk menggambarkan kegagalan produk cake dengan karakteristik bagian tepi cake sudah gosong tetapi bagian tengahnya masih mentah?

Keterangan: Jawaban hanya terdiri atas satu kata saja, soal tidak menuntut jawaban analisis proses kejadiannya

- 6) Ada petunjuk yang jelas tentang cara pengerjaan soal

Baik	: Buatlah buku catatan penerimaan dan penyimpanan bahan makanan berdasarkan data yang tersedia dengan menggunakan format berikut ini.					
	Tanggal	Nama bahan	Penerimaan	Penggunaan	Stok	Status
Kurang baik	: Buatlah administrasi penerimaan dan penyimpanan bahan makanan menggunakan sistem periodik.					

Keterangan: Soal kedua menimbulkan banyak penafsiran karena petunjuk yang kurang jelas.

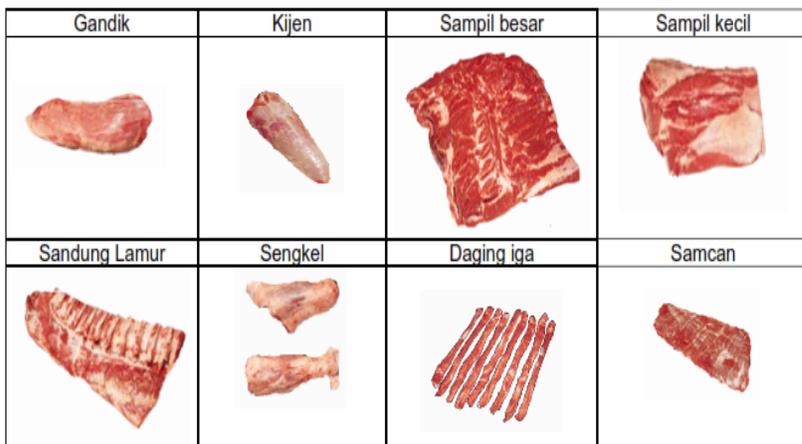
7) Ada pedoman penyekoran

Pedoman penyekoran menjadi satu dengan kunci jawaban. Skor setiap butir tidak sama tergantung pada tingkat kesulitan butir soal. Skor total seluruh tes yang telah digabung menggunakan rentang skor standar yaitu 1-10 atau 10-100. Contoh rubrik penilaian soal uraian:

No	Soal	Kunci jawaban	Rincian Skor
			10 - jika menjawab lengkap dan benar
			8 - jika hanya menjawab 4 poin yang benar
			6 - jika hanya menjawab 3 poin yang benar
			4 - jika hanya menjawab 2 poin yang benar
			2 - jika hanya menjawab 1 poin yang benar
			0 - jika tidak menulis jawaban

8) Grafik, gambar, tabel, peta, diagram dan yang sejenisnya disajikan dengan jelas dan terbaca.

Perhatikan dua gambar berikut ini, kemudian pilihlah bagian daging yang cocok untuk memasak rawon dan sup. Jelaskan alasannya!



Gambar 3.7 Klasifikasi Potongan Daging SNI: 3932: 2008

Sumber: <http://bbpkhcinagara.bppsdp.pertanian.go.id/wp-content/uploads/2015/10/who-is-butcher4.jpg>

9) Perumusan kalimat soal komunikatif

Baik : Susunlah menu makan siang yang memenuhi unsur gizi seimbang, memiliki variasi rasa, warna, tekstur dan teknik olah

Kurang baik : Susunlah menu makan siang yang tidak **mengulangi konsistensi** dan **dominasi cita rasa sama**

Keterangan: Soal memiliki bahasa yang kurang komunikatif, padahal terdapat kata-kata yang lebih tepat digunakan yaitu “variasi”

10) Butir soal menggunakan bahasa Indonesia yang baik dan benar

Baik : Sebutkan 5 contoh hidangan dessert

Kurang baik : Hidangan ini merupakan hidangan yang disajikan setelah makanan utama. Bagi **kita** lebih dikenal dengan istilah **pencuci mulut**. Berilah 5 contoh hidangan pencuci mulut!

Keterangan: Bahasa ilmiah tidak lazim menggunakan nama diri (**kita**). Pencuci mulut bahasa kiasan yang dapat memiliki makna lain

11) Rumusan soal tidak menggunakan kata-kata/kalimat yang menimbulkan penafsiran ganda atau salah pengertian.

Baik : Hitunglah kebutuhan gizi pada penderita diabetes mellitus yang memiliki umur 46 tahun, aktivitas ringan, berat badan 64 kg dan tinggi badan 165 cm. Susunlah menu diet yang sesuai dengan penyakit dan kebutuhan gizinya.

Kurang baik : Susunlah menu diet pada penderita diabetes mellitus yang mengalami kekurangan gizi. Bagaimana karakteristik wisatawan nusantara dan domestik?

Keterangan: (1) Diet perlu disesuaikan dengan kebutuhan energi yang ditentukan oleh berat badan, tinggi badan, umur dan aktivitas fisik. Jumlah kebutuhan energi yang belum diketahui dapat menyebabkan kesalahan pengukuran porsi makanan; (2) Wisatawan domestic memiliki pengertian yang sama dengan nusantara (asli daerah, dalam negeri).

12) Tidak menggunakan bahasa yang berlaku setempat

Baik : Susunlah menu dan rancang kebutuhan bahan tumpeng 7 bulanan untuk 40 orang.

Kurang baik : Ibu Ani mau mengadakan acara **mitoni**, rancanglah biaya **ugo rampe** konsumsi untuk 40 orang tamu.

Keterangan: Mitoni dan ugo rampe merupakan bahasa khas Yogyakarta yang mungkin kurang dikenal di daerah lain.

- 13) Rumusan soal tidak mengandung kata-kata yang menyinggung perasaan siswa

Baik : **Jelaskan bagaimana strategi pemasaran dalam usaha pastry dan cake yang berorientasi ke masa depan?**

Kurang baik : Mengapa usaha *pastry* dan *cake* yang dikelola oleh ibunya Maria saat ini mengalami kebangkrutan?

Keterangan: Maria adalah salah satu siswa di sekolah tersebut, sehingga dengan pertanyaan itu bisa membuatnya sedih.

Tes lisan

Tes lisan atau *personal communication* dilaksanakan melalui komunikasi langsung dan bertatap muka antara peserta tes dengan seorang atau beberapa pemberi tes (penguji). Pertanyaan dan jawaban diberikan secara lisan dan spontan. Tes lisan memerlukan daftar pertanyaan utama dan pedoman pensekoran. Penguji sering melakukan *probing* atau *in depth interviewing* ketika mendapatkan jawaban yang belum memuaskan. Dengan teknik ini, siswa yang menguasai materi mendapat kesempatan untuk menunjukkan kemampuannya.

Tes lisan digunakan selama proses pembelajaran maupun waktu khusus untuk tes lisan. Selama proses pembelajaran, pendidik (guru/dosen) dapat mengajukan pertanyaan-pertanyaan ketika mahasiswa sedang pemaparan tugas dan hasil diskusi sehingga peserta didik (mahasiswa) dituntut menunjukkan keterampilan mengemukakan jawaban/gagasan. Tes lisan sangat efektif digunakan untuk menguji tugas penulisan makalah, laporan, dan penguasaan materi pelajaran. Tes lisan dapat mengungkap kejujuran akademik dan tingkat pemahaman mahasiswa terhadap materi yang ditulis atau dipaparkan. Jika mahasiswa dapat mempresentasikan materi tanpa membaca power point, dan menjelaskan kembali isi makalah yang ditulis maka dia sudah dapat memperoleh skor yang tinggi. Sebaliknya, ketika mahasiswa hanya membaca tulisan di power point dan tidak dapat menjelaskan isi makalah yang ditulisnya, sudah pasti dia hanya melakukan plagiarisme, tanpa mau memahami isinya.

Tes Kinerja

Dalam penilaian kompetensi dikenal asesmen alternatif, asesmen otentik atau asesmen kinerja. Assesmen alternatif digunakan untuk memberi penilaian kinerja (*performance assessment*) atau hasil belajar peserta didik secara multidimensional pada situasi nyata (otentik). Asesmen ini dilaksanakan dengan cara mengobservasi dan mengevaluasi suatu proses, kinerja, perilaku dimana dalam proses tersebut akan muncul keterampilan, sikap, dan produk secara bersama-sama. *Performance assessment* dan *personal communication assessment* termasuk pada kategori *alternative assessment*. Dua alat asesmen tersebut memiliki ciri-ciri pengukuran secara langsung (*direct*) dan otentik selama proses pembelajaran.

Objek *performance assessment* (asesmen kinerja) ini adalah segala sesuatu kemampuan yang dapat diobservasi. Kinerja tidak harus berwujud keterampilan tetapi bisa berupa penyampaian gagasan yang membutuhkan proses kognitif yang kompleks seperti: menganalisis masalah, memecahkan masalah, melakukan percobaan, membuat keputusan, mengukur, bekerja sama dengan teman lain, membuat produk (benda, barang, jasa), dll. *Assessment* alternatif dapat juga digunakan untuk menilai kebiasaan berpikir (*habit of mind*), cara bekerja, nilai-nilai.

Penilaian unjuk kerja (*performance*) dilakukan dengan cara mengamati kegiatan peserta didik pada saat melakukan pekerjaan/tugas praktik di laboratorium, bengkel atau tempat kerja. Tingkat kemampuan diukur berdasarkan standar kompetensi kerja yang dicapai peserta didik pada saat melaksanakan pekerjaan tertentu. Oleh sebab itu, dalam rubrik penilaian keterampilan perlu diuraikan indikator pencapaian kompetensi kerja secara jelas. Penilaian unjuk kerja dilakukan dengan mengisi daftar cek ("ya"/"tidak") pada indikator-indikator pencapaian kompetensi kerja tertentu. Peserta didik dinyatakan "kompeten" apabila seluruh indikator terpenuhi (ya) atau mendapat skor nilai ≥ 7.00 dan "tidak kompeten" diberi nilai < 7.00 yaitu apabila ada indikator yang tidak terpenuhi. Jika peserta memiliki semua indikator pencapaian kompetensi keterampilan, maka mereka berhak mendapat sertifikat kompetensi.

Rubrik penilaian unjuk kerja keterampilan terdiri atas lima komponen yaitu: persiapan kerja, proses, hasil kerja, sikap kerja dan waktu. Kegiatan yang dinilai pada komponen persiapan antara lain: persiapan diri, bahan dan alat kerja. Kegiatan yang dinilai pada komponen proses adalah

pelaksanaan pekerjaan yang sesuai dengan *standard operating procedure* (SOP). Penilaian hasil kerja memasak biasanya mencakup karakteristik organoleptik makanan (citarasa, warna, tekstur, aroma) dan kreativitas penyajian. Penilaian sikap kerja ditetapkan pada sikap kerja esensial untuk melakukan pekerjaan tertentu. Dalam praktikum memasak, sikap kerja esensial yang perlu dinilai antara lain: kedisiplinan, kerapian, kerja keras, dan tanggung jawab. Dalam praktikum pelayanan makan dan minum, sikap kerja esensial adalah etika, keramah-tamahan, dan komunikasi dengan tamu. Komponen waktu penting untuk dinilai karena indikator seseorang yang sudah kompeten harus dapat menyelesaikan pekerjaan dengan baik, tepat, dan cepat.

Setiap mata pelajaran keterampilan memiliki kriteria unjuk kerja yang berbeda. Dalam bidang keahlian Tata Boga, dikenal beberapa mata pelajaran yang memiliki karakteristik berbeda misalnya mata pelajaran praktik pelayanan makan dan minum, pengolahan makanan, seni penyajian makanan dan manajemen usaha. Penilaian kinerja praktik keterampilan perlu memperhatikan unsur-unsur sebagai berikut: (1) kelengkapan dan ketepatan aspek yang dinilai dalam kinerja tersebut; (2) kemampuan-kemampuan khusus yang diperlukan untuk menyelesaikan tugas; (3) hanya kemampuan esensial saja yang diamati; (4) kemampuan yang dinilai diurutkan berdasarkan prosedur kerja. Berikut ini ada beberapa contoh rubrik penilaian kompetensi di bidang *food service*, *basic culinary art*, dan *pastry*.

Tabel 3.8 Lembar Penilaian Ujian Praktik Kejuruan
Mata pelajaran : Melayani Makan dan Minum di Restoran

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
I	Persiapan Kerja			15
	1. Bahan			
	2. Alat			
	3. Penampilan/performance			
	a) Pakaian kerja pramusaji			
	b) Sepatu bersih dan disemir			
	c) Rambut disisir rapi			

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
	d) Kuku dipotong pendek dan bersih			
	Rerata skor komponen I			
II	Proses (Sistematika & Cara Kerja)			30
	2.1. Cara Kerja persiapan ruang restoran			
	2.1.1. Menyiapkan ruang/area restoran untuk layanan			
	2.1.2. Polishing peralatan			
	2.1.3. Membuat lipatan napkin			
	2.1.4. Membuat table set up			
	2.1.5. Menyiapkan daftar menu yang akan diberikan			
	1.2. Menyambut tamu			
	2.2.1. Menyambut kedatangan tamu			
	2.2.2. Mempersiapkan tamu untuk duduk			
	2.2.3. Membuka napkin			
	2.2.4. Menuangkan air ke dalam water goblet (<i>pooring the water</i>)			
	2.2.5. Menawarkan menu segera setelah tamu duduk			
	2.2.6. Menulis pesanan sesuai pesanan tamu			
	2.2.7. Mengulang kembali pesanan yang dipesan			
	2.3. Memproses dan mengambil pesanan			
	2.3.1. Memeriksa pesanan tamu agar tidak terjadi kesalahan			
	2.3.2. Menyampaikan pesanan tamu ke bagian dapur			
	2.3.3. Mengambil pesannan yang sudah dicatat			
	2.4. Menyajikan makanan dan minuman			
	2.4.1. Melakukan <i>adjusting cutleries</i>			
	2.4.2. Menyajikan bread and butter			

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
	2.4.3. Menyiapkan makanan dan minuman sesuai pesanan			
	2.4.4. Menata hidangan di side board sebelum disajikan kepada tamu			
	2.4.5. Memeriksa kembali hidangan sebelum dibawa tamu			
	2.4.6. Menyajikan makanan dan minuman kepada tamu (4 giliran makan)			
	2.4.7. Teknik mengangkat piring kotor (clear up)			
	2.4.8. Teknik <i>crumbing down</i>			
	2.4.9. Menyajikan <i>tea or coffee</i>			
	2.5. Memproses bill			
	2.5.1. Memberikan bill kepada tamu			
	2.5.2. Memproses rekening sesuai prosedur			
	2.5.3. Memeriksa kembali bill yang sudah diterima			
	2.6. Mengantar tamu keluar area restoran dan menutup area restoran			
	2.6.1. Menyapa dan mengantar tamu dengan ramah ketika keluar dari area restoran/ruang makan			
	2.6.2. Membersihkan dan menyimpan peralatan yang sudah dibersihkan			
	2.6.3. Membersihkan area restoran			
	2.7. Cara Kerja Skirting			
	2.7.1. Kain skirting tidak menyentuh lantai			
	2.7.2. Skirting dapat dilihat dari 4 sisi			
	2.7.3. Kreatifitas bentuk skirting			
	2.7.4. Teknik skirting			
	2.8. Cara Kerja Napkin Folding			

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
	2.8.1. Bentuk lipatan napkin rapi			
	2.8.2. Variasi bentuk napkin folding			
	2.8.3. Kreatifitas bentuk napkin folding			
	Rerata skor komponen II			
III	Hasil Kerja			30
	3.1. Mise en Place			
	3.2. Kelengkapan			
	3.3. Tampilan Table set up			
	3.4. Tampilan napkin folding			
	3.5. Pelayanan			
	• Kebersihan dan kerapihan			
	• Penampilan keseluruhan			
	Rerata skor komponen III			
IV	Sikap Kerja			15
	4.1. Etika menghadapi tamu			
	4.2. Penggunaan alat			
	4.3. Kemampuan berkomunikasi dengan tamu			
	4.4. Kontak mata dengan tamu			
	4.5. Tanggap terhadap kebutuhan tamu			
	4.6. K3			
	Rerata skor komponen IV			
V	Waktu			10
	Waktu penyelesaian semua tahapan praktik			

Dalam penilaian kompetensi *food service*, unsur yang dinilai sangat rinci, sesuai dengan prosedur atau urutan melayani tamu di restoran.

Siswa yang baru latihan mungkin belum dapat menunjukkan kegiatan sesuai prosedur sehingga mereka harus mengulangi lagi dari awal. Dalam uji kompetensi, siswa yang tidak mampu menunjukkan kegiatan sesuai prosedur dianggap gagal. Dalam praktik penilaian keterampilan, skor masing-masing komponen penilaian ditetapkan berdasarkan perolehan rerata skor dari sub komponen penilaian tersebut.

Pada mata pelajaran praktik pengolahan makanan, unsur kompetensi yang dinilai berbeda dengan pelayanan makan dan minum. Oleh sebab itu, diperlukan rubrik penilaian yang berbeda. Pada tabel 3.9 terdapat contoh rubrik penilaian UKK (Uji Kompetensi Keterampilan) membuat kue (patiseri).

Tabel 3.9 Contoh Lembar Penilaian UKK Patiseri

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
I	Persiapan Kerja			20
	1.1 Perencanaan menu, resep, design, bahan dan alat, tertib kerja, perhitungan harga jual			
	1.2. Pakaian kerja dikenakan sesuai dengan standar yang berlaku			
	1.3. Mice en place			
	1.4. Hygiene dan sanitasi			
	Rerata skor komponen I			
II	Proses (Sistematika & Cara Kerja)			30
	2.1 Sistematika kerja			
	2.2. Cara kerja			
	2.2.1. Membuat adonan			
	2.2.2. Membentuk adonan			
	2.2.3. Mengolah			
	2.2.4. Menghias			
	2.2.5. Menata/mendisplay			
	Rerata skor komponen II			

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
III	Hasil Kerja			25
	3.1. Bentuk			
	3.2. Warna			
	3.3. Rasa			
	3.4. Tekstur			
	3.5. Penataan kue			
	3.5.1. Kreativitas penataan kue			
	3.5.2. Garnish			
	3.5.3. Kebersihan dan kerapihan			
	3.5.4. Penampilan keseluruhan/Display			
	Rerata skor komponen III			
IV	Sikap Kerja			15
	4.1. Berpedoman pada K3			
	4.2. Kejujuran			
	4.3. Kedisiplinan			
	4.4. Tanggung jawab			
	Rerata skor komponen IV			
V	Waktu			10
	Penggunaan waktu yang tersedia			

Dalam pengolahan kue (patiseri) persiapan dan proses mendapat bobot yang tinggi karena jika dua kegiatan ini sudah dilakukan dengan baik dan benar, hasil praktik juga akan baik dan benar. Contoh lembar penilaian yang sudah dituliskan di atas merupakan penilaian keterampilan tunggal. Dalam lomba kompetensi siswa (LKS), satu orang peserta kemungkinan dituntut menunjukkan beberapa produk sekaligus misalnya memasak hidangan menu makan siang, menu diet, atau menu table d'hote yang terdiri atas beberapa masakan seperti makanan pembuka, pokok dan penutup. Berikut contoh lembar penilaian LKS untuk hidangan menu

kontinental yang terdiri atas tiga masakan yaitu masakan untuk makanan pembuka, pokok dan penutup.

Tabel 3.10 Lembar Penilaian Lomba Kompetensi Siswa

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)
		<7	≥7	
I	Persiapan Kerja			5
	1.1. Bahan			
	1.2. Alat			
	1.3. Pakaian Kerja			
	• Chef jacket dan apron			
	• Safety shoes			
	• Rambut ditutup dengan topi atau karpus			
	Rerata skor komponen I			
II	Proses (Sistematika & Cara Kerja)			50
	2.1 Sistematika kerja			
	2.2. Cara Kerja Salad/Pembuka			
	2.2.1 Menyiapkan bumbu/dressing			
	2.2.2. Mengolah dengan teknik yang tepat			
	2.2.3. Menata hidangan			
	2.2.4. Menghias/garnish			
	2.3. Cara Kerja Hidangan Utama			
	2.3.1 Menyiapkan bumbu			
	2.3.2. Menyiapkan condiment			
	2.3.2. Mengolah dengan teknik yang tepat			
	2.3.3. Menata hidangan			
	2.3.4. Menghias/garnish			
	2.4. Cara Kerja Hidangan penutup			
	2.4.1 Menyiapkan saus/bahan garnish			
	2.4.2. Mengolah dengan teknik yang tepat			
	2.4.3. Menata hidangan			

No	Komponen/Sub komponen Penilaian	Skor nilai		Bobot (%)	
		<7	≥7		
	2.4.4. Menghias/garnish				
	Rerata skor komponen II				
III	Hasil Kerja	Skor Komponen Hidangan			
		Pembuka	Utama	Penutup	Rerata
	3.1.1 Bentuk				
	3.2.2 Warna				
	3.1.3 Rasa				
	3.1.4 Tekstur				
	3.1.5. Suhu				
	3.1.6 Kreasi				
	Rerata skor komponen III				
	3.4. DISPLAY				
	3.4.1. Keserasian				
	3.4.2. Penataan				
	3.4.3. Centre piece				
	3.4.4. Penampilan keseluruhan				
	3.4.5. Kebersihan dan Kerapihan				
	Rerata skor komponen III				
IV	Sikap Kerja				15
	4.1. Etika				
	4.2. Penggunaan alat				
	4.3. K3 dan sanitasi				
	Rerata skor komponen IV				
V	Waktu				10
	Penggunaan waktu yang tersedia				

Sumber : Lembar Penilaian LKS SMK

Seleksi pegawai dalam bidang keahlian Boga (hotel, restoran) dilakukan menggunakan sertifikat kompetensi. Untuk memperoleh sertifikat kompetensi, seseorang harus mengikuti uji kompetensi keahlian. Pada tabel 3.11 terdapat indikator yang digunakan untuk menilai kompetensi keahlian seni kuliner yang dikelola oleh the American Culinary Federation untuk sertifikasi keahlian seni kuliner (*The Culinary Arts Certification assessment*).

Tabel 3.11 Indikator Uji Kompetensi Keahlian Basic *Culinary Art*

No	Indikator	Sub Indikator	Bobot
1.	Recipes	<ul style="list-style-type: none"> – Menunjukkan pemahaman tentang istilah-istilah memasak – Membaca, mengikuti dan mengubah resep baku – Menetapkan mise en place (mis : menetapkan komposisi bahan, menyiapkan alat dan meracik bumbu) 	4
2.	Knife Skills	<ul style="list-style-type: none"> – Mengidentifikasi dan membedakan jenis pisau dan istilah yang berkaitan – Menunjukkan kemampuan mengasah pisau dan perawatan pisau yang tepat – Menunjukkan dasar memotong menggunakan pisau 	4
3.	Large and Small Equipment	<ul style="list-style-type: none"> – Mengidentifikasi dan menunjukkan penggunaan alat ukur yang tepat – Identifikasi peralatan dapur besar dan kecil – Memanfaatkan peralatan besar dan kecil 	3
4.	Sanitation and Safety	<ul style="list-style-type: none"> – Menunjukkan pengetahuan tentang metode pembersihan dan sanitasi – Menjaga prinsip-prinsip dasar sanitasi yang diakui secara nasional – Menunjukkan pemahaman tentang pencegahan, penyebab, dan cara mengatasi kecelakaan di tempat kerja – Menjelaskan alat pemadam kebakaran dan prosedur pemadaman kebakaran – Memahami peraturan <i>occupational safety and health administration</i> – Memelihara semua peralatan dapur – Memelihara keamanan dapur dan lingkungannya 	8

No	Indikator	Sub Indikator	Bobot
5.	Cold Food Preparation	<ul style="list-style-type: none"> – Menyiapkan salad dan dressing – Menyiapkan <i>cold soups, appetizers, entrees, and desserts</i> – Menghias dan menata menu per item – Menyiapkan sandwich, mengoles dan mengisi sandwich – Menjelaskan prosedur untuk persiapan buffet dan perawatannya – Mengidentifikasi metode penyimpanan makanan dingin 	8
6.	Fruits, Vegetables, and Starches	<ul style="list-style-type: none"> – Mengidentifikasi buah dan sayuran – Menyiapkan buah dan sayuran untuk memasak – Mengidentifikasi dan menyiapkan masakan kentang – Mengidentifikasi dan menyiapkan pasta/ biji-bijian/kacang-kacangan – Mengidentifikasi metode penyimpanan buah dan sayuran 	5
7.	Bakery Products	<ul style="list-style-type: none"> – Menyiapkan cookies – Menyiapkan quick breads and cakes – Menyiapkan produk adonan beragi – Menyiapkan icings, fillings, sauces, and toppings – Menyiapkan pies, tarts, and pastries – Menyiapkan creams, custards, and mousse – Identifikasi metode penyimpanan produk roti 	10
8.	Stocks, Sauces, and Hot Soups	<ul style="list-style-type: none"> – Menyiapkan kaldu dan glace – Menyiapkan mother (foundation) sauces – Menyiapkan roux and thickening agents – Mengidentifikasi dan menyiapkan sup – Mengidentifikasi metode pendinginan dan penyimpanan, saus, dan sup panas untuk persediaan 	7

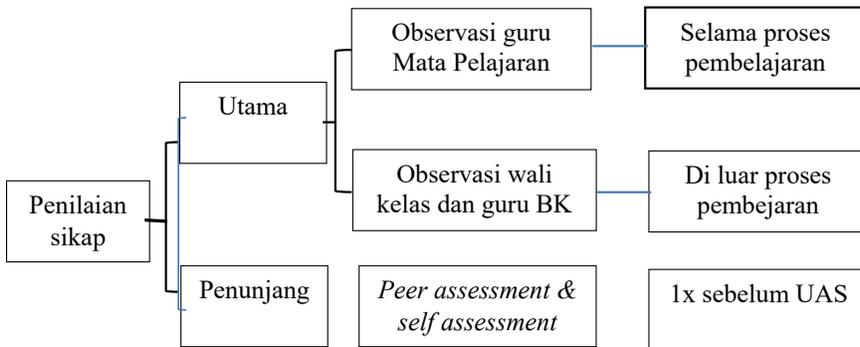
No	Indikator	Sub Indikator	Bobot
9.	Meats, Poultry, Fish, and Seafood	<ul style="list-style-type: none"> – mengidentifikasi dan memilih potongan daging, unggas, ikan, dan seafood dengan benar – menyiapkan daging, unggas, ikan, dan makanan laut – mengidentifikasi berbagai metode memasak – menentukan derajat kematangan yang tepat – mengidentifikasi metode penyimpanan daging, unggas, ikan, dan makanan laut. 	6
10.	Breakfast Foods	<ul style="list-style-type: none"> – Mengidentifikasi dan menyiapkan sarapan pagi – Mengidentifikasi dan menyiapkan daging untuk sarapan – Mengidentifikasi dan menyiapkan masakan telur yang berbeda – Mengidentifikasi dan menyiapkan sereal – Mengidentifikasi metode penyimpanan makanan sarapan pagi. 	6
11.	Dairy Products and Alternatives	<ul style="list-style-type: none"> – Mengidentifikasi berbagai produk susu – Mengidentifikasi alternatif produk susu – Mengidentifikasi varietas keju – Mengidentifikasi dan membedakan mentega dan alternatif mentega – Mengidentifikasi metode penyimpanan untuk produk susu dan alternatifnya 	6
12.	Receiving and Storage	<ul style="list-style-type: none"> – Menginventaris persediaan bahan makanan basah dan kering – Menjelaskan kebijakan dan prosedur HACCP – Membandingkan biaya, jumlah, dan harga bahan-bahan makanan 	5
13.	Nutritional Values	<ul style="list-style-type: none"> – mendiskripsikan kelompok makanan pokok; – menentukan pilihan makanan yang sehat – menjelaskan teknik memasak yang sehat – menjelaskan sumber nutrisi yang bervariasi dan efeknya 	7

No	Indikator	Sub Indikator	Bobot
14.	Management and Employability Skills	<ul style="list-style-type: none"> – menunjukkan pemahaman tentang perilaku profesional, penampilan, dan keterampilan dalam wawancara pekerjaan – menunjukkan pengetahuan tentang berbagai profil pekerjaan dan rantai komando – menghitung gaji – menghitung biaya produk makanan dan persentasinya 	5
15.	Customer Service	<ul style="list-style-type: none"> – Menunjukkan pemahaman tentang prinsip-prinsip dasar pelayanan kepada pelanggan – Menunjukkan kemampuan menata meja, melayani, dan membersihkan meja makan yang sesuai dengan standar – Menghitung jumlah uang yang harus dibayar tamu, termasuk pajak dan bonus – Menunjukkan pengetahuan tentang kepadatan jadwal penggunaan ruang makan, termasuk sistem procedure operating standard (POS) – Memahami item menu dan persiapan bahan khusus yang dipesan – Mengantar dan menghidangkan pesanan tamu 	9
16.	Menu Planning	<ul style="list-style-type: none"> – Menunjukkan pemahaman tentang perencanaan menu. 	2
17.	Sustainability	<ul style="list-style-type: none"> – Menunjukkan pemahaman tentang daur ulang di industri jasa boga – Menunjukkan pemahaman tentang pengomposan dan <i>biodegradability</i> (proses pengomposan sampah-sampah organik yang terurai dengan bantuan Mikroba) di industri jasa boga – Menunjukkan pemahaman tentang efisiensi energi dan keberlanjutan dalam industri jasa boga – Menjelaskan pro dan kontra penggunaan produk lokal atau musiman di industri jasa boga 	5
	Jumlah		100

Sumber : (American Culinary Federation, 2013)

2. Alat Ukur non Tes

Penilaian sikap dalam kurikulum nasional adalah kegiatan untuk mengetahui kecenderungan perilaku spiritual (Kompetensi Inti-1) dan sosial (Kompetensi Inti-2) siswa dalam kehidupan sehari-hari sebagai hasil pendidikan. Penilaian sikap dilakukan dengan menggunakan teknik observasi oleh guru mata pelajaran, guru bimbingan konseling (BK), wali kelas berdasarkan informasi yang relevan dan valid dari berbagai sumber. Hasil penilaian sikap ditulis dalam buku jurnal, catatan anekdot (*anecdotal record*), catatan kejadian tertentu (*incidental record*), Penilaian sikap merupakan akumulasi dari semua kejadian dalam satu semester. Pada umumnya, penilaian sikap yang lebih detil hanya dilakukan terhadap siswa yang menonjol kebaikannya dan keburukannya, sedangkan siswa lain yang memiliki sikap biasa mendapat nilai yang sama dengan nilai rerata kelasnya. Skema penilaian sikap terdapat pada gambar berikut ini



Gambar 3.8 Skema Penilaian Sikap (Direktorat PSMK, 2017)

Penilaian sikap selama proses pembelajaran dilakukan untuk membiasakan siswa agar memiliki sikap-sikap positif terhadap belajar dan lingkungan belajarnya seperti semangat mengikuti pelajaran, bertanggung jawab, jujur, disiplin, santun, mau bekerja sama, menghargai orang lain, dll. Seperti telah dijelaskan pada sub bagian analisis kompetensi sikap bahwa sikap dibentuk dari proses penerimaan stimulus, pemberian respon, pemberian nilai (+/-), pengorganisasian dan internalisasi. Peran guru dalam proses pembentukan sikap sangat penting oleh sebab itu guru wajib memberi teladan, nasehat dan stimulus lain yang positif. Penilaian sikap diharapkan dapat mengendalikan dan mengubah perilaku siswa yang negatif menjadi positif melalui stimulus yang berulang-ulang.

Ranah afektif tidak dapat diukur dengan alat tes karena menyangkut sikap, kepribadian dan motivasi yang melekat pada pribadi seseorang. Meskipun demikian, informasi afektif tetap diperlukan dalam penilaian kompetensi atau hasil belajar. Oleh sebab itu, untuk mengukur ranah afektif diperlukan beberapa alat ukur yang bersifat non tes. Selain untuk mengukur ranah afektif, alat ukur non tes ini dapat pula digunakan untuk melihat kemajuan belajar, pengalaman belajar dan lain-lain. Alat ukur non tes meliputi observasi, penugasan, portofolio, jurnal, inventori, penilaian diri, dan penilaian teman sebaya.

Observasi

Observasi dilakukan untuk mengumpulkan data kualitatif dan kuantitatif sesuai dengan kompetensi yang dinilai, baik dilakukan secara formal maupun informal. Observasi formal dilakukan dengan menggunakan instrumen yang sudah dirancang, sedangkan observasi informal dilakukan tanpa menggunakan instrumen yang sudah dirancang. Dalam penilaian hasil belajar yang terstruktur, guru/dosen sebaiknya menggunakan lembar observasi dan rubrik penilaian kinerja yang telah dirancang sehingga pada saat pelaksanaan tinggal membubuhkan tanda cek (√) jika sikap dan kinerja yang diobservasi muncul.

Tabel 3.12 Contoh Penilaian Sikap dalam Proses Pembelajaran

Nama siswa: Anjar

NIS: ...02

No	Sikap yang dinilai	Self assessment				Teacher assessment			
		1	2	3	4	1	2	3	4
1	Menyelesaikan tugas dengan tekun				√				√
2	Mengumpulkan tugas tepat waktu			√				√	
3	Tidak mencontek pekerjaan orang lain			√				√	
4	Mempresentasikan tugas dengan penuh percaya diri				√			√	
5	Membersihkan tempat kerja dengan penuh tanggung jawab				√			√	

Keterangan:

1. Tidak pernah,
3. Sering

2. Kadang-kadang
4. Selalu

Guru pada umumnya menggunakan lembar observasi sikap untuk kepentingan penelitian tindakan kelas atau quasi eksperimen yang menuntut adanya perubahan sikap selama dilakukan tindakan/ perlakuan. Dalam praktik pembelajaran sehari-hari, format penilaian yang terlalu banyak dapat menambah beban administrasi bagi guru sehingga mengurangi perhatian guru kepada siswa selama proses pembelajaran. Dapat dibayangkan jika satu guru mata pelajaran mengajar 6 kelas dan setiap kelasnya terdiri atas 32 siswa, maka guru harus menilai $6 \times 32 = 192$ siswa yang belum tentu hapal nama-namanya. Oleh sebab itu, penilaian sikap dalam proses pembelajaran dipilih pada sikap-sikap esensial atau sikap kerja yang harus ditunjukkan pada saat uji kompetensi keahlian (UKK) karena satu orang siswa minimal dinilai oleh 2 orang assessor UKK. Jika pembelajaran berbasis riset, penilaian sikap ini menjadi penting untuk menambah data yang bisa diinformasikan kepada orang lain. Hasil penilaian sikap pada tabel 3.13, dapat direkap dalam tabel seperti contoh berikut ini

Tabel 3.13 Contoh Rekap Hasil Penilaian Sikap

No	Nama	Nomor butir													
		Self Assessment						Teacher Assessment							
		1	2	3	4	5	Σ	1	2	3	4	5	Σ		
1	Abid														
2	Anjar	4	3	3	4	4	18	4	3	3	3	3	3	16	
3	Badi														
4	Cecep														
...	dst														

Keterangan nomor butir:

- | | | |
|-------------|-----------------|-------------------|
| 1. Tekun | 3. Jujur | 5. Tanggung jawab |
| 2. Disiplin | 4. Percaya diri | |

Penugasan

Penugasan dapat dilaksanakan dalam bentuk proyek, investigasi, *field study*, atau tugas rumah. Proyek adalah sejumlah kegiatan yang dirancang, dilakukan, dan diselesaikan oleh peserta didik di luar kegiatan kelas dan harus dilaporkan secara tertulis maupun lisan dalam waktu tertentu.

Metode proyek sering digunakan apabila peserta didik diberi tugas membuat produk tertentu melalui proses perancangan, implementasi, monitoring, dan evaluasi. Investigasi dan *field study* juga merupakan tugas belajar di luar kelas yang harus dilaporkan tetapi topik pembelajaran lebih cenderung ke arah penyelidikan, mencari data dan informasi tentang produk/materi yang sedang dipelajari, misalnya investigasi penggunaan *food additive* pada makanan jajanan. Tugas rumah adalah tugas yang harus diselesaikan peserta didik di luar kegiatan kelas, misalnya menyelesaikan soal-soal dan melakukan latihan di luar jam pelajaran. Penugasan yang dilakukan setelah proses pembelajaran (*assessment of learning*) sedangkan penugasan yang diberikan sebelum dan/atau selama proses pembelajaran (*assessment for learning*).

Tanggung jawab profesi pendidik adalah menilai tugas-tugas yang telah diberikan kepada peserta didik dengan menggunakan rubrik penilaian yang sederhana. Tugas yang telah dinilai pendidik kemudian diberi umpan balik dan dikembalikan kepada siswa untuk memberi motivasi belajar. Jika pendidik ketahuan tidak mengoreksi tugas, maka ketika pendidik memberi tugas lagi, peserta didik akan mengumpulkan tugas seadanya, bukan yang sebaik-baiknya. Oleh sebab itu, guru harus mengukur kemampuannya mengoreksi tugas setiap kali akan memberi tugas kepada peserta didik.

Contoh tugas dan penilaian proyek dalam mata pelajaran Manajemen Usaha Boga.

Tugas:

- 1) Carilah order nasi box untuk makan siang, rapat, arisan, dll minimal 25 pack.
- 2) Buatlah rancangan menu, daftar belanja, biaya operasional, perhitungan laba.
- 3) Buat laporan kegiatan dengan format yang telah tersedia.

Tabel 3.14 Contoh Rubrik Penilaian Tugas Proyek

No	Aspek yang dinilai	Skor				Bobot	Nilai
		1	2	3	4		
I	Perencanaan					25	75
	Pemilihan menu			√			
	Perencanaan belanja bahan			√			
	Pembagian tugas			√			

No	Aspek yang dinilai	Skor				Bobot	Nilai
		1	2	3	4		
II	Pelaksanaan					30	120
	Persiapan bahan dan alat				√		
	Pengolahan hidangan				√		
	Pengemasan				√		
III	Hasil					30	105
	Ketepatan rasa,warna, tekstur			√			
	Ketepatan waktu delivery				√		
IV	Pelaporan					15	52,5
	Kelengkapan laporan proyek				√		
	Kerapian penulisan			√			
	Jumlah nilai						352,5/4 =88,1

Keterangan Skor:

1. Kurang 2. Cukup 3. Baik 4. Sangat baik

Tugas proyek dalam mata pelajaran teori berbeda dengan mata pelajaran praktik. Tugas proyek pada mata pelajaran teori dapat berbentuk penelitian, pembuatan media cetak/eletronik, pembuatan poster, dll. Model pembelajaran proyek secara umum dilakukan melalui langkah-langkah: (1) perencanaan tujuan; (2) persiapan, (3) pelaksanaan dan monitoring, (4) penyusunan laporan dan evaluasi. Pelaksanaan kegiatan proyek dapat dilakukan melalui kegiatan workshop, penelitian lapangan, tugas rumah, atau kegiatan lain yang dilakukan di bengkel dan laboratorium. Peran penilaian dalam evaluasi proyek adalah untuk mengendalikan agar langkah-langkah kegiatan berjalan baik dan lancar, oleh sebab itu penilaian proyek dilakukan sejak tahap perencanaan sampai penulisan laporan. Tahap pelaksanaan proyek yang berbeda-beda menuntut format penilaian yang berbeda.

Dalam penilaian proyek terdapat beberapa hal yang perlu dipertimbangkan: (1) kemampuan peserta didik dalam memilih topik/mencari informasi, melaksanakan tugas/proyek, mengelola waktu, dan penulisan laporan; (2) urgensi masalah dalam pencapaian standar kompetensi dan kebutuhan kompetensi tersebut di masyarakat; (3)

kualitas produk secara keseluruhan sesuai dengan kriteria; dan (4) keaslian karya terbebas dari unsur plagiarisme. Proyek yang berkualitas bagus adalah proyek yang memiliki topik menarik, baru dan penting dalam pencapaian kompetensi, dilakukan dengan langkah kerja yang sistematis. Penilaian proyek menggunakan rubrik sesuai langkah kerja yang meliputi perencanaan, pelaksanaan kegiatan, hasil dan pelaporan. Berikut ini terdapat contoh rubrik penilaian proyek penelitian. Penilaian dapat dilakukan dengan cara memberikan skor x bobot atau langsung menggunakan skor maksimal yang dapat diperoleh. Total skor maksimal jika dijumlahkan menjadi 100.

Tabel 3.15 Contoh Rubrik Penilaian Proyek Penelitian

Tahap	Deskripsi	Nilai max
Perencanaan	a) Memilih topik yang menarik dan urgen untuk dilaksanakan b) Merumuskan tujuan yang SMART (<i>specific, measurable, achievable, rational, timely</i>) c) Menyusun langkah kerja dengan jelas dan sistematis d) Menyusun daftar pertanyaan sesuai dengan tujuan.	40
Pengumpulan data/informasi	a) Mengumpulkan data/informasi yang sesuai dengan tujuan b) Mengumpulkan data/informasi yang lengkap isi dan jumlahnya c) Mengentry data/informasi dengan teliti	20
Pengolahan data	a) Mengklasifikasikan data sesuai dengan tujuan b) Mendeskripsikan data dengan jelas c) Menginterpretasikan data secara ilmiah	20
Penulisan laporan	a) Menulis laporan secara sistematis, b) Menggunakan bahasa yang komunikatif. c) Menyimpulkan hasil dan memberi saran.	20
	Total Skor	100

Portofolio

Portofolio adalah kumpulan dokumen, karya-karya pribadi dan hasil-hasil pelaksanaan tugas dalam bidang tertentu yang menunjukkan bukti tentang kompetensi dan pencapaian prestasi belajar siswa pada mata pelajaran tertentu. Menurut buku Penilaian Hasil Belajar SMK,

penilaian portofolio merupakan kegiatan penilaian yang dilakukan dengan menggunakan sekumpulan bukti-bukti hasil belajar (*evidence*) yang relevan dengan kompetensi keahlian yang dipelajari. Evidence tersebut dapat berupa karya peserta didik (hasil pekerjaan) dari proses pembelajaran yang dianggap terbaik, atau bentuk informasi lain yang terkait dengan kompetensi keahlian tertentu. Menurut panduan penilaian SMK (Direktorat PSMK, 2017), penilaian berbasis portofolio merupakan penilaian yang dilaksanakan untuk menilai keseluruhan proses belajar peserta didik termasuk penugasan perseorangan dan/atau kelompok di dalam dan/atau di luar kelas khususnya pada sikap/perilaku dan keterampilan.

Penilaian hasil belajar menggunakan portofolio dilakukan pada saat guru memberi tugas secara berkelanjutan. Langkah-langkah penggunaan penilaian portofolio: (1) Guru menjelaskan tujuan penilaian portofolio, (2) Guru menjelaskan cara pembuatan portofolio; (3) guru menentukan bentuk dokumen/tugas yang dikumpulkan; (4) guru menentukan kriteria penilaian yang digunakan; (5) siswa mengumpulkan dokumen sesuai dengan permintaan; (6) siswa menyusun dokumen sesuai klasifikasinya, menulisi dokumen dan kolom nilai; (7) siswa menjilid dokumen yang telah disusun dan menilai portofolionya sendiri (*self assessment*) menggunakan pensil pada kolom nilai yang telah dibuat; (8) guru memberikan umpan balik terhadap portofolio yang belum sesuai dengan kriteria.

Dokumen portofolio adalah dokumen yang menjadi bukti pelaksanaan kegiatan, tugas atau pekerjaan tertentu. Guru dalam memberi tugas pengumpulan dokumen portofolio perlu memikirkan bahwa tugas yang diberikan merupakan tugas yang dapat dilaksanakan oleh siswa itu sendiri (tidak tergantung pada kegiatan yang disediakan orang lain), siswa memiliki potensi dan kemampuan untuk mengerjakannya. Dokumen portofolio dapat berupa kumpulan tugas mingguan, nilai tes formatif, laporan kegiatan rutin, tugas merangkum, penilaian dari sumber belajar lain pada saat studi lapangan, sertifikat kompetensi, piagam penghargaan, karya ilmiah, bukti-bukti keikutsertaan dalam kegiatan ilmiah/organisasi, dsb. Pada akhir semester kumpulan dokumen tersebut digunakan sebagai referensi tambahan untuk mendeskripsikan kompetensi yang telah dicapai siswa. Contoh rubrik penilaian portofolio pada mata pelajaran Catering yang dilaksanakan melalui metode *project based learning*. Portofolio kompetensi diklasifikasikan menjadi 3 komponen kognitif, afektif, dan psikomotor. Sedangkan substansi materi Catering dinilai melalui kepuasan konsumen.

Tabel 3.16 Contoh Rubrik Penilaian Portofolio

Bagian A. Ranah kognitif

No	Kegiatan	Skor				Dokumen
		4	3	2	1	
1.	Merancang menu dan resep	√				Persiapan proyek
2.	Merancang biaya produksi		√			
3.	Merancang harga jual		√			
4.	Menghitung laba yang akan diperoleh		√			
	Total skor	13				

Bagian B. Ranah Afektif

No	Kriteria	Skor				Dokumen
		4	3	2	1	
1.	Mengantar pesanan tepat waktu	√				Bukti penerimaan dan pengiriman pesanan
2.	Membuat produk sesuai pesanan	√				Nota pesanan dan foto produk
	Total skor	8				

Bagian C. Penilaian Kepuasan Konsumen

No	Penilaian kepuasan konsumen terhadap:	Skor				Dokumen
		4	3	2	1	
1.	Mutu produk		√			Data kuesioner yang diisi pelanggan
2.	Mutu layanan	√				
3.	Harga jual	√				
4.	Keramahtamahan	√				
	Total skor	15				

Bagian C. Ranah Psikomotor

No	Kegiatan Penerimaan Pesanan dan penjualan	Satuan	Skor			
			60	50	40	30
1.	Kue	Biji	<100	71-100	40-70	>40
2.	Lauk-pauk	Biji	<100	71-100	40-70	>40
3.	Sayur	Porsi	<100	71-100	40-70	>40
4.	Minuman	Gelas	<100	71-100	40-70	>40
5.	Snack	Box	<70	51-70	31-50	>30
6.	Menu makan siang/malam	Box	<70	51-70	31-50	>30
7.	Menu prasmanan	Porsi	<70	51-70	31-50	>30

Tabel 3.17 Contoh Penilaian Dokumen Portofolio Project ke 1

Tanggal	Deskripsi kegiatan/tugas	Jumlah	Skor Bagian				Total
			A	B	C	D	
01/02/18	Menerima pesanan snack box untuk acara rapat sekolah	40 pack	13	8	15	60	86
08/02/18	Menerima pesanan lunch box untuk acara pengajian	60 pack	14	7	12	50	83

Proyek ke dua dan seterusnya dilakukan dengan cara yang sama. Dalam praktik penilaian sehari-hari, portofolio bisa disederhanakan menjadi kumpulan nilai tugas mingguan. Tugas mingguan yang telah dinilai dikumpulkan kemudian dijilid dalam bentuk buku portofolio. Penilaian portofolio sudah pernah digunakan untuk sertifikasi guru, pemilihan pendidik dan tenaga kependidikan berprestasi, penilaian angka kredit untuk kenaikan jabatan guru dan dosen, dll. Berikut ini terdapat lembar komponen dokumen portofolio yang harus dikumpulkan guru untuk memperoleh sertifikat pendidik dan lomba pemilihan guru berprestasi.

Tabel 3.18 Contoh Komponen Penilaian Portofolio Guru

NO	KOMPONEN PORTOFOLIO
1.	Kualifikasi akademik
2.	Pendidikan dan pelatihan
3.	Pengalaman mengajar
4.	Perencanaan dan pelaksanaan pembelajaran
5.	Penilaian dari atasan dan pengawas
6.	Prestasi akademik
7.	Karya pengembangan profesi
8.	Keikutsertaan dalam forum ilmiah
9.	Keikutsertaan dalam organisasi di bidang kependidikan dan sosial
10.	Penghargaan yang relevan dengan bidang pendidikan

Komponen penilaian beban kerja dosen (BKD) dan angka kredit kumulatif (*cumulative credit point*) pada prinsipnya mengacu pada penilaian dokumen portofolio. Komponen penilaian BKD terdiri atas tiga komponen utama dan satu penunjang. Komponen utama dikenal dengan nama Tri Dharma Perguruan Tinggi yang terdiri atas: (A) pendidikan dan pengajaran; (B) penelitian dan publikasi; (C) pengabdian kepada masyarakat. Komponen penunjang yang dapat dinilai antara lain adalah: kegiatan dosen dalam forum ilmiah seperti seminar, *workshop* kepanitiaan dan kegiatan akademik lainnya.

Contoh penggunaan penilaian portofolio dalam pembelajaran adalah sebagai berikut.

- 1) Laporan pelaksanaan kegiatan pembelajaran di luar kelas, antara lain pembelajaran yang menggunakan metode PjBL, investigasi, dan *field study*.
- 2) Dokumentasi persiapan dan hasil praktik siswa di laboratorium/ bengkel kerja.
- 3) Kumpulan tugas untuk pendalaman materi yang telah dipelajari di sekolah

- 4) Kumpulan dokumen elektronik pada tugas pembuatan media video, audio, dan pemrograman komputer.
- 5) Fotokopi piagam atau tanda penghargaan yang pernah diterima oleh siswa yang mendukung pencapaian kompetensi.
- 6) Kumpulan peta konsep (mindmap), diagram alir, rangkuman atau review journal yang menunjukkan bukti pemahaman terhadap materi yang telah dibaca.

Jurnal

Jurnal merupakan catatan pendidik selama proses pembelajaran yang berisi informasi kekuatan dan kelemahan peserta didik yang terkait dengan kinerja ataupun sikap peserta didik yang dipaparkan secara deskriptif. Dalam kurikulum 2013, perilaku siswa yang dicatat di dalam jurnal adalah perilaku yang sangat baik dan/atau kurang baik yang berkaitan dengan indikator kompetensi inti (KI) sikap spiritual (KI-1) dan sikap sosial (KI-2), Setiap catatan memuat deskripsi perilaku yang dilengkapi dengan waktu dan tempat kejadiannya.

Tabel 3.19 Contoh Jurnal dari Observasi Selama Satu Semester

Tanggal	Tempat	Nama/NIM	Perilaku	Butir sikap
14/03/18	Lapangan baktet	Santo/15	Berkelahi dengan siswa yang menjadi lawan main basket	Sosial
04/04/18	Kelas Xa	Kirno/21	Memprovokasi keributan pada saat kelas lain sedang ulangan	Toleransi
05/05/18	Kelas Xb	Mada/20	Menyontek pekerjaan temannya	Jujur

Prosedur pembuatan dan penggunaan jurnal menurut buku panduan penilaian SMK (Direktorat PSMK, 2017) adalah sebagai berikut:

- 1) Jurnal digunakan oleh wali kelas, guru mata pelajaran, dan guru BK selama periode satu semester;
- 2) Jurnal oleh wali kelas digunakan untuk satu kelas, jurnal oleh guru mata pelajaran digunakan untuk seluruh siswa yang mengikuti mata pelajarannya, dan bagi guru BK untuk semua siswa di bawah bimbingannya;
- 3) Hasil observasi guru mata pelajaran dan guru BK diserahkan kepada wali kelas untuk diolah lebih lanjut;

- 4) Indikator yang diamati dapat dikembangkan sesuai dengan kebutuhan, kondisi, atau ciri khas satuan pendidikan;
- 5) Catatan dilakukan selama satu semester hanya pada siswa-siswa yang menunjukkan perilaku sangat baik atau kurang baik, sehingga ada kemungkinan dalam satu hari hanya ada beberapa orang atau bahkan tidak ada yang menunjukkan perilaku yang sangat baik dan/atau kurang baik sesuai dengan indikator perilaku yang diamati;
- 6) Perilaku sangat baik atau kurang baik yang dicatat dalam jurnal tidak terbatas pada butir-butir sikap (perilaku) yang hendak ditanamkan melalui pembelajaran yang saat itu sedang berlangsung sebagaimana dirancang dalam RPP, tetapi dapat mencakup butir-butir nilai sikap lainnya yang ditanamkan dalam semester itu selama butir nilai sikap tersebut ditunjukkan oleh siswa melalui perilakunya; Perilaku siswa yang sudah sesuai dengan harapan tidak perlu dicatat karena menunjukkan tujuan sudah tercapai.

Inventori

Inventori merupakan skala psikologis yang dipakai untuk mengungkapkan sikap, minat, dan persepsi peserta didik terhadap sesuatu objek. Inventori dapat ditulis dalam bentuk angket dengan skala jawaban yang dikembangkan oleh: Guttman (ya-tidak), skala Likert (sangat setuju s/d sangat tidak setuju), atau skala semantik diferensial (menggunakan lawan kata misalnya: menerima >< menolak). Berdasarkan hasil pengisian angket ini akan diketahui kecenderungan sikap psikologis seseorang dalam menanggapi suatu kejadian. Contoh inventori untuk mengetahui tanggapan siswa terhadap penerapan kurikulum 2013 pada salah satu guru mata pelajaran.

Contoh:

Bagaimana tanggapan saudara terhadap model pembelajaran *saintific* yang diterapkan guru mata pelajaran Pengolahan Makanan Kontinental?

No	Tanggapan positif	4	3	2	1	Tanggapan negatif
1.	Menyenangkan					Membosankan
2.	Metode mengajar bervariasi					Metode mengajar monoton
3.	Materi cepat dipahami					Materi lambat dipahami
4.	Mudah dilaksanakan					Sulit dilaksanakan
5.	Bebas belajar mandiri					Belajar tergantung guru

Penilaian Diri (self assessment)

Penilaian diri adalah suatu teknik penilaian di mana peserta didik diminta untuk menilai dirinya sendiri yang berkaitan dengan kemampuannya, tingkat partisipasinya, kelebihan dan kekurangan dirinya dalam berbagai hal. Hasil penilaian persepsi diri kemudian dibandingkan dengan hasil penilaian antar teman untuk mengungkap kejujuran dan meningkatkan kemampuan refleksi diri. Contoh lembar penilaian diri untuk mengungkap tingkat partisipasi dan kejujuran dalam tugas diskusi kelompok penyelesaian masalah pengolahan limbah catering ditunjukkan pada Tabel 3.19.

Tabel 3.20 Contoh Instrumen *Self Evaluation*

No	Partisipasi dalam Kegiatan Diskusi	Jawaban			
		SS	S	KS	TS
1.	Mengidentifikasi masalah				
2.	Mengumpulkan data, fakta, dan informasi yang berkaitan dengan masalah				
3.	Mencari teori pendukung untuk mengatasi masalah				
4.	Mencari solusi untuk mengatasi masalah				
5.	Mengusulkan solusi untuk mengatasi masalah				
6.	Merangkum hasil diskusi				
7.	Mempresentasikan hasil diskusi				
8.	Menjawab pertanyaan pada saat presentasi				

Keterangan jawaban

SS: sangat sesuai

S: sesuai

KS: kurang sesuai

TS: tidak sesuai

Catatan:

Guru menyarankan kepada peserta didik untuk menyatakan kondisi yang dimilikinya secara jujur, karena hasil penilaian tidak berpengaruh terhadap nilai akhir tetapi hanya bertujuan untuk perbaikan proses pembelajaran.

Hasil penilaian diri tidak digunakan sebagai alat untuk mengambil keputusan ketuntasan belajar tetapi hanya untuk mengetahui perkembangan/kemajuan belajar siswa. Pada umumnya siswa terdorong untuk mendapatkan nilai yang baik sehingga guru perlu mencocokkan hasil penilaian diri dengan hasil penilaian teman sebaya atau lainnya. Guru harus sering mengingatkan kepada siswa untuk selalu menjaga integritas sampai

siswa terlatih melakukan penilaian diri secara baik, objektif, dan jujur. Jika siswa sudah bisa dipercaya dalam melakukan penilaian diri, maka hasilnya dapat digunakan untuk melengkapi hasil penilaian guru lainnya.

Penilaian teman sebaya (peer assessment)

Penilaian teman sebaya (*peer assessment*) merupakan teknik penilaian dengan cara meminta peserta didik untuk mengemukakan kelebihan dan kekurangan temannya dalam berbagai hal. Penilaian teman sebaya sering digunakan pada metode pembelajaran kelompok seperti *peer teaching* dan *learning together*. Penilaian teman sebaya digunakan untuk mengecek kejujuran penilaian diri (*self assessment*) dan membantu guru dalam menilai hasil kerja kelompok yang heterogen. Aspek yang dinilai adalah kontribusi anggota kelompok dalam penyelesaian tugas-tugas kelompok. Tujuan penilaian agar pembagian tugas anggota kelompok merata, semua anggota kelompok turut berpartisipasi mengerjakan tugas dan jika beban kerja anggota kelompok tidak sama akan diperoleh rasa keadilan karena nilai masing-masing anggota kelompok tidak sama. Kelemahan penilaian teman sebaya jika terlalu sering digunakan adalah siswa dapat bekerjasama dengan temannya untuk saling memberi nilai yang bagus, oleh sebab itu teman yang dinilai dan yang menilai sebaiknya tidak saling tahu atau sering diganti.

Penjaminan mutu hasil penilaian non test dilakukan melalui verifikasi internal dan eksternal. Verifikasi internal dilakukan oleh guru sejawat, ketua kompetensi keahlian, dan wakil kepala sekolah bidang kurikulum di sekolahnya masing-masing. Verifikasi eksternal dilakukan oleh penilai dari lembaga resmi di luar sekolah misalnya *assessor* sertifikasi profesi, praktisi dunia usaha/industri atau asosiasi profesi. Komitmen yang harus dimiliki oleh penilai adalah memahami kriteria unjuk kerja yang harus dikuasai peserta didik sehingga dapat membedakan peserta didik yang kompeten dan tidak kompeten.

3. *Authentic assessment*

Menurut Herrington and Herrington (Herrington & Herrington, A, 1998) penilaian autentik adalah penilaian yang digunakan dalam konteks kegiatan nyata. Penilaian otentik banyak digunakan untuk menilai proses belajar yang menghasilkan produk rumit dan penuh tantangan sehingga membutuhkan beberapa indikator untuk menilai keberhasilannya. Penilaian dilakukan dalam situasi yang realistis, kontekstual, dan kompleks.

Penilaian otentik dikenal dengan nama penilaian kinerja, penilaian alternatif, atau penilaian langsung. Penilaian otentik terdapat pada teknik penilaian non tes menggunakan portofolio, daftar cek, observasi guru, dan proyek kelompok. Semua bentuk penilaian otentik dapat dikuantitatifkan dengan menggunakan rubrik penilaian yang memiliki skala numerik. Hasil penilaian otentik kemudian digabung dengan hasil-hasil penilaian lainnya untuk menetapkan nilai final.

Penilaian autentik dikembangkan sebagai kritik terhadap tes pilihan ganda yang hanya memberi gagasan dangkal tentang apa yang telah dipelajari siswa. Tes pilihan ganda tidak menunjukkan apa yang dapat dilakukan siswa dari kegiatan belajar yang sudah diikutinya (Aiken L. R., 1996). Menurut Aiken (1996) penilaian autentik lebih dapat dipertanggungjawabkan. Dalam penilaian otentik ini, siswa menggunakan proses kognitif mereka untuk menunjukkan apa yang telah mereka pelajari dan apa yang dapat mereka lakukan (Eby, 1998). Penilaian autentik menuntut siswa untuk menggunakan pengetahuan sebelumnya, hasil belajar terbaru, dan keterampilan dalam mengatasi masalah nyata dan rumit. Melalui penilaian otentik pada metode proyek, siswa didukung untuk dapat membuat proyek tentang topik yang dipilih, menyiapkan laporan penelitian dan mempresentasikan produk akhir mereka kepada guru/dosen. (DiMartino & Castaneda, A, 2007).

BAB IV

PENGEMBANGAN TES BAKU

A. Perspektif Penggunaan Tes Baku

Pengembangan tes banyak dilakukan oleh ahli pengukuran psikologi (psikometri) dan tidak menutup kemungkinan untuk dilakukan oleh guru/dosen yang akan mengembangkan perangkat tes baru untuk keperluan evaluasi pembelajaran atau penyusunan bank soal. Butir soal yang dapat dimasukkan ke dalam bank soal adalah butir-butir soal yang sudah baku atau teruji kualitasnya. Untuk mendapatkan butir soal baku yang berkualitas, maka perlu dilakukan pengujian baik secara teoritis/kualitatif maupun empiris/kuantitatif.

Tes baku sering digunakan untuk beberapa keperluan seperti tes prestasi belajar dan tes potensi belajar. Tes prestasi belajar baku sering digunakan oleh sekolah untuk menentukan kelulusan, pemetaan peringkat sekolah dan seleksi masuk sekolah. Ujian nasional merupakan salah satu contoh tes baku yang pernah dipakai untuk penentuan kelulusan dan seleksi masuk ke jenjang sekolah menengah pertama dan atas. Tes potensi belajar dalam serial tes psikologi (kecerdasan, potensi, dan bakat sekolah) sering digunakan untuk seleksi pegawai, seleksi masuk sekolah dan mengetahui bakat serta potensi seseorang. McIntire dan Miller (McIntire & Miller, 2000) menuliskan bahwa tes prestasi digunakan untuk mengukur pencapaian hasil belajar siswa pada wilayah akademik khusus. Tes bakat (*aptitude test*) didesain untuk mengukur potensi belajar atau kemampuan yang akan ditunjukkan dalam wilayah akademik yang tidak dilatih sebelumnya. Penggunaan Tes Potensi Belajar (TPB) sebagai alat seleksi memiliki keunggulan karena materi yang diujikan tidak perlu dihapal, tetapi TPB juga memiliki kelemahan sebagai alat prediksi prestasi

belajar karena dimensi yang diukur pada tes potensi belajar berbeda dengan tes prestasi belajar. Apabila kriteria hasil belajar ditetapkan dengan tes prestasi, maka prediktor yang terbaik juga berupa tes prestasi belajar.

Tes baku menjadi kebijakan dan strategi populer yang digunakan untuk menilai akuntabilitas sekolah sejak tahun 1984 di Carolina Utara. Negara bagian tersebut menerapkan *High-Stakes Testing* (tes yang digunakan untuk menentukan keputusan penting). Prestasi siswa diukur sebagai sarana promosi siswa untuk memperoleh kelulusan dan menuju peringkat kelas berikutnya. Substansi tes terdiri atas tes membaca (*reading*) dan matematika. Kebijakan tersebut telah memberi efek yang cukup luas pada siswa, guru, orang tua dan sekolah. Kebijakan tersebut secara cepat dapat meningkatkan skor prestasi, siswa terdorong untuk bekerja keras dan menuntut guru untuk memberi perhatian yang lebih tinggi pada pencapaian prestasi belajar siswa (Roderick, Jacob, & Bryk, 2002).

High stake test adalah tes sumatif yang menjadi bagian dari penentu tindak lanjut siswa. Pada umumnya *high stake test* menggunakan tes baku seperti *Scholastic Aptitude Test* (SAT) dan *Computerized Adaptive Test* (CAT) (Johnson & Johnson, 2002). Tes baku merupakan strategi penilaian yang mengevaluasi semua siswa dan semua sekolah pada keterampilan dasar yang sama. Hasil tes baku dapat menjadi alasan untuk mengindikasikan sekolah yang mempunyai kinerja tinggi atau tidak. Tes yang dibakukan didesain untuk menetapkan jaminan pendidikan bagi siswa yang akan melanjutkan ke sekolah berikutnya. Nilai tes baku dapat digunakan untuk menentukan letak kemampuan masing-masing siswa dan memprediksi kesuksesan siswa pada masa yang akan datang (Nicola, 2004).

Johnson & Johnson (Johnson & Johnson, 2002) memaparkan keuntungan dan kelemahan penggunaan tes baku. Beberapa keuntungan tes baku adalah: (1) mudah diselenggarakan dan hanya memerlukan waktu yang relatif pendek; (2) menyediakan situasi standar yaitu semua siswa dituntut untuk menjawab pertanyaan yang sama dan menjamin semua siswa dapat dievaluasi pada kriteria yang sama; (3) menyediakan catatan perilaku permanen ketika tes dikerjakan secara tertulis; (4) skor tes baku dapat dibandingkan antar siswa, sekolah, dan kabupaten; (5) Jika tes mempunyai tingkat kesulitan sedang cenderung mempunyai validitas prediktif tinggi. Kelemahan tes baku adalah: (1) isi tes baku hanya mengukur informasi faktual atau deklaratif yang dekat dengan kemampuan verbal dan cenderung tidak mengungkap pemahaman mendalam yang

mengintegrasikan berbagai pengetahuan; (2) tes baku tidak memadai dalam menilai kapabilitas yang dihasilkan siswa; (3) tes baku sangat terbatas untuk mengukur dampak hasil belajar siswa; (4) tes baku tidak mengarah pada audit kurikulum eksternal, dan (5) penggunaan hasil tes sangat terbatas. Nicole (Nicola, 2004) tidak menyetujui penyelenggaraan tes baku karena penilaian tunggal beresiko dan tidak akurat untuk melihat pengetahuan dan keterampilan siswa. Pada saat pelaksanaan tes, banyak siswa yang merasa sakit, kelelahan, depresi atau kondisi lain yang tidak menyenangkan. Stres dapat menyebabkan siswa gagal dalam menempuh ujian karena kecemasan yang tinggi dapat membuat mereka lupa semua informasi yang sudah diketahui.

James Popham dalam Nicole (2004) mengundurkan diri dari kegiatan membuat tes baku setelah mendengar banyak permasalahan yang muncul. Pada umumnya, penulis tes baku membuat klaim positif tentang reliabilitas tes yang dibuatnya agar tes tersebut dibeli dan digunakan. Teori respon butir yang menyatakan sebaran skor bervariasi untuk memperbaiki reliabilitas telah dipenuhi oleh penulis soal dengan membuat sebaran jawaban dari pertanyaan pada beberapa siswa yang mewakili saja (tidak pada semua siswa yang menjawab). Hasil ujian dari tes yang telah dibakukan tidak dapat digunakan untuk mengevaluasi kualitas siswa sebab pertanyaan tertutup dan cenderung mengeliminasi isi yang lebih penting dari pengukuran kematangan siswa yang lebih mendalam.

Teitelbaum (Teitelbaum, 2003) dalam sebuah studinya menemukan upaya pengukuran prestasi melalui tes standar (baku) mata pelajaran matematika dan sains sebagai persyaratan kelulusan telah membuat frustrasi sebagian warga dan sekolah Amerika yang tidak mampu dalam bidang tersebut. Menurut hasil penelitian itu, tes baku lebih menguntungkan ras Asian karena mereka lebih unggul dalam bidang matematika dan sains di banding ras yang lain. Berdasarkan evaluasi kebijakan yang dilakukan ternyata banyak sekolah yang menginginkan alat ukur prestasi yang lebih ringan karena siswa yang cocok masuk sekolah kejuruan selalu kalah bersaing dengan siswa yang hendak masuk ke sekolah umum. Hasil penelitian tersebut menunjukkan bahwa tes baku tidak tepat sebagai penentu kelulusan dan seleksi masuk karena kemampuan masing-masing siswa dan sekolah yang dituju tidak sama. Materi tes baku dari mata pelajaran matematika dan sains tidak dapat diikuti dengan baik oleh siswa yang kemampuan kognitifnya rendah. Prediksi kesuksesan siswa dalam hidup di masa yang akan datang tidak hanya ditentukan oleh kemahiran

dalam ilmu eksakta tetapi juga didukung oleh keterampilan lain seperti keterampilan teknis, komunikasi, emosi, dan kepribadian.

Penggunaan pengukuran ganda (*multiple measurement*) telah berkembang menjadi perspektif baru dalam dunia pendidikan. Menurut Montero (2003: 8) pendekatan pengukuran ganda berusaha untuk meningkatkan reliabilitas dengan cara menurunkan sejumlah kesalahan yang berhubungan dengan evaluasi pada siswa, guru, sekolah dan *local educational agency* (LEAs). Pengukuran ganda yang dilakukan secara terus menerus memiliki beberapa manfaat yaitu: (1) sistem penilaian yang dilakukan secara terus menerus dan berkelanjutan dapat mengukur kemampuan siswa secara lebih komprehensif; (2) alat penilaian ganda dan berkelanjutan yang reliabel atau sudah menunjukkan hasil konsisten dapat menjadi alat evaluasi keberlanjutan siswa yang akurat; dan (3) kemampuan siswa setelah lulus atau meninggalkan kelas sudah dapat diprediksi berdasarkan hasil penilaian ganda yang berkelanjutan.

Berdasarkan beberapa hasil studi yang dipaparkan di atas, tes baku masih perlu dikembangkan asalkan tidak digunakan sebagai satu-satunya alat penentu kelulusan. Ujian akhir semester dapat menggunakan tes baku karena tes baku sudah teruji kualitasnya. Ada beberapa tips untuk dapat membuat butir tes memiliki kualitas baku yaitu: (1) soal tes mengukur ranah kognitif minimal pada level C3 atau aplikasi; (2) butir soal memiliki tingkat kesulitan sedang; (3) materi yang diujikan tidak menyimpang dari materi yang sudah diajarkan.

B. Prosedur Pengembangan Tes Baku

Pembuatan perangkat tes baku harus melewati proses pengembangan dan pengujian yang cukup panjang. McIntire (2000) menetapkan 10 langkah pengembangan tes yang harus dilalui yaitu:

- 1) Defining the test universe, audience, and purpose
- 2) Developing a test plan
- 3) Composing the test items
- 4) Writing the administration instructions
- 5) Conduct piloting test
- 6) Conduct item analysis
- 7) Revising the test

- 8) Validation the test TAHAP
- 9) Developing norms
- 10) Complete test manual

Keterangan

1. *Defining the test universe, audience, and purpose*

Tahap pertama yang harus dilakukan oleh seorang pengembang tes adalah mendefinisikan domain tes secara umum, siapa sasarannya dan untuk keperluan apa tes tersebut dikembangkan. Domain tes perlu dipertegas supaya tes benar-benar mengukur kompetensi yang diharapkan dapat ditunjukkan oleh peserta tes. Sasaran peserta tes menjadi bahan pertimbangan supaya penyusunan butir-butir pertanyaan sesuai dengan karakteristik peserta tes. Sebagai contoh, tes yang ditujukan untuk peserta yang memiliki tingkat pendidikan rendah harus disusun dalam bahasa yang sederhana supaya peserta tes mudah memahaminya.

Tujuan tes menjadi pertimbangan dalam penyusunan tes. Tujuan tes diklasifikasikan menjadi tes seleksi, tes penempatan (*placement*), tes diagnostik, dan tes hasil belajar. Tes seleksi dan penentu kelulusan membutuhkan kualitas butir yang bagus dan memiliki daya pembeda tinggi dibandingkan dengan tes diagnostik yang hanya untuk mengukur potensi kemampuan karena tes tersebut digunakan untuk pengambilan keputusan penting (*high stake testing*). Pengembangan tes masih banyak digunakan untuk tes seleksi atau penentu kelulusan.

Mehrens (1973) mendefinisikan macam-macam tes berdasarkan tujuannya. Tes yang digunakan dalam pembelajaran mempunyai tujuan untuk mengevaluasi dampak belajar, evaluasi mengajar, evaluasi kurikulum, *diagnosis*, kelulusan, dan motivasi. Tes bimbingan (*guidance*) bertujuan untuk mengadakan bimbingan kejuruan, pendidikan, dan personal. Berdasarkan administrasi, tes ada yang bertujuan untuk seleksi, klasifikasi, penempatan, *public relational/ information*, perencanaan dan evaluasi kurikulum, evaluasi guru, penyediaan informasi, dan kelulusan (Mehrens W. A., 1973).

Dalam pasal 17, Peraturan Menteri Pendidikan dan Kebudayaan Republik Indonesia Nomor 4 Tahun 2018 tentang Penilaian Hasil Belajar oleh Satuan Pendidikan dan Penilaian Hasil Belajar oleh Pemerintah. Hasil tes standar nasional atau ujian nasional (UN) digunakan sebagai

dasar untuk: (a) pemetaan mutu program dan/atau Satuan Pendidikan; (b) pertimbangan seleksi masuk jenjang pendidikan berikutnya; dan (c) pembinaan dan pemberian bantuan kepada Satuan Pendidikan dalam upayanya untuk meningkatkan mutu pendidikan. Salah satu fungsi yang telah dihapus adalah hasil UN digunakan untuk penentuan kelulusan. Hal ini terjadi karena beberapa pertimbangan antara lain: mutu pendidikan belum merata di seluruh pelosok Indonesia, tidak ada keadilan karena proses belajar selama bertahun-tahun hanya diukur menggunakan satu alat ukur dan satu kali pengujian; terjadi banyak kecurangan selama proses ujian sehingga validitas prediksi hasil tes rendah.

2. *Developing a test plan*

Setelah informasi yang diperlukan pada tahap pendefinisian tes diperoleh, pengembang tes melanjutkan kegiatan dengan merencanakan tes secara keseluruhan. Hal-hal yang direncanakan meliputi menyusun konstruk/kisi-kisi, format pertanyaan atau jawaban, bentuk penyelenggaraan dan cara penyekorannya.

a. *Konstruk test*

Seperangkat soal tes terdiri atas beberapa konstruk atau indikator/kisi-kisi tes. Konstruk tes dirancang berdasarkan analisis kompetensi dan kompetensi dasar. Kompetensi adalah seperangkat pengetahuan, sikap, dan keterampilan yang harus dimiliki seseorang ketika mereka bekerja pada profesi tertentu. Kompetensi hasil belajar adalah seperangkat pengetahuan, sikap dan keterampilan yang harus dikuasai siswa setelah proses pembelajaran. Rumusan kompetensi hasil belajar sudah disusun pada tujuan pembelajaran. Kompetensi hasil belajar diukur dari komponen KSAOs (*knowledge, skills, abilities and other characteristics*). *Knowledge* diukur dengan tes tertulis dan *skill* diukur menggunakan tes kinerja (*performance test*/uji kompetensi). Penilaian kinerja menggunakan rubrik yang berisi kriteria-kriteria kerja yang standar dan dinilai oleh asesor atau penilai (*rater*). Dalam buku ini dibahas pengembangan tes baku komponen kognitif yang diselenggarakan secara tertulis dengan bentuk jawaban pilihan ganda.

Setiap komponen kompetensi memiliki tingkatan mulai dari yang mudah sampai ke sulit. Pada bab sebelumnya telah ditulis bahwa terdapat 6 tingkatan kompetensi kognitif (*knowledge*) menurut Bloom yaitu *remembering, comprehension, application, analysis, evaluation*, dan

creation. Untuk mendapat hasil tes yang baik, susunlah butir tes pada level *comprehension*, *application*, dan *analysis* supaya kualitas butir soal bagus, memiliki tingkat kesulitan sedang dan daya pembeda tinggi. Soal yang mengukur level ingatan (*remembering*) dapat memicu siswa untuk menyontek dan tidak mengukur kemampuan kognitif tingkat tinggi.

Konstruksi tes untuk penelitian berbeda dengan tes hasil belajar. Pada penelitian, konstruk disusun berdasarkan teori yang di dalamnya terdapat indikator pengukur variabel. Variabel penelitian belum tentu diukur dengan tes tetapi lebih sering diukur menggunakan kuesioner yang tidak memiliki kriteria jawaban benar atau salah. Pada tes hasil belajar, konstruk disusun berdasarkan cakupan materi yang telah diajarkan. Satu indikator/ konstruk dapat dikembangkan menjadi beberapa butir test.

b. Format test

Format tes merujuk pada tipe-tipe pertanyaan dan tipe jawaban pada sebuah tes. Menurut karakteristik jawaban, dikenal tes yang jawabannya sudah tersedia dan peserta tes tinggal memilih serta tes yang jawabannya ditulis sendiri oleh peserta tes. Tes tertulis yang jawabannya sudah tersedia dikategorikan pada tipe *objective test* karena hanya ada satu jawaban benar yang langsung dapat diberi skor dan jawaban salah tidak mendapat skor, tanpa ada pertimbangan lain dari penilai. Tes tertulis yang jawabannya dikembangkan sendiri oleh peserta tes (*constructed-response*: CR) dinamakan *subjective test* karena pada saat penskoran atau pemberian nilai tes, penilai sangat dipengaruhi oleh suasana hati dan pertimbangan lain yang bersifat subjektif.

Tes objektif terdiri atas tes pilihan ganda atau MC (*multiple choice*), benar-salah (*true-false items*), menjodohkan atau mencocokkan (*matching exercises*), isian singkat (*short answer, fill-in items*). Tes subjektif, atau tes yang jawabannya diisi sendiri oleh peserta tes (*constructed-response*: CR) dapat berbentuk jawaban singkat, melengkapi kalimat, menghitung, dan lain-lain. Dalam tes subjektif ini, pokok pertanyaan diberikan dalam beberapa bentuk persoalan kompleks yang menuntut jawaban tertulis berupa paparan, atau solusi terhadap persoalan tersebut. Jawaban maupun penilaian tes semacam ini bersifat subjektif.

Tes baku lebih banyak menggunakan tes objektif yaitu tes yang hanya memiliki satu jawaban benar. Jawaban soal tes berbentuk pilihan ganda dengan 4 sampai 5 opsi/pilihan jawaban. Bentuk tes pilihan ganda ini

mudah dianalisis dan terdapat banyak program komputer yang digunakan untuk menganalisisnya. Respon (jawaban) butir tes pilihan ganda bisa digunakan untuk berbagai kepentingan antara lain: analisis kualitas butir tes, pengukuran validitas, dan reliabilitas secara kuantitatif.

c. Bentuk Penyelenggaraan dan Cara Penskoran

Tes dapat diselenggarakan dalam beberapa bentuk yaitu tes tertulis dengan menggunakan *paper and pencil* (pp), tes lisan atau wawancara, tes kinerja/uji kompetensi, tes interaktif menggunakan komputer, dan tes dikerjakan individu atau kelompok. Bentuk penyelenggaraan tes ini akan berpengaruh terhadap penulisan bentuk pertanyaan/butir tesnya.

Cara penskoran tes juga menjadi bahan pertimbangan dalam pengembangan tes. Model penskoran kumulatif (*cumulative model of scoring*) merupakan model penskoran yang paling umum digunakan untuk penentuan skor tes terakhir tiap-tiap individu. Melalui model ini, peserta tes akan mendapat skor total yang diperoleh dari hasil penjumlahan skor butir.

Bentuk penyelenggaraan tes dapat mempengaruhi kinerja peserta tes dan interpretasinya. Sebuah penelitian yang dilakukan oleh Lunz, & Bergstrom mencoba empat kondisi yang diterapkan untuk menjawab pertanyaan yaitu meloncat (*skip*), melihat kembali (*review*), menunda (*defer*) dan tanpa kondisi (*none*) dalam *Computerized Adaptive Test*. Hasil penelitian menemukan siswa yang diberi perlakuan *skip* menunjukkan kinerja yang lebih baik dari siswa yang tidak mengontrol kembali jawaban mereka. Rerata estimasi menurun dari *skip* ke *review*, ke *defer* dan terakhir ke *none* (Lunz & Bergstrom, 1994). Hasil penelitian lain yang dilakukan oleh Oshima menemukan kondisi tes kecepatan (*speed test*) dapat mempengaruhi estimasi butir dan parameter kemampuan. Hasil mengindikasikan bahwa di bawah kondisi kecepatan, ada korelasi tinggi antara jawaban benar dan parameter estimasi kemampuan (Oshima, 1994). Dalam hal ini, peserta tes yang memiliki kemampuan rendah pada umumnya kurang berhasil jika mengikuti *speed test* karena kemampuan membacanya juga rendah. *Speed test* tepat digunakan untuk seleksi dan kurang tepat jika digunakan untuk test kelulusan.

Oshima (Oshima, 1994) dan Wollack (Wollack, Cohen, & Wells, C. A, 2003) meneliti pengaruh kecepatan terhadap parameter estimasi dalam IRT dan pemeliharaan stabilitas skala dalam tes kecepatan. Penelitian

tersebut menemukan bahwa tes kecepatan tidak mampu mengestimasi kemampuan seseorang karena banyak butir-butir yang tidak mempunyai kesempatan untuk dibaca terutama butir-butir terakhir. Akibat dari tes kecepatan ini, estimasi teori respon butir menjadi kurang tepat. Pengujian yang tidak menggunakan batas waktu yang dipercepat untuk menyamakan dan mengestimasi parameter butir dalam skala multidimensi mempunyai efek yang lebih kecil jika dibandingkan dengan tes kecepatan.

Hasil review jurnal penelitian pengembangan alat tes kemampuan banyak diwarnai oleh strategi pelaksanaan tes dan pengujian instrumen menggunakan software baru. Satu temuan menarik evaluasi penggunaan CAT (*Computer Adaptive Test*) dibandingkan dengan versi kertas dan pena dari penelitian Wang adalah CAT tidak dapat digunakan sebagai pembanding untuk mengestimasi kemampuan berdasarkan skor *Item Response Theory* (IRT). Perubahan dalam komponen CAT, seperti pengendalian nomor butir yang ditampilkan, keseimbangan isi, panjang test, dan ukuran kelompok butir tidak mampu dibedakan. Berdasarkan hasil penelitian yang ditampilkan dalam bentuk grafik diketahui perolehan skor responden lebih tinggi pada tes yang menggunakan kertas dan pena dari pada tes adaptif yang menggunakan komputer (Wang, 2001).

Pitkin & Vispoel mengumpulkan hasil penelitian dalam topik perbedaan antara CAT dan SAT melalui sebuah penelitian meta-analisis. Hasil penelitian menemukan SAT (*Scholastic Aptitude Test*) lebih peka digunakan dalam membuat keputusan bagi pengembang tes. SAT memiliki efek yang lebih baik sebab pilihan butir yang sulit dalam ujian dapat diberi umpan balik. Sedangkan CAT memiliki hasil yang cukup stabil. Reliabilitas CAT tinggi sebab seleksi butir CAT menggunakan algoritma yang mempunyai tipe pilihan butir yang dapat menurunkan kesalahan pengukuran secara maksimal. Antara SAT dan CAT sama-sama mempunyai kelemahan yaitu *overestimate* apabila butir sangat mudah dikerjakan, dan *underestimate* apabila butir yang diujikan terlalu sulit (Pitkin & Vispoel, 2001).

Masih dalam konteks penyelenggaraan tes, bentuk tes yang sama dapat dilaksanakan dalam berbagai kondisi dan persyaratan. *Booklet* tes dapat diaplikasikan dalam kondisi tes tradisional menggunakan *paper and pencil* (PP) dan dapat pula diterapkan dalam *software* komputer. Tes PP sering menuntut siswa menjawab mendekati 70% butir secara benar agar lulus ujian. Panjang tes yang memadai dapat menjamin kualitas tes. Tes

yang memiliki jumlah butir lebih panjang dapat menurunkan kesalahan pengukuran dan meningkatkan reliabilitas tes menjadi lebih tinggi (Lunz & Bergstrom, 1994).

Estimasi kemampuan yang dihasilkan dari program statistik selalu mengandung kesalahan. Hasil penelitian yang melacak pengaruh *Omits* (isian yang dihapus) terhadap estimasi kemampuan. Ketika *omit* dicoba diisi sebagai butir salah menampakkan indikasi bahwa kemampuan estimasi lebih akurat daripada ketika dianggap sebagai jawaban benar, walaupun hubungan linier antara teta dan teta estimasi tidak dipengaruhi khusus oleh omit yang dicoba sebagai isian yang salah atau benar (Ayala, Plake., & Impara, 2001). Jawaban yang dihapus menunjukkan peserta tes masih ragu-ragu terhadap pilihan jawabannya. Jika peserta test merasa yakin pilihan jawabannya benar, maka tidak mungkin jawabannya dihapus.

3. *Composing the test items*

Tahap yang paling penting dalam pengembangan tes adalah menyusun butir-butir soal tes. Dalam penulisan butir tes, pengembang tes perlu menyusun kisi-kisi. Di dalam kisi-kisi terdapat informasi tentang identitas dan matriks/format isi materi yang diujikan. Di dalam identitas memberi informasi: (1) Nama institusi/lembaga; (2) Kompetensi Keahlian/Mata Pelajaran/Standar Kompetensi; (3) Alokasi waktu. Di dalam matriks/format isi, antara lain berisi:

- 1) Standar Kompetensi (SK)
- 2) Kompetensi Dasar (KD)
- 3) Indikator kinerja/pencapaian kompetensi
- 4) Bentuk soal
- 5) Nomor urut soal

Kisi-kisi berfungsi sebagai petunjuk teknis dalam penulisan butir soal. Dengan adanya petunjuk teknis ini, penyusun soal akan dapat menghasilkan butir-soal yang sesuai dengan tujuan penilaian. Jika tersedia sebuah kisi-kisi yang baik, maka pengembang soal yang berbeda pun akan dapat menghasilkan perangkat soal yang relatif sama, baik dari tingkat kedalaman maupun cakupan materi yang diukur (ditanyakan).

Kriteria SK/KD yang baik untuk dipilih adalah sebagai berikut :

- (1) Materi yang penting dan mutlak harus dikuasai oleh peserta tes.
- (2) Materi yang berkelanjutan dan menjadi persyaratan untuk belajar pada kompetensi berikutnya.
- (3) Materi yang relevan dengan kompetensi kerja atau kehidupan sehari-hari.

Kriteria indikator soal yang baik adalah :

- (1) Mencerminkan pencapaian SK/KD yang hendak diukur.
- (2) Memuat satu kata kerja operasional yang dapat diukur/diamati.
- (3) Kata kerja operasional adalah kata kerja yang dapat diubah menjadi pertanyaan/perintah yang menuntut peserta menunjukkan kompetensi yang diminta pada soal.

Contoh kisi-kisi soal:

Nama Sekolah/lembaga : Sekolah Menengah Kejuruan

Kompetensi Keahlian :

Alokasi Waktu :

Jumlah Soal :

No	Kompetensi Dasar	Indikator Pencapaian Kompetensi	Bentuk soal	No butir
1	Menjelaskan menu seimbang	Siswa dapat menganalisis unsur-unsur zat gizi yang terkandung dalam susunan menu	Pilihan ganda	1-10
		Siswa dapat memilih bahan makanan yang mengandung zat tenaga, pembangun dan pengatur		
2	Menyusun menu untuk remaja	Siswa dapat menyusun menu seimbang untuk remaja	Uraian terbatas	11

Setelah butir-butir tes disusun, perancang tes kemudian menelaah kualitas butir secara teoritis/kualitatif dengan kriteria penilaian seperti pada Tabel 26. Analisis butir secara kualitatif dapat melibatkan beberapa orang ahli (*expert judgement*) yaitu ahli materi dari bidang studi yang

sama, ahli pengukuran dan ahli bahasa. Ahli materi berkewajiban mereviu substansi keilmuan soal tes, yaitu tes sudah mengukur semua indikator atau semua sub kompetensi. Ahli materi juga mempertimbangkan tingkat kematangan berpikir sasaran peserta tes. Misalnya tes matematika pada sub materi perkalian bilangan, materi tes yang digunakan untuk siswa SD lebih sederhana yaitu menggunakan bilangan di bawah 100 sedangkan untuk siswa SMA, materi perkalian dapat menggunakan bilangan lebih dari 100. Dengan demikian, domain tes yang diukur sama tetapi tingkat kesulitannya berbeda. Ahli konstruksi tes berkewajiban mengkoreksi teknik penulisan soal yang benar. Ahli bahasa antara lain mengkoreksi kejelasan hal yang ditanyakan, penggunaan bahasa baku, dan struktur kalimat mudah dipahami.

Pusat Pengembangan Sistem Pengujian (Pusbangsisjian:1996) Jakarta menyusun kartu telaah soal pilihan ganda. Ada tiga komponen (materi, konstruksi, dan bahasa) yang diuraikan menjadi 13 butir penelaahan soal seperti tercantum pada Tabel 4.1.

Tabel 4.1 Kartu Telaah Soal Pilihan Ganda

NO	KRITERIA PENILAIAN
A	MATERI
1	Soal sesuai dengan indikator.
2	Pengecoh berfungsi.
3	Mempunyai satu jawaban yang benar atau paling benar.
B	KONSTRUKSI
4	Pokok soal dirumuskan secara jelas dan tegas.
5	Rumusan soal dan rumusan jawaban hanya merupakan pernyataan yang diperlukan saja.
6	Pokok soal tidak menunjuk ke arah jawaban yang benar.
7	Pokok soal tidak mengandung pernyataan negatif ganda.
8	Pilihan jawaban homogen dan logis ditinjau dari sisi materi.
9	Panjang rumusan jawaban relatif sama.
10	Pilihan jawaban tidak mengandung pernyataan "semua jawaban di atas benar atau semua jawaban di atas salah".

NO	KRITERIA PENILAIAN
11	Pilihan jawaban yang berbentuk angka disusun berdasarkan urutan, sedangkan pilihan jawaban yang berbentuk waktu kejadian disusun secara kronologis.
12	Grafik, gambar, tabel dan diagram yang terdapat pada soal jelas dan berfungsi.
13	Butir soal tidak tergantung pada jawaban 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/frasa yang bukan merupakan satu kesatuan pengertian.

Contoh penulisan soal:

1. Soal sesuai dengan indikator:

Indikator: Siswa dapat menganalisis unsur-unsur zat gizi yang terkandung dalam susunan menu

KURANG BAIK	BAIK
Salah satu menu prasmanan pada pesta pernikahan antara lain menyajikan sambal goreng kentang dan krecek. Sambal goreng kentang menggunakan bumbu dasar A. merah B. putih C. kuning D. orange E. hijau	Sambal goreng kentang dan krecek menjadi salah satu hidangan dalam menu prasmanan di pesta pernikahan. Sambal goreng kentang banyak mengandung zat gizi A. protein dan lemak B. vitamin dan lemak C. mineral dan lemak D. karbohidrat dan lemak E. vitamin larut dalam lemak
Keterangan: Soal tidak sesuai indikator, materi tentang zat gizi tetapi soal menanyakan pengetahuan Boga Dasar.	

2. Pengecoh berfungsi/pilihan jawaban homogen dan logis

Semua pilihan jawaban harus berasal dari materi yang sama, pilihan jawaban memiliki kemiripan sehingga diperkirakan peserta ujian yang tidak menguasai materi dengan benar akan terkecoh memilih jawaban yang mirip tetapi semua pilihan jawaban harus logis/masuk akal.

KURANG BAIK	BAIK
<p>Merebus telur menggunakan teknik pengolahan panas basah (<i>moist heat cooking</i>) yaitu:</p> <p>A. Boilling B. Baking C. Boling D. Bowling E. Blender</p>	<p>Merebus telur menggunakan teknik pengolahan panas basah (<i>moist heat cooking</i>) yaitu:</p> <p>A. Boilling B. Stewing C. Simmering D. Poaching E. Blanching</p>
<p>Keterangan: Pilihan jawaban C dan D adalah cabang olahraga dan E alat memasak</p>	

3. *Memiliki satu kunci jawaban yang benar*

Setiap soal harus mempunyai satu jawaban yang benar atau yang paling benar menurut referensi yang tersedia bukan menurut keyakinan pembuat soal.

KURANG BAIK	BAIK
<p>Berikut ini yang termasuk teknik pengolahan panas kering (<i>dry heat cooking</i>) adalah:</p> <p>A. Grilling B. Roasting C. Shallow frying D. Stir frying E. Deep frying</p>	<p>Teknik pengolahan makanan yang digunakan pada saat memasak kentang goreng adalah yaitu:</p> <p>A. Baking B. Stewing C. Shallow frying D. Stir frying E. Deep frying</p>
<p>Keterangan: Pilihan jawaban A dan B benar</p>	
<p>Sekolah memiliki program kerja yang baru. Anda diberi tugas kepala sekolah untuk melakukan sosialisasi kepada semua warga sekolah. Apa yang akan anda lakukan?</p> <p>A. Memberikan leaflet untuk dibagikan ke warga sekolah B. Memasang benner pada sudut sekolah yang strategis C. Meminta kepala sekolah untuk menjelaskan program tersebut pada saat upacara D. Menyampaikan program kegiatan saat rapat guru E. Mengundang seluruh warga sekolah untuk berdiskusi bersama mengenai pelaksanaan program</p> <p>Keterangan: semua jawaban bisa benar menurut keyakinan masing-masing Sumber: soal tryout UTN PPG SM3T tahun 2017</p>	

4. *Pokok soal dirumuskan secara singkat, jelas, dan tegas*

Rumusan soal jelas, menggunakan kalimat tunggal dan tidak berbelit-belit.

KURANG BAIK	REVISI
Sekolah anda berada di lingkungan yang memiliki status sosial ekonomi bawah. Dalam rangka tahun pendidikan, kepala sekolah memiliki program kegiatan berbagi kepada masyarakat lingkungan sekitar, sebagai guru yang akan anda lakukan adalah...	Kegiatan sosial yang tepat dilakukan untuk membantu warga miskin di sekitar sekolah adalah
Peserta didik adalah salah satu peserta didik yang memiliki hasil belajar yang kurang baik sehingga dikhawatirkan tidak bisa masuk sekolah lanjutan yang diinginkan . Wali murid peserta didik C menghubungi anda untuk memberikan tambahan nilai pada peserta didik C dengan imbalan yang berjumlah besar. Apa yang anda lakukan.	Bagaimana sikap saudara jika ada orangtua siswa yang akan memberi gratifikasi supaya anaknya mendapat nilai lebih baik?
<p>Sumber: Soal try out PPG SM3T</p> <p>Keterangan: pokok soal berisi kalimat yang terlalu panjang dan kurang tegas jawaban yang diharapkan.</p>	

5. *Rumusan pokok soal dan pilihan jawaban harus merupakan pernyataan yang diperlukan saja*

KURANG BAIK	BAIK
Anda dinilai memiliki komitmen dalam menjalankan tugas sehingga kepala sekolah meng- gaskan anda sebagai ketua dalam penerimaan peserta didik baru. Ternyata jumlah peserta didik yang mendaftar jauh dari kuota yang di tetapkan sekolah. Apa yang anda lakukan untuk dapat me- penuhi kuota peserta didik yang telah ditentukan sekolah?	Tindakan apa yang tepat dilakukan oleh Kepala Sekolah jika jumlah pendaftar lebih sedikit dari kuota yang di tetapkan?
<p><i>Penjelasan:</i> Pokok soal di atas mengandung pernyataan yang tidak diperlukan sehingga menyita waktu untuk membaca dan memahami maksud soal.</p> <p>Sumber: Soal try out PPG SM3T tahun 2017</p>	

6. *Pokok soal tidak memberi petunjuk ke arah kunci jawaban*

KURANG BAIK	BAIK
<p>Salah satu nama kue yang terbuat dari beras ketan adalah:</p> <p>A. Ketan srikaya B. Klepon C. Mendut D. Nagasari E. Lapis</p>	<p>Salah satu nama kue yang terbuat dari beras ketan adalah:</p> <p>A. Wajik B. Klepon C. Mendut D. Nagasari E. Lapis</p>
<p>Keterangan: Ketan sudah ditulis pada pokok soal dan menjadi alternatif jawaban yang benar.</p>	

7. *Pokok soal tidak mengandung pernyataan yang bersifat negatif ganda*

KURANG BAIK	BAIK
<p>Vitamin yang tidak larut di dalam lemak tidak harus dikonsumsi setiap hari kecuali:</p> <p>A. Vitamin A B. Vitamin C C. Vitamin D D. Vitamin E E. Vitamin K</p>	<p>Vitamin larut di dalam air harus dikonsumsi setiap hari adalah:</p> <p>A. Vitamin A B. Vitamin C C. Vitamin D D. Vitamin E E. Vitamin K</p>
<p>Keterangan: Menggunakan kata tidak dua kali, dan kecuali yang sama-sama memiliki makna negatif</p>	

8. *Panjang rumusan pilihan jawaban relatif sama*

KURANG BAIK
<p>Saat baru sampai disekolah anda mendapat kabar bahwa putra anda mengalami kecelakaan padahal hari itu waktunya ulangan harian yang sudah dijadwalkan. Sebagai guru tindakan anda adalah...</p> <p>A. Memberi tugas dan meminta mengumpulkan hari itu juga B. Berupaya masuk kelas sebentar memberi pengantar dan meminta tolong teman guru yang sama-sama mampu mata pelajaran untuk menjaga ulangan harian C. Meminta guru piket untuk mengisi dengan materi pelajaran yang lain D. Menunda ulangan harian E. Meminta teman guru menjaga ulangan agar tetap berlangsung</p>

Sumber: Soal try out PPG SM3T

Keterangan:

Ada kecenderungan jawaban yang paling panjang dan lengkap merupakan kunci jawaban sehingga banyak dipilih peserta tes

9. *Pilihan jawaban tidak mengandung pernyataan, “Semua pilihan jawaban di atas salah”, atau “Semua pilihan jawaban di atas benar, dan sejenisnya*

KURANG BAIK	BAIK
<p>Vitamin yang larut di dalam lemak tidak harus dikonsumsi setiap hari yaitu:</p> <p>A. Vitamin A B. Vitamin D C. Vitamin E D. Vitamin K E. Semua jawaban benar</p>	<p>Vitamin yang larut di dalam lemak tidak harus dikonsumsi setiap hari yaitu:</p> <p>(1) Vitamin A (2) Vitamin B (3) Vitamin C (4) Vitamin D</p> <p>Jawaban:</p> <p>A. (1) dan (2) benar B. (1) dan (3) benar C. (1) dan (4) benar D. (2) dan (3) benar E. (2) dan (4) benar</p>

10. *Pilihan jawaban yang berbentuk angka disusun berdasarkan urutan besar ke kecil atau sebaliknya, demikian pula jawaban kronologi waktu juga disusun berurutan dari awal sampai akhir.*

KURANG BAIK	BAIK
<p>Satu gelas sirup buah naga membutuhkan bahan 150 cc air dan 2 sdt gula. Berapa jumlah kebutuhan air dan gula untuk membuat 10 gelas minuman sirup buah naga jika 1 sdt gula sama dengan 10 gram?</p> <p>A. 1,5 lt air dan 200 gram gula B. 1,5 lt air dan 100 gram gula C. 150 cc air dan 10 gram gula D. 150 cc air dan 20 gram gula E. 1500 cc air dan 100 gram gula</p>	<p>Satu gelas sirup buah naga membutuhkan bahan 150 cc air dan 2 sdt gula. Berapa jumlah kebutuhan air dan gula untuk membuat 10 gelas minuman sirup buah naga jika 1 sdt gula sama dengan 10 gram?</p> <p>A. 150 cc air dan 10 gram gula B. 150 cc air dan 20 gram gula C. 1500 cc air dan 100 gram gula D. 1,5 lt air dan 200 gram gula E. 1,5 lt air dan 400 gram gula</p>

11. *Gambar, grafik, tabel, diagram, dan sejenisnya jelas dan berfungsi.*

Gambar dan sejenisnya yang menyertai suatu butir soal harus jelas, terbaca, dan dapat dimengerti. Apabila soal tersebut tetap bisa dijawab tanpa melihat gambar, grafik, tabel atau sejenisnya yang terdapat pada soal, berarti gambar, grafik, atau tabel tersebut tidak berfungsi

KURANG BAIK	BAIK
	
<p>Nama potongan sayur pada gambar di atas adalah</p> <p>A. Slice B. Paysanne C. Macedoine D. Brunoise , E. Matignone</p>	<p>Nama potongan sayur pada gambar di atas adalah</p> <p>A. Slice B. Paysanne C. Macedoine D. Brunoise , E. Matignone</p>
<p>Gambar kurang jelas</p>	

12. *Butir soal tidak bergantung pada jawaban soal sebelumnya.*

Soal disusun secara independen dan unidimensi karena siswa yang tidak dapat menjawab soal sebelumnya juga tidak akan bisa menjawab soal berikutnya

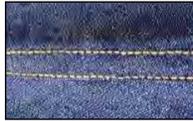
BAIK	KURANG BAIK
<p>1. Sauce turunan yang menggunakan mother sauce mayonaise adalah:</p> <p>A. Béarnaise sauce B. Mouseline sauce C. Morney sauce D. Andalouse sauce E. Bourdelaisesauc</p>	

<p>a. Bahan dasar sauce pada jawaban nomor 1 adalah:</p> <p>A. Egg yolk, salad oil, vinegar/lime juice, mustard, tomato paste, capsicum</p> <p>B. Butter, chopped onion, garlic, chopped parsley, anchovy essence, S/P.</p> <p>C. Salad oil, french mustard, egg yolk, lime juice, vinegar, S/P, L/P.</p> <p>D. Vinegar, lime juice, salad oil, chopped onion, chopped garlic, S/P, L/P</p> <p>E. Red chili, brown sugar, vinegar, dried shrimp, salt, garlic, shrimp paste, hot water</p>
<p><i>Keterangan:</i> soal nomor 2 sebaiknya disusun secara independen, tidak bergantung pada jawaban soal nomor 1</p>

13. *Menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia*

KURANG BAIK	BAIK
<p>Yang menjadi penyebab terjadinya pencoklatan non enzimatis karena ada reaksi antara gula pereduksi dengan asam aminoseperti pada proses pemanggangan roti adalah</p> <p>A. Browning</p> <p>B. Maillard</p> <p>C. Karamelisasi</p> <p>D. Case hardening</p> <p>E. Fenolase</p>	<p>Reaksi pencoklatan non enzimatis yang terjadi antara gula pereduksi dengan aminoseperti pada saat pemanggangan roti adalah</p> <p>A. Browning</p> <p>B. Maillard</p> <p>C. Karamelisasi</p> <p>D. Case hardening</p> <p>E. Fenolase</p>
<p><i>Keterangan:</i> Pokok soal 1 tidak menggunakan struktur kalimat baku yaitu SPOK, kata "yang" merupakan kata sambung tidak layak menjadi pokok kalimat</p>	

14. *Menggunakan bahasa yang kurang komunikatif dan efektif*

KURANG BAIK	BAIK
<p>Kampuh yang mempunyai bekas jahitan pada satu sisi sebanyak dua setikan dan sisi yang sebelahnya satu setikan disebut</p> <p>A. Kampuh sarung</p> <p>B. Kampuh balik</p> <p>C. Kampuh pipih</p> <p>D. Kampuh buka</p> <p>E. Kampuh yang di rubah</p>	<div style="display: flex; flex-wrap: wrap;"> <div style="width: 50%; text-align: center;">  <p>Kampuh Balik <input type="checkbox"/></p> </div> <div style="width: 50%; text-align: center;">  <p>Kampuh Pipih <input type="checkbox"/></p> </div> <div style="width: 50%; text-align: center;">  <p>Kampuh Perancis <input type="checkbox"/></p> </div> <div style="width: 50%; text-align: center;">  <p>Kampuh Sarung <input type="checkbox"/></p> </div> </div>

	<p>Gambar jahitan ini menunjukkan bentuk kampuh ...</p> <p>A. Sarung B. balik C. pipih D. buka E. tutup</p>
<p>Tusuk dasar yang hasilnya sama seperti menjahit disebut....</p> <p>A. Tusuk jelujur B. Tusuk feston C. Tusuk tikam jejak D. Tusuk rantai E. Tusuk flanel</p>	<div data-bbox="575 395 1018 690" data-label="Image"> </div> <p>Tusuk dasar menjahit pada gambar ini adalah tusuk</p> <p>A. jelujur B. feston C. tikam jejak D. rantai E. flanel</p>
<p>Bahasa kurang efektif dan sulit dipahami, istilah lebih mudah dijelaskan dengan gambar</p>	

15. *Tidak menggunakan bahasa yang berlaku setempat (bias budaya)*

Penggunaan bahasa yang tidak berlaku secara nasional akan mengakibatkan kesalahan/perbedaan penafsiran.

KURANG BAIK	BAIK
<p>Buah yang banyak mengandung vitamin A dan tinggi serat sehingga dapat melancarkan BAB adalah</p> <p>A. kates B. pelem C. nanas D. sirsak E. duren</p>	<p>Buah yang banyak mengandung vitamin A dan tinggi serat sehingga dapat melancarkan buang air besar adalah</p> <p>A. pepaya B. mangga C. nenas D. sirsak E. durian</p>
<p>Keterangan: Jawaban banyak menggunakan bahasa Jawa.</p>	

16. *Pilihan jawaban tidak mengulang kata/kelompok kata yang bukan merupakan satu kesatuan pengertian*

KURANG BAIK	<p>Prosedur pembuatan adonan sponge cake yang tepat adalah:</p> <ol style="list-style-type: none"> a. (1) kocok putih telur sampai berbusa dan kaku, sisihkan, (2) kocok margarin dan gula sampai putih kekuningan; (3) Masukkan kuning telur satu demi satu sambil dikocok terus sampai rata; (4) Masukkan terigu dan putih telur yang sudah dikocok, aduk sampai rata b. (1) kocok telur, gula dan ovalet sampai kental, (2) tambahkan tepung terigu, aduk sampai rata. (3) tambahkan margarin cair, aduk sampai rata c. (1) kocok margarin bersama gula pasir sampai berwarna putih; (2) masukkan satu butir telur dan sebagian tepung sambil terus dikocok; (3) ulangi prosedur ke-2, sampai telur dan terigu habis, d. (1) campur kuning telur, minyak, gula, dan terigu sisihkan; (2) kocok putih dan cream or tartar sampai mengembang; (3) sambil terus dikocok, tambahkan gula pasir sedikit demi sedikit; (4) campur bagian (1) dan (3) menjadi satu e. (1) kocok telur sampai mengental, sisihkan (2) kocok margarin dan gula sampai pucat; (3) Masukkan adonan telur dan tepung terigu ke dalam adonan margarine sedikit demi sedikit, sambil dikocok terus sampai rata;
BAIK	<p>Urutkan prosedur pembuatan sponge cake yang tepat:</p> <ol style="list-style-type: none"> (1) Kocok putih telur sampai berbusa dan kaku, sisihkan, (2) kocok margarin dan gula sampai pucat dan mengembang; (3) tambahkan tepung terigu, aduk sampai rata. (4) tambahkan margarin cair, aduk sampai rata (5) Masukkan kuning telur satu demi satu sambil di kocok terus sampai rata; (6) Masukkan putih telur, aduk sampai rata (7) Kocok telur, gula dan ovalet sampai kental, <ol style="list-style-type: none"> A. (7), (4) dan (3) B. (7), (3), dan (4) C. (2), (5) dan (6) D. (1), (2), (5) dan (6) E. (2), (3) dan (5)
<p>Jawaban terlalu panjang mengungkap prosedur yang hampir sama sehingga menyita banyak waktu untuk membaca</p>	

4. *Writing the administration instructions*

Setelah butir-butir tes tertulis dan mewakili semua kisi-kisi, pengembang tes selanjutnya menulis petunjuk penyelenggaraan. Petunjuk yang harus disiapkan oleh pengembang tes minimal ada dua yaitu: petunjuk untuk penyelenggara dan pengawas ujian serta petunjuk untuk peserta tes itu sendiri. Petunjuk untuk penyelenggara antara lain berisi kondisi lingkungan yang harus disiapkan oleh penyelenggara pada saat tes dilaksanakan. Kondisi yang diharapkan antara lain:

- 1) Tes dilaksanakan kelompok atau individu.
- 2) Kondisi-kondisi khusus yang dituntut seperti: kerahasiaan soal, penataan ruang, kenyamanan kursi dan meja.
- 3) Tuntutan perlengkapan yang harus disediakan seperti lembar jawaban, kertas folio, pensil, alas tulis, program komputer atau aplikasi ponsel.
- 4) Batas waktu yang sesuai untuk menyelesaikan soal yang bisa diramalkan dari tingkat kesulitan butir soal. Soal yang membutuhkan pemecahan masalah, menghitung dengan rumus membutuhkan waktu lebih lama untuk mengerjakan.
- 5) Teks yang harus dibaca oleh pengawas kepada peserta tes termasuk jawaban-jawaban yang dapat diberikan apabila ada pertanyaan dari peserta tes. Pengawas pada umumnya perlu menyampaikan tata tertib peserta selama mengikuti ujian.
- 6) Petunjuk untuk peserta tes dapat dibacakan oleh pengawas atau ditulis pada lembar soal. Petunjuk ini meliputi petunjuk penulisan identitas peserta tes dan petunjuk umum cara mengerjakan soal tes, petunjuk khusus untuk mengerjakan tiap-tiap bagian soal tes, dsb.

5. *Conduct Piloting Test*

Apabila perangkat tes yang dikembangkan merupakan tes baru, maka pengembangan tes tidak selalu menunjukkan hasil seperti yang diharapkan. Oleh sebab itu, perangkat tes baru harus diuji cobakan pada skala kecil (*piloting test*). Langkah ini dilakukan untuk menyediakan data empiris yang digunakan untuk analisis kualitas butir tes secara empiris yaitu dari tingkat kesulitan, daya pembeda, reliabilitas dan validitas tes. Subjek yang menjadi sasaran uji coba tes harus memiliki karakteristik yang sama dengan sasaran tes yang sebenarnya. Jika soal tes akan digunakan untuk ujian nasional pada siswa kelas IX SMP, maka sasaran uji cobanya juga siswa kelas IX SMP

dari berbagai wilayah supaya hasil uji coba mencerminkan karakteristik kemampuan siswa yang sebenarnya akan diuji.

6. *Conduct item analysis*

Pengembangan tes pada tahap ini adalah melakukan uji coba tes pada sampel yang sesuai dengan sasaran sesungguhnya. Respon butir hasil uji coba tes digunakan untuk membuktikan kualitas butir secara kuantitatif berdasarkan data empiris. Analisis data respon butir dapat dilakukan dengan bantuan program ITEMAN, ANATES, dan IRT (*item response theory*), dll. Pada umumnya, analisis butir dilakukan untuk mengetahui tingkat kesulitan, daya pembeda, korelasi point biserial, pengecoh, dan reliabilitas tes. Berikut ini diberikan contoh cara menganalisis respon butir secara manual.

a) *Tingkat kesulitan Butir*

Tingkat kesulitan butir (*item difficulty*) adalah angka yang menunjukkan besarnya proporsi peserta tes yang menjawab benar pada suatu butir. Tingkat kesulitan butir merentang mulai dari 0,00 sampai 1,00. Jika suatu butir soal memiliki tingkat kesukaran 0,00 berarti tidak ada peserta tes yang menjawab butir soal tersebut dengan benar. Dengan kata lain butir soal tersebut terlalu sulit, tidak terdapat jawaban yang benar atau memiliki jawaban benar lebih dari satu. Sebaliknya, jika butir soal memiliki tingkat kesulitan 1,00 berarti semua peserta tes dapat menjawab butir soal dengan benar. Dengan kata lain, butir soal tersebut terlalu mudah. Cara menghitung tingkat kesulitan butir dapat menggunakan rumus sebagai berikut:

$$\text{Tingkat kesulitan} = \frac{\text{jumlah peserta tes yang menjawab benar}}{\text{jumlah seluruh kelompok tes}}$$

$$\text{atau } p = \frac{\sum B}{N}$$

Berikut ini terdapat contoh tabel untuk menganalisis butir dari unsur tingkat kesulitan dan korelasi point biserial yang akan dijelaskan pada bagian berikutnya. Jawaban benar mendapat nilai 1 dan salah 0

Tabel 4.1 Contoh Input Data Respon Butir untuk Menghitung Tingkat Kesulitan dan Korelasi Point Biserial

No	Nama	1	2	3	4	5	6	7	8	9	10	ΣX	ΣX^2
1	A	1	1	1	1	1	1	1	0	1	1	9	81
2	B	1	1	1	1	1	1	1	0	1	1	9	81
3	C	1	0	1	1	1	1	1	0	1	1	8	64
4	D	1	1	0	1	0	0	0	0	1	1	5	25
5	E	1	0	1	0	0	1	1	0	1	1	6	36
6	F	1	0	1	0	0	0	1	0	0	1	4	16
7	G	1	1	1	0	0	1	0	0	0	1	5	25
8	H	1	1	1	1	1	1	1	0	1	1	9	81
9	I	1	0	0	0	1	1	0	0	0	0	3	9
10	J	1	1	0	1	1	1	1	0	1	1	8	64
11	K	1	1	1	1	1	1	0	0	0	0	6	36
12	L	1	1	1	1	1	1	1	0	1	1	9	81
13	M	1	0	1	1	1	1	1	0	1	0	7	49
14	N	1	1	0	1	0	1	0	0	0	1	5	25
15	O	1	0	1	0	0	1	1	0	0	0	4	16
16	P	1	0	0	0	1	1	1	0	0	1	5	25
17	Q	1	1	1	1	1	1	1	0	1	1	9	81
18	R	1	0	1	1	1	1	1	0	1	1	8	64
19	S	1	1	0	0	1	1	0	0	0	0	4	16
20	T	1	1	0	1	0	1	0	0	1	0	5	25
Σ		20	12	13	13	13	18	13	0	12	14	128	900
p		1	0.6	0.65	0.65	0.65	0.9	0.65	0	0.6	0.7		
r_{pb}		0	0.31	0.51	0.72	0.6	0.3	0.62	0	0.77	0.51		

Keterangan: p = Indeks kesulitan butir

R_{pb} : Korelasi point biserial

Tingkat kesulitan butir dikategorikan menjadi 3 yaitu sulit, sedang, dan mudah. Kriteria yang digunakan untuk menginterpretasikan hasil analisis tingkat kesulitan butir mengacu pada Tabel 4.1.

Tabel 4.2 Kriteria Indeks Tingkat Kesulitan Butir (p)

Proporsi benar	Kategori
$p > 0,7$	Mudah
$0,3 \leq p \leq 0,7$	Sedang
$p < 0.3$	Sulit

Sumber: Bahrul Hayat (1997: 18)

Contoh hasil analisis butir soal nomor 2 yaitu terdapat 12 peserta tes yang menjawab benar dari 20 peserta tes yang ikut, maka Indeks Kesulitan (p) butirnya adalah:

$$p = \frac{12}{20} = 0,6$$

Hasil analisis pada tabel 4.2 menunjukkan ada dua butir (nomor 1 dan 6) termasuk kategori butir yang mudah dengan indeks kesulitan butir (p) $> 0,7$ sedangkan butir nomor 8 termasuk butir yang paling sulit karena tidak ada satupun peserta didik yang menjawab benar $p = 0$. Respon butir yang sama (salah semua atau benar semua) tidak dapat dianalisis korelasi. Soal yang dijawab salah oleh semua peserta tes perlu dicek kembali karena kesalahan kemungkinan besar dilakukan oleh pembuat soal. Analisis penyebab kesalahan antara lain pilihan jawaban benar lebih dari satu atau tidak ada jawaban yang benar.

Tingkat kesulitan suatu butir soal dipengaruhi oleh tingkat kemampuan dari anggota kelompok peserta tes. Butir soal yang sama ketika diujikan pada kelas yang beranggotakan siswa pandai seperti kelas akselerasi dapat direspon mudah tetapi ketika diujikan ke kelas yang kurang pandai dapat direspon sulit. Hal ini berarti bahwa tingkat kesulitan butir soal tidak semata-mata menunjukkan ukuran kesulitan butir soal, tetapi juga menunjukkan kemampuan rata-rata peserta tes.

b) Daya Pembeda Butir

Daya pembeda butir (**D**) soal berfungsi untuk melihat kemampuan butir soal dalam membedakan antara peserta tes yang berkemampuan tinggi dengan peserta tes yang berkemampuan rendah. Untuk menganalisis

daya pembeda dilakukan dengan membandingkan kinerja kelompok atas, atau kelompok yang memperoleh skor tes sangat tinggi (*Upper*) dengan kinerja kelompok bawah atau kelompok yang memperoleh skor tes sangat rendah (*Lower*) pada masing-masing butir. Untuk menghitung indeks daya pembeda butir (*a discrimination index*) digunakan rumus sebagai berikut:

$$Upper = \frac{\text{jumlah kelompok atas yang menjawab benar}}{\text{jumlah seluruh kelompok atas}}$$

$$Lower = \frac{\text{jumlah kelompok bawah yang menjawab benar}}{\text{jumlah seluruh kelompok bawah}}$$

$$D = Upper - Lower$$

Kelompok atas dan kelompok bawah dibentuk berdasarkan skor akhir tes. McIntire (McIntire & Miller, 2000) menyarankan untuk mengambil sekitar 25% s/d 35% ranking atas menjadi kelompok atas dan sebaliknya 25% s/d 35% ranking bawah menjadi kelompok bawah. Contoh input data analisis respon butir soal pada table 27 dapat ditindak lanjuti untuk menganalisis daya pembeda butir dengan cara mengurutkan (meranking) skor total peserta tes. Dalam contoh terdapat 20 orang peserta tes sehingga kelompok atas ditetapkan pada ranking 1 – 5 (25% X 20) dan kelompok bawah ditetapkan pada ranking 15 – 20. Berikut ini contoh table pembagian kelompok atas dan bawah untuk menghitung indeks daya pembeda butir.

Tabel 4.3 Contoh Input Data Respon Butir Untuk Menghitung Daya Pembeda

No	Nama	1	2	3	4	5	6	7	8	9	10	Ranking
1	A	1	1	1	1	1	1	1	0	1	1	9
2	B	1	1	1	1	1	1	1	0	1	1	9
8	H	1	1	1	1	1	1	1	0	1	1	9
12	L	1	1	1	1	1	1	1	0	1	1	9
17	Q	1	1	1	1	1	1	1	0	1	1	9
3	C	1	0	1	1	1	1	1	0	1	1	8

No	Nama	1	2	3	4	5	6	7	8	9	10	Ranking
10	J	1	1	0	1	1	1	1	0	1	1	8
18	R	1	0	1	1	1	1	1	0	1	1	8
13	M	1	0	1	1	1	1	1	0	1	0	7
5	E	1	0	1	0	0	1	1	0	1	1	6
11	K	1	1	1	1	1	1	0	0	0	0	6
4	D	1	1	0	1	0	0	0	0	1	1	5
7	G	1	1	1	0	0	1	0	0	0	1	5
14	N	1	1	0	1	0	1	0	0	0	1	5
16	P	1	0	0	0	1	1	1	0	0	1	5
20	T	1	1	0	1	0	1	0	0	1	0	5
6	F	1	0	1	0	0	0	1	0	0	1	4
15	O	1	0	1	0	0	1	1	0	0	0	4
19	S	1	1	0	0	1	1	0	0	0	0	4
9	I	1	0	0	0	1	1	0	0	0	0	3
Atas		5	5	5	5	5	5	5	0	5	5	
Bawah		5	2	2	1	2	4	2	0	1	1	
Selisih		0	3	3	4	3	1	3	0	4	4	
Daya beda		0	0.6	0.6	0.8	0.6	0.2	0.6	0	0.8	0.8	

Contoh soal nomor 2 yaitu terdapat 5 orang kelompok atas menjawab benar dan 2 orang kelompok bawah menjawab benar, maka dapat dihitung indeks daya pembedanya (*discrimination index*) sebagai berikut:

$$Upper = \frac{5}{5};$$

$$Lower = \frac{2}{5};$$

$$D = \frac{5}{5} - \frac{2}{5} = \frac{3}{5};$$

$$D = 0,6$$

Daya beda butir dipengaruhi oleh kemampuan rata-rata kelompok peserta tes. Butir soal yang memiliki daya beda tinggi pada peserta tes kelas A tidak mungkin sama dengan dengan daya beda kelas B meskipun soal mengukur indikator yang sama. Soal yang terlalu mudah memiliki daya pembeda rendah. Soal yang sulit juga bisa memiliki daya pembeda rendah jika peserta tes kelompok atas menjawab salah dan peserta tes kelompok bawah justru menjawab benar karena menebak-nebak. Soal yang terlalu sulit ini bisa disebabkan karena jawaban benar hanya menurut pengalaman penulis soal tetapi tidak terdapat satu teoripun yang dapat dirujuk. Soal yang memiliki daya pembeda tinggi dapat diprediksi dari soal yang memiliki pilihan jawaban hampir sama sehingga jawaban pengecoh berfungsi, soal yang mengukur kemampuan berpikir tinggi yaitu pada tingkat aplikasi, analisis, dan evaluasi. Indikator untuk menilai kualitas butir dari unsur daya pembeda butir menurut Dali S. Naga (Naga, 1992) ditetapkan dengan kriteria pada Tabel 4.4.

Tabel 4.4 Kriteria Indeks Daya Pembeda Butir (D)

Indeks pembeda (D/r_{pb})	Kategori
$D \geq 0,4$	Sangat baik
$0,3 \leq D \leq 0,39$	Baik, tanpa revisi
$0,2 \leq D \leq 0,29$	Perbatasan atau perlu revisi
$D \leq 0,19$	Dibuang atau diganti

Daya pembeda butir juga dapat dianalisis menggunakan rumus korelasi *Point-Biserial* karena data dikotomi yaitu benar = 1 dan salah = 0. Koefisien korelasi (r_{pb}) diperoleh dengan mengorelasikan skor pada satu butir tes tertentu (X) terhadap skor seluruh butir tes (Y). Rumus korelasi Point-Biserial (Hinkle, Wiersma, & Jurs, 1979) yang digunakan adalah:

$$r_{pbis} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{\{ N \sum X^2 - (\sum X)^2 \} \{ N \sum Y^2 - (\sum Y)^2 \}}}$$

Soal nomor 3, 4, 5, dan 7 memiliki indeks kesulitan sama (0,65) tetapi tidak menunjukkan korelasi (r_{pb}) yang sama. Hal ini dipengaruhi posisi jawaban salah/benar. Jika kelompok atas bisa menjawab benar pada butir

soal yang sulit, maka daya pembeda naik dan korelasi butir juga naik. Jika kelompok atas menjawab salah sedangkan kelompok asor menjawab benar pada butir soal yang sulit, maka daya pembeda turun dan korelasi butir juga turun. Persiapan analisis korelasi tercantum pada tabel 30 di bawah ini. Skor butir diberi simbol X dan skor total butir diberi simbol Y.

Tabel 4.5 Contoh Tabel Persiapan Analisis Korelasi Point Biserial

No	SOAL NOMOR 3					SOAL NOMOR 4				
	X	X ²	Y	Y ²	XY	X	X ²	Y	Y ²	XY
1.	1	1	9	81	9	1	1	9	81	9
2.	1	1	9	81	9	1	1	9	81	9
3.	1	1	8	64	8	1	1	8	64	8
4.	0	0	5	25	0	1	1	5	25	5
5.	1	1	6	36	6	0	0	6	36	0
6.	1	1	4	16	4	0	0	4	16	0
7.	1	1	5	25	5	0	0	5	25	0
8.	1	1	9	81	9	1	1	9	81	9
9.	0	0	3	9	0	0	0	3	9	0
10.	0	0	8	64	0	1	1	8	64	8
11.	1	1	6	36	6	1	1	6	36	6
12.	1	1	9	81	9	1	1	9	81	9
13.	1	1	7	49	7	1	1	7	49	7
14.	0	0	5	25	0	1	1	5	25	5
15.	1	1	4	16	4	0	0	4	16	0
16.	0	0	5	25	0	0	0	5	25	0
17.	1	1	9	81	9	1	1	9	81	9
18.	1	1	8	64	8	1	1	8	64	8
19.	0	0	4	16	0	0	0	4	16	0
20.	0	0	5	25	0	1	1	5	25	5
Σ	13	13	128	900	93	13	13	128	900	97

Hasil penjumlahan skor butir tes yang terdapat pada Tabel 4.5 kemudian dimasukkan pada rumus korelasi. Langkah-langkah analisis korelasi poin biserial adalah sebagai berikut.

Soal nomor 3

Input data korelasi pada soal nomor 3

Skor butir nomor 3		Skor total seluruh butir		
X	X ²	Y	Y ²	XY
13	13	128	900	93

Jumlah peserta tes 20

$$r_{pbis} = \frac{20.93 - (13 \times 128)}{\sqrt{\{ 20.13 - (13)^2 \} \{ 20.900 - (128)^2 \}}}$$

$$r_{pbis} = \frac{1860 - (1664)}{\sqrt{\{ 260 - 169 \} \{ 18000 - 16384 \}}}$$

$$r_{pbis} = \frac{196}{\sqrt{(91)(1616)}} = \frac{196}{\sqrt{147056}} = \frac{196}{383,48} = 0,51$$

Input data korelasi pada soal nomor 4

Skor butir nomor 4		Skor total seluruh butir		
X	X ²	Y	Y ²	XY
13	13	128	900	97

$$r_{pbis} = \frac{20.97 - (13 \times 128)}{\sqrt{\{ 20.13 - (13)^2 \} \{ 20.900 - (128)^2 \}}}$$

$$r_{pbis} = \frac{1940 - (1664)}{\sqrt{\{ 260 - 169 \} \{ 18000 - 16384 \}}}$$

$$r_{pbis} = \frac{276}{\sqrt{(91)(1616)}} = \frac{276}{\sqrt{147056}} = \frac{276}{383,48} = 0,72$$

Hasil analisis tersebut kemudian dikonsultasikan dengan kriteria yang terdapat pada Tabel 4.6.

Tabel 4.6 Interpretasi Koefisien Korelasi

Koefisien Korelasi		Interpretasi
Positif	Negatif	
>0,9 s/d 1,0	-0,9 s/d -1,0	korelasi sangat tinggi
>0,7 s/d 0,9	-0,7 s/d -0,9	korelasi tinggi
>0,5 s/d 0,7	-0,5 s/d -0,7	korelasi sedang
>0,3 s/d 0,5	-0,3 s/d -0,5	korelasi rendah
0,0 s/d 0,3	-0,0 s/d -0,3	apabila ada korelasi maka sangat kecil

Sumber (Hinkle: 1979:85)

Hasil analisis butir dapat dimanfaatkan untuk dua kepentingan yaitu membuktikan daya pembeda dan validitas butir. Interpretasi daya pembeda menggunakan tabel 4.4 dan interpretasi validitas butir menggunakan tabel 4.6 Penggunaan hasil analisis korelasi butir soal sebagai alat pembuktian validitas butir masih menjadi perdebatan beberapa ahli pengukuran karena data dikotomi (1,0) sulit mendapat koefisien yang tinggi. Korelasi butir sering digunakan untuk mengukur validitas kuesioner yang memiliki data berskala interval (1, 2, 3, 4). Jika akan menggunakan rumus korelasi pada pembuktian validitas soal, sebaiknya menggunakan validitas konstruks dengan cara menggabungkan beberapa butir yang mengukur indikator sama sebagai X dan skor total menjadi Y. Hasil analisis korelasi butir nomor 3 memperoleh koefisien r_{pb} sebesar 0,51 yang menunjukkan ada korelasi sedang antara skor butir dengan skor totalnya. dan butir nomor 4 menunjukkan koefisien r_{pb} sebesar 0,72 menunjukkan ada korelasi tinggi.

Selain analisis butir secara kuantitatif seperti di atas, pengembang tes juga diharapkan menggali informasi kualitatif kepada peserta tes. Pengembang tes dapat menyiapkan pedoman wawancara dan observasi untuk mengetahui persepsi dan reaksi peserta tes setelah mengikuti atau menjawab butir-butir soal tes. Aspek yang ditanyakan seperti jumlah waktu yang disediakan, pemahaman terhadap butir-butir pertanyaan, dan isi tes secara umum.

7. *Revising the test*

Setelah dilakukan analisis butir tes, butir yang kurang baik kemudian diganti atau direvisi. Penentuan keputusan butir tersebut baik atau kurang baik dapat menggunakan kriteria analisis butir kuantitatif yaitu tingkat kesulitan (p), daya pembeda (D) dan korelasi (r_{pb}) butir. Apabila dua dari tiga kriteria butir tes yang baik dapat terpenuhi atau konsisten, maka butir tes tersebut dapat digunakan. Butir yang tidak dapat memenuhi kualitas butir yang baik perlu diganti atau direvisi. Contoh penentuan keputusan seleksi butir berdasarkan rangkuman hasil analisis butir tes dapat disimak pada Tabel 4.7.

Tabel 4.7 Contoh Rangkuman Hasil Analisis Butir

NO	p	r_{pb}	D	Keterangan
1.	1	#DIV/0!	0	Perlu revisi
2.	0.6	0.31	0.6	Tidak perlu revisi
3.	0.65	0.51	0.6	Tidak perlu revisi
4.	0.65	0.72	0.8	Tidak perlu revisi
5.	0.65	0.56	0.6	Tidak perlu revisi
6.	0.9	0.32	0.2	Perlu revisi
7.	0.65	0.62	0.6	Tidak perlu revisi
8.	0	#DIV/0!	0	Perlu revisi
9.	0.6	0.77	0.8	Tidak perlu revisi
10.	0.7	0.51	0.8	Tidak perlu revisi

Pada umumnya soal yang mudah akan memiliki daya pembeda butir yang rendah. Korelasi butir konsisten dengan daya pembeda, sehingga jika soal memiliki daya pembeda rendah secara otomatis korelasi butirnya juga rendah (tidak signifikan). Butir soal nomor 1, 6, dan 8 perlu direvisi atau dibuang karena terdapat korelasi yang sangat rendah dan daya pembeda yang rendah pula, meskipun tingkat kesulitan ada yang mudah dan ada yang sulit.

8. Validation the test

Setelah melakukan revisi, pengembang selanjutnya melakukan pengukuran validasi dan reliabilitas tes pada sasaran yang lebih luas. Validasi tes dapat dilakukan dengan cara mengorelasikan skor tes individu yang dikembangkan saat ini dengan skor tes individu pada tes yang pernah diikuti sebelumnya (*concurrent validity*). Reliabilitas hasil tes diukur dengan cara membelah skor tes menjadi dua bagian (ganjil dan genap) kemudian diolah menggunakan analisis korelasi. Rumus yang digunakan untuk mengukur validitas dan reliabilitas sama yaitu menggunakan analisis korelasi *product moment*.

$$r_{pbis} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{\{ N \sum X^2 - (\sum X)^2 \} \{ N \sum Y^2 - (\sum Y)^2 \}}}$$
 atau

$$r_{xy} = \frac{\sum xy}{\sqrt{(\sum x^2)(\sum y^2)}}$$

Keterangan:

Simbol	Concurrent Validity	Split half reliability
X	Skor tes baru yang dikembangkan	Ganjil
Y	Skor tes sebelumnya pada materi yang sama	Genap
rx _y	Koefisien validitas konkuren	Koefisien reliabilitas
N	Jumlah subjek	

Tabel 4.8 Persiapan Analisis Reliabilitas Belah Dua Ganjil Genap

No	Ganjil					ΣX	ΣX ²	Genap					ΣY	ΣY ²	ΣXY
	1	3	5	7	9			2	4	6	8	10			
1	1	1	1	1	1	5	25	1	1	1	0	1	4	16	20
2	1	1	1	1	1	5	25	1	1	1	0	1	4	16	20
3	1	1	1	1	1	5	25	0	1	1	0	1	3	9	15
4	1	0	0	0	1	2	4	1	1	0	0	1	3	9	6

No	Ganjil					ΣX	ΣX^2	Genap					ΣY	ΣY^2	ΣXY
	1	3	5	7	9			2	4	6	8	10			
5	1	1	0	1	1	4	16	0	0	1	0	1	2	4	8
6	1	1	0	1	0	3	9	0	0	0	0	1	1	1	3
7	1	1	0	0	0	2	4	1	0	1	0	1	3	9	6
8	1	1	1	1	1	5	25	1	1	1	0	1	4	16	20
9	1	0	1	0	0	2	4	0	0	1	0	0	1	1	2
10	1	0	1	1	1	4	16	1	1	1	0	1	4	16	16
11	1	1	1	0	0	3	9	1	1	1	0	0	3	9	9
12	1	1	1	1	1	5	25	1	1	1	0	1	4	16	20
13	1	1	1	1	1	5	25	0	1	1	0	0	2	4	10
14	1	0	0	0	0	1	1	1	1	1	0	1	4	16	4
15	1	1	0	1	0	3	9	0	0	1	0	0	1	1	3
16	1	0	1	1	0	3	9	0	0	1	0	1	2	4	6
17	1	1	1	1	1	5	25	1	1	1	0	1	4	16	20
18	1	1	1	1	1	5	25	0	1	1	0	1	3	9	15
19	1	0	1	0	0	2	4	1	0	1	0	0	2	4	4
20	1	0	0	0	1	2	4	1	1	1	0	0	3	9	6
	20	13	13	13	12	71	289	12	13	18	0	14	57	185	213

Data dari tabel 4.8 selanjutnya dimasukkan ke dalam rumus korelasi, Jumlah skor ganjil menjadi X dan jumlah skor genap menjadi Y. Hasil penjumlahan dirangkum sebagai berikut.

$$\Sigma X = 71 \quad \Sigma X^2 = 289 \quad \Sigma Y = 57 \quad \Sigma Y^2 = 185 \quad \Sigma XY = 213$$

Setelah data dimasukkan ke dalam rumus, diperoleh koefisien korelasi 0.37 atau termasuk kategori rendah. Peningkatan reliabilitas dapat dilakukan melalui penambahan jumlah butir.

$$r_{xy} = \frac{20 \cdot 213 - (71 \cdot 57)}{\sqrt{\{ 20 \cdot 289 - (71)^2 \} \{ 20 \cdot 185 - (57)^2 \}}}$$

$$r_{pbis} = \frac{4260 - (4047)}{\sqrt{\{5780 - 5041\} \{3700 - 3249\}}}$$

$$r_{pbis} = \frac{213}{\sqrt{(739)(451)}} = \frac{213}{\sqrt{333289}} = \frac{213}{577,32} = 0,37$$

Selain konsistensi internal menggunakan split half, reliabilitas soal dapat juga dianalisis menggunakan rumus Alpha Cronbach's dan KR 20. Alpha Cronbach's lebih banyak digunakan untuk data berskala interval sedangkan KR 20 digunakan untuk data dikotomis (1,0). Jika akan menggunakan Alpha Cronbach's untuk mencari reliabilitas butir soal tes, maka varian butir soal yang bersifat dikotomi (0 dan 1), diperoleh dengan rumus $\sigma_i^2 = p_i q_i$. Kriteria soal dinyatakan reliabel apabila mempunyai koefisien **Alpha** > 0,70 (McIntire & Miller, 2000). Berikut perbandingan rumus Alpha Cronbach's dan KR 20 dalam perhitungan reliabilitas soal tes.

Alpha Cronbach's

$$r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right\}$$

k = Jumlah butir pertanyaan

$\sum \sigma_i^2$ = varian butir soal i

σ_t^2 = varians total

$$\sigma_i^2 = \frac{\sum X_t^2}{n} - \frac{(\sum X_t)^2}{n^2}$$

$$\sigma_i^2 = \frac{JK_i}{n} - \frac{JK_s}{n^2}$$

Kuder Richardson (KR) 20

$$r_i = \frac{k}{(k-1)} \left\{ \frac{\sigma_t^2 - \sum p_i q_i}{\sigma_t^2} \right\}$$

k = jumlah butir pertanyaan

p_i = tingkat kesulitan butir

$q_i = 1 - p_i$

σ_t^2 = varians total

$$\sigma_i^2 = \frac{\sum X_t^2}{n} - \frac{(\sum X_t)^2}{n^2}$$

Contoh perhitungan reliabilitas menggunakan rumus Alpha Cronbach's dan KR 20 diawali dengan membuat tabel persiapan seperti contoh berikut ini

Tabel 4.9 Persiapan Analisis Reliabilitas Menggunakan Rumus Alpha Cronbach's dan KR 20

		1	2	3	4	5	6	7	8	9	10	X	X ²
1	A	1	1	1	1	1	1	1	0	1	1	9	81
2	B	1	1	1	1	1	1	1	0	1	1	9	81
3	C	1	0	1	1	1	1	1	0	1	1	8	64
4	D	1	1	0	1	0	0	0	0	1	1	5	25
5	E	1	0	1	0	0	1	1	0	1	1	6	36
6	F	1	0	1	0	0	0	1	0	0	1	4	16
7	G	1	1	1	0	0	1	0	0	0	1	5	25
8	H	1	1	1	1	1	1	1	0	1	1	9	81
9	I	1	0	0	0	1	1	0	0	0	0	3	9
10	J	1	1	0	1	1	1	1	0	1	1	8	64
11	K	1	1	1	1	1	1	0	0	0	0	6	36
12	L	1	1	1	1	1	1	1	0	1	1	9	81
13	M	1	0	1	1	1	1	1	0	1	0	7	49
14	N	1	1	0	1	0	1	0	0	0	1	5	25
15	O	1	0	1	0	0	1	1	0	0	0	4	16
16	P	1	0	0	0	1	1	1	0	0	1	5	25
17	Q	1	1	1	1	1	1	1	0	1	1	9	81
18	R	1	0	1	1	1	1	1	0	1	1	8	64
19	S	1	1	0	0	1	1	0	0	0	0	4	16
20	T	1	1	0	1	0	1	0	0	1	0	5	25
Total/Σ		20	12	13	13	13	18	13	0	12	14	128	900
p		1	0.6	0.65	0.65	0.65	0.9	0.65	0	0.6	0.7		
q = (1-p)		0	0.4	0.35	0.35	0.35	0.1	0.35	1	0.4	0.3		
pq		0.00	0.24	0.23	0.23	0.23	0.09	0.23	0.00	0.24	0.21	1.69	

Penggunaan rumus Alpha Cronbach's

$$\text{Rumus: } r_i = \frac{k}{(k-1)} \left\{ 1 - \frac{\sum \sigma_i^2}{\sigma_t^2} \right\}$$

Varian total

$$\sigma_i^2 = \frac{\sum X_i^2}{n} - \frac{(\sum X_i)^2}{n^2} = \frac{900}{20} - \frac{(128)^2}{400} = 45 - \frac{16384}{400} = 45 - 40,96 = 4,04$$

Varians butir soal i

$$\sum \sigma_i^2 = p_i q_i = 0,00 + 0,24 + 0,23 + 0,23 + 0,23 + 0,09 + 0,23 + 0,00 + 0,24 + 0,21 = 1,69$$

Dimasukkan ke dalam rumus Alpha Cronbach's

$$= \frac{20}{20-1} \left\{ 1 - \frac{1,69}{4,04} \right\} = 1,05 - \{1 - 0,42\} = 1,05 - 0,58 = 0,61$$

Rumus untuk varians butir data berskala interval adalah sebagai berikut.

$$\sum \sigma_i^2 = \frac{JK_i}{n} - \frac{JK_s}{n^2}$$

Dimana

JK_i = jumlah kuadrat skor item

JK_s = jumlah skor item

Data respon butir soal termasuk dalam kategori data dikotomi (1, 0). Pengujian reliabilitas data dikotomi menggunakan yang lazim menggunakan rumus KR 20 (Kuder Richardson) sedangkan rumus Alfa Cronbach lebih banyak digunakan untuk jenis data berskala interval (1, 2, 3, 4) dari kuesioner.

Cara penggunaan **rumus KR 20** adalah sebagai berikut:

$$r_i = \frac{k}{(k-1)} \left\{ \frac{\sigma_i^2 - \sum p_i q_i}{\sigma_t^2} \right\}$$

Varian total

$$\begin{aligned}\sigma_i^2 &= \frac{\sum X_i^2}{n} - \frac{(\sum X_i)^2}{n^2} \\ &= \frac{900}{20} - \frac{(128)^2}{400} = 45 - \frac{(16384)}{400} = 45 - 40,96 = 4,04\end{aligned}$$

Varians butir soal i

$$\sum p_i q_i = 0.00+0.24+0.23+0.23+0.23+0.09+0.23+0.00+0.24+0.21 = \mathbf{1.69}$$

Dimasukkan ke dalam rumus KR 20

$$\begin{aligned}r_i &= \frac{k}{(k-1)} \left\{ \frac{\sigma_i^2 - \sum p_i q_i}{\sigma_i^2} \right\} \\ &= \frac{20}{20-1} \left\{ \frac{4,04 - 1,69}{4,04} \right\} = 1,05 \left\{ \frac{2,35}{4,04} \right\} = 1,05 - 0,58 = 0,61\end{aligned}$$

Hasil analisis reliabilitas menggunakan rumus Alpha Cronbach's dan KR 20 menunjukkan indeks reliabilitas yang sama. Hasil analisis reliabilitas menggunakan rumus Alpha Cronbach's/KR 20 ternyata lebih tinggi daripada korelasi belah dua ganjil genap. Dalam output program analisis ITEMAN reliabilitas dihitung menggunakan Alpha Cronbach sedangkan program ANATES reliabilitas dihitung menggunakan korelasi belah dua ganjil genap, peneliti dapat menggunakan salah satu metode analisis yang memenuhi persyaratan.

Setelah tes memenuhi persyaratan validitas dan reliabilitas, pengembang tes dapat melanjutkan validasi silang untuk putaran terakhir proses pengembangan tes dengan cara menyelenggarakan tes pada sampel lain yang mewakili karakteristik sasaran peserta tes yang sebenarnya.

9. *Developing norms*

Setelah validasi lengkap, pengembang tes dapat menetapkan norma acuan dari distribusi skor tes untuk menginterpretasikan posisi skor tes individu dibandingkan dengan skor tes peserta tes yang lain. Selain itu, pengembang tes juga dituntut untuk menetapkan skor potong (*cut score*)

yaitu batas skor kelulusan yang digunakan untuk menetapkan keputusan seseorang termasuk dalam kategori kelompok peserta yang lolos atau gagal. Dalam ujian akhir semester (UAS) skor potong ditetapkan menggunakan batas nilai ketuntasan minimal (KKM).

10. Complete test manual

Akhir dari kegiatan pengembangann tes adalah menyusun buku petunjuk penggunaan tes (*test manual*). Isi buku petunjuk menjelaskan latar belakang pembuatan tes, sejarah proses pengembangan, hasil-hasil studi validasi, deskripsi target sasaran yang sesuai, petunjuk penyelenggaraan, cara penyekoran tes, dan informasi tentang cara menginterpretasikan skor individu. Petunjuk tes menekankan pada proses pengembangan yang terus menerus yang dimulai dari konsep pengembangan tes itu sendiri. Informasi yang terdapat pada petunjuk tes disajikan dalam tulisan yang mudah dibaca dan bahasa yang mudah dipahami.

BAB V

PENGENDALIAN KUALITAS ALAT PENILAIAN

Kebenaran, kredibilitas, dan kepercayaan hasil penilaian ditentukan oleh kualitas alat ukurnya. Kualitas alat pengukuran atau instrumen penilaian yang berupa tes maupun non tes ditentukan oleh validitas dan reliabilitasnya. Apabila alat-alat pengukuran tidak memenuhi kriteria valid dan reliabel, maka alat tersebut dianggap tidak layak digunakan untuk menilai dan guru harus mengulangi penilaian dengan alat penilaian yang telah diperbaiki. Pengukuran validitas dan reliabilitas banyak digunakan pada butir soal tes objektif. Cara-cara pemeriksaan validitas dan reliabilitas soal ujian dipaparkan sebagai berikut.

A. Validitas

Validitas mengarah pada kebenaran pengukuran dan penilaian terhadap apa yang seharusnya diukur dan dinilai. Isi soal dinyatakan valid jika terdapat kesepadanan antara materi yang diajarkan dengan soal yang diujikan. Isi soal ujian harus mewakili seluruh materi yang diajarkan. Apabila soal ujian hanya berisi sebagian kecil materi yang diajarkan maka isi alat penilaian tersebut dinyatakan tidak valid. Contoh: Tes pengetahuan gizi harus mewakili 5 (lima) unsur zat gizi, yaitu: karbohidrat, lemak, protein, mineral, dan vitamin. Apabila soal ujian hanya menanyakan zat mineral dan vitamin, maka dari sisi isi tes tersebut belum dapat dinyatakan valid karena belum mewakili unsur zat gizi karbohidrat, lemak, dan protein.

Validitas merupakan istilah yang sering digunakan untuk memberi arti 'benar', (*true or correct*) pada seperangkat alat ukur yang mampu mengukur apa yang seharusnya diukur (tepat). Validitas menuntut ide pemikiran yang tertulis dalam alat ukur mempunyai kecocokan dengan hasil pengamatan kongkret. Nitko (Nitko & Brookhart, 2011) menjelaskan

hasil asesmen dinyatakan valid jika hasil-hasil asesmen sesuai dengan kondisi yang senyatanya. Dengan kata lain, asesmen yang valid tergantung pada bagaimana hasil pengukuran diinterpretasikan dan digunakan. Dalam pengukuran validitas perlu dipertimbangkan beberapa kejadian yang dapat mempengaruhi hasil pengukuran validitas, antara lain:

- 1) Validitas dipakai untuk membantu asesor/pengembang tes menginterpretasikan dan menginformasikan kualitas soal tes yang digunakan untuk mengukur kemampuan siswa. Contoh: soal ujian pengetahuan penataan meja (*table set up*) dapat mengukur pemahaman siswa tentang peralatan, prosedur, menu, dan pelayanan kepada tamu dengan baik. Pemahaman teori tersebut dibuktikan pada saat praktik melayani tamu, siswa dapat melakukan *table set up* dengan prosedur yang benar dan alat-alat yang digunakan pun juga tepat. Dalam konteks ini, alat ukur yang valid dapat menghasilkan informasi yang sesuai dengan kondisi nyata. Dengan alat ukur yang valid, siswa yang pandai mendapat skor yang tinggi dan siswa yang kurang pandai mendapat skor yang rendah.
- 2) Hasil uji coba soal test mempunyai tingkat validitas yang berbeda jika diterapkan pada situasi dan lingkungan yang berbeda. Contoh: soal ujian pengetahuan *table set up* valid untuk mengukur pemahaman pelayanan makanan dan minuman (*food and beverage service*) pada siswa SMK Tata Boga yang memiliki fasilitas Tata Hidang lengkap, tetapi soal tersebut mungkin validitasnya rendah jika diujikan pada siswa SMK Tata Boga yang memiliki fasilitas Tata Hidang tidak lengkap. Hal ini antara lain disebabkan karena soal Tata Hidang untuk siswa SMK yang fasilitas praktiknya belum lengkap bisa dianggap sulit sehingga siswa tidak memiliki pengalaman konstektual yang membantu pemahamannya. Soal yang sulit cenderung dijawab dengan cara menebak-nebak sehingga jawaban betul belum tentu menunjukkan kemampuan siswa yang sebenarnya tetapi mungkin secara kebetulan cocok. Validitas soal yang tinggi dapat diperoleh dari soal yang memiliki daya beda tinggi. Daya beda ditentukan oleh kemampuan siswa kelompok atas menjawab benar pada soal-soal yang sulit.
- 3) Hasil pengukuran yang telah valid perlu dibandingkan dengan hasil-hasil pengukuran lain seperti hasil ulangan harian atau hasil belajar sebelumnya. Jika terdapat konsistensi skor tes yang diperoleh sekarang dengan skor tes sebelumnya, maka validitas soal saat ini sudah benar-

benar bagus. Ada beberapa situasi yang perlu dikendalikan agar hasil tes dapat dipercaya antara lain: apakah benar-benar nyata hasil kerja sendiri, tidak ada kecurangan, dan kondisi fisik sedang sehat atau tidak *bad mood*. Sebelum menyimpulkan hasil validitas tes maka kumpulkan dulu sejumlah informasi yang berkaitan dengan hasil tes yang diperoleh siswa tersebut,.

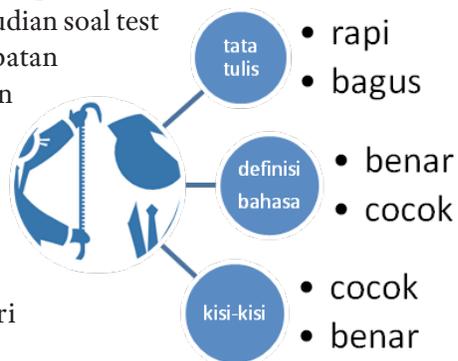
Ada empat prinsip validitas menurut Messick dalam (Linn R. L., 1989) yaitu kesesuaian interpretasi, kesesuaian penggunaan, kesesuaian nilai dan kesesuaian konsekuensi. Hasil asesmen hanya valid jika menunjukkan kesesuaian dengan kondisi yang sebenarnya. Skor test hanya dapat digunakan dari hasil asesmen yang valid. Interpretasi dan penggunaan hasil asesmen hanya valid ketika nilai-nilai yang diperoleh cocok dengan kemampuan yang sebenarnya. Artinya: siswa yang di kelas tergolong pandai akan mendapat skor lebih tinggi dibanding teman-teman lain di kelasnya. Interpretasi dan penggunaan hasil asesmen valid hanya ketika memberi konsekuensi nilai yang konsisten.

Pengukuran validitas dilakukan secara berjenjang, mulai dari tampak, isi kemudian konstruk teorinya. Soal tes minimal memiliki validitas tampak (*face*) dan isi (*content*) validity.

1) Face validity

Face validity paling mudah dicapai dan menjadi dasar bagi bermacam-macam pengukuran validitas lainnya. *Face validity* atau validitas tampak dimulai dari melihat penampilan fisik soal test seperti tata tulis, penggunaan kalimat, tata bahasa, dan pencetakan.

Setelah hal tersebut tercapai, kemudian soal test disimak kebenaran definisi, ketepatan kisi-kisi, kebenaran penulisan, dan kecocokan istilah yang digunakan dengan tingkat pendidikan dan usia peserta tes. Validitas tampak dalam penilaian soal tes dilakukan dengan cara telaah teoritis oleh ahli (*expert*) materi pelajaran yang sama.



Gambar 5.1 Ilustrasi *Face Validity*

2) *Content validity*

Content validity atau validitas isi dicapai setelah validitas tampang terpenuhi. Validitas isi dicapai melalui cara memeriksa kesesuaian isi tes dengan materi pelajaran. Setiap kisi-kisi materi dikembangkan menjadi beberapa butir soal. Apabila terdapat beberapa kisi-kisi, jumlah butir soal yang mewakili kisi-kisi perlu dicek keseimbangannya. Validitas mudah dicapai apabila isi tes merata pada seluruh materi. Pembuktian validitas isi dapat dilakukan dengan meminta pertimbangan ahli (*expert judgment*) yaitu guru yang mengampu mata pelajaran yang sama misalnya soal kontinental ditelaah oleh guru mata pelajaran kontinental soal pengetahuan gizi dimintakan pertimbangan/telaah butir kepada ahli gizi. Melalui pertimbangan ahli tersebut diharapkan tidak akan terjadi kesalahan konsep dari pertanyaan maupun pilihan jawaban.

Validitas isi dapat dicapai melalui tiga tahap yaitu:

- a) Memfokuskan isi sesuai dengan cakupan materi/ kompetensi dasarnya.
- b) Mengembangkan indikator (kisi-kisi), yang mencakup semua bagian materi.
- c) Membuat butir pertanyaan/ pernyataan yang mewakili semua kisi-kisi materi.

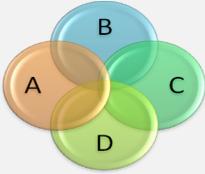
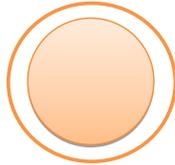
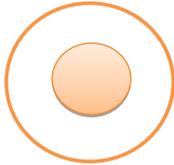
Supaya dapat memenuhi validitas isi maka penyusunan butir tes dapat dirancang melalui penjabaran materi ke dalam kisi-kisi. Contoh: ruang lingkup materi mata pelajaran ilmu gizi terdiri atas 5 unsur zat gizi. Contoh kisi-kisi soal ujian Ilmu Gizi dapat dilihat pada Tabel 5.1 berikut ini.

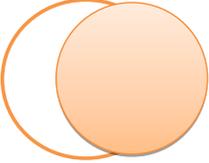
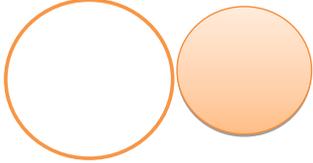
Tabel 5.1 Kisi-kisi Butir Soal Ilmu Gizi

No	Materi	Kisi-kisi	Butir-butir soal
1	Karbohidrat	a. Klasifikasi karbohidrat b. Fungsi karbohidrat c. Makanan sumber karbohidrat d. Metabolisme karbohidrat	
2	Protein	a. Klasifikasi protein b. Fungsi protein c. Sumber protein nabati dan hewani d. Penyakit kekurangan atau kelebihan protein e. Diet tinggi/rendah protein	

No	Materi	Kisi-kisi	Butir-butir soal
3	Lemak	a. Jenis-jenis lemak b. Fungsi lemak c. Sumber lemak jenuh dan tidak jenuh d. Diet rendah lemak	
4	Vitamin	a. Klasifikasi vitamin b. Fungsi vitamin yang larut dalam lemak (A, D, E, K) dan larut dalam air (B, C) c. Bahan pangan sumber vitamin yang larut dalam lemak (A, D, E, K) dan larut dalam air (B, C)	
5	Mineral	a. Klasifikasi mineral b. Bahan pangan sumber mineral makro (Na, Cl, Ca, K, Mg, P, dan S) dan mineral mikro (Fe, Zn, Y, Mn, Cu, Se, F, Co) c. Fungsi mineral makro dan mikro d. Penyakit kekurangan mineral	

Dalam validitas isi, soal dinyatakan valid jika mewakili seluruh materi. Jika materi terdiri atas 4 kompetensi dasar, maka soal juga mencakup semua kompetensi dasar. Jika ujian menggunakan soal esai, ada kemungkinan tidak semua kompetensi dasar dapat diujikan. Oleh sebab itu, pengukuran hasil belajar harus menggunakan berbagai macam alat pengukuran (*multiple measurement*). Berikut ini terdapat ilustrasi soal yang memenuhi kriteria validitas isi terdapat pada soal yang mewakili seluruh materi.

Materi/KD	Kisi-kisi	Soal
		A: B: C: D: Valid, karena mewakili semua materi
		

A. Soal mewakili seluruh materi	B. Soal hanya berisi sebagian materi
	
C. Sebagian soal yang diujikan berada di luar materi yang diajarkan	D. Soal yang diujikan berbeda dengan materi yang diajarkan

Gambar 5.2 Ilustrasi validitas Isi

Validasi para ahli (*expert*) dapat diberikan dalam bentuk skor nilai dari 1-5 dengan ketentuan sebagai berikut:

1. Tidak baik
2. Kurang baik
3. Cukup baik
4. Baik
5. Sangat baik

Hasil penilaian ahli dapat dianalisis untuk mengetahui indeks validitasnya. Rumus untuk menghitung indeks validitas menurut Aiken (Aiken L. R., 1991) adalah sebagai berikut:

$$V = \frac{\sum S}{n(c - 1)}$$

Keterangan:

$s = r - Lo$

Lo = nilai minimum yang bisa diberikan validator (misalnya 1)

c = nilai maksimum yang bisa diberikan validator (misalnya 5)

n – jumlah ahli/validator

r = nilai yang diberikan oleh *expert*

Contoh:

Kelayakan butir soal dimintakan pendapat tiga orang ahli (*expert*) atau sebagai validator memperoleh data sebagai berikut:

No	Komponen yang dinilai	Nilai validator		
		1	2	3
1	Penulisan soal rapidan menarik	4	5	4
2	Isi soal mencakup seluruh kisi-kisi materi	5	4	5
3	Konstruksi soal sesuai persyaratan	3	3	4
4	Bahasa yang digunakan baku dan mudah dipahami	4	3	4

Keterangan Skor:

1. Tidak baik
2. Kurang baik
3. Cukup baik
4. Baik
5. Sangat baik

Cara analisis:

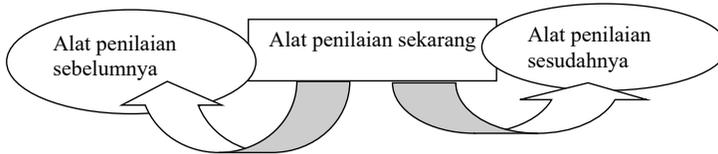
Butir	Nilai Validator			Σs	$V = \frac{\Sigma s}{n(c-1)}$
	1	2	3		
1	4	5	4	10	$\frac{10}{3(5-1)} = 0,83$
s	3	4	3		
2	5	4	5	11	$\frac{11}{3(5-1)} = 0,92$
s	4	3	4		
3	3	3	4	7	$\frac{7}{3(5-1)} = 0,58$
s	2	2	3		
4	4	3	4	8	$\frac{8}{3(5-1)} = 0,67$
s	3	2	3		

Kriteria: Butir soal dinyatakan valid jika memiliki koefisien validitas isi (V) $\geq 0,5$. Berdasarkan kriteria tersebut, maka soal sudah memiliki validitas dari unsur tampilan, isi, konstruksi dan bahasa karena indeks validitas yang paling rendah $0,58 > 0,5$.

3) Validitas kriteria

Validitas kriteria (*criterion validity*) ditemukan dengan cara mengorelasikan hasil penilaian sekarang dengan beberapa hasil penilaian yang sudah dilakukan atau hasil penilaian yang akan datang. Validitas kriteria dapat digunakan jika kisi-kisi dari dua perangkat soal yang

dikorelasikan sama. Ada dua tipe validitas ini yaitu: *concurrent validity* dan *predictive validity* yang diilustrasikan pada gambar berikut ini



Gambar 5.3 Konsep Validitas Kriteria

- a) ***Concurrent validity* (validitas kongruence)** dapat dicapai apabila hasil penilaian yang sedang dilakukan mempunyai korelasi positif (hubungan linear) dengan hasil penilaian sebelumnya yang telah dinyatakan valid. Soal tes baru dinyatakan valid apabila memperoleh hasil yang konsisten dengan hasil pengukuran yang telah dilakukan sebelumnya pada subjek yang sama. Contoh: guru yang membuat soal tes baru untuk mengukur kemampuan mata pelajaran Ilmu Gizi agar soal tes baru yang dibuat memenuhi validitas kriteria (*concurrent validity*) peneliti harus mengorelasikan satu set skor tes Ilmu Gizi dengan skor tes mata pelajaran lain yang sudah pernah diuji sebelumnya (dengan asumsi kisi-kisi yang diujikan setara). Hal ini berarti bahwa kebanyakan orang yang mempunyai skor tinggi pada pengukuran sebelumnya juga akan mendapat skor yang tinggi pada pengukuran yang baru dan sebaliknya. Dua alat pengukuran dapat berasosiasi secara sempurna apabila dimensi yang diukur sama, secara logis pengukuran akan memperoleh hasil yang serupa.

Validitas kriteria dapat diterapkan pada semua indikator atau hanya beberapa indikator yang memiliki kesesuaian saja. *Concurrent validity* banyak digunakan dalam pengembangan alat tes baru misalnya tes IQ, tes potensi belajar, TOEFL, tes prestasi belajar, soal ujian nasional dan sebagainya.

- b) ***Predictive validity* (validitas prediktif)** termasuk validitas kriteria tetapi kriteria yang digunakan berbeda dengan validitas kongruence dan berfungsi sebaliknya. Pada validitas kongruence, kriteria yang digunakan adalah hasil penilaian sebelumnya sedangkan kriteria yang digunakan dalam validitas prediksi adalah hasil penilaian berikutnya. Validitas prediksi sering digunakan untuk mengukur kekuatan prediksi alat tes seleksi dengan prestasi setelah lolos seleksi dan mengikuti

program pendidikan. Prediksi akan efektif apabila dimensi yang diukur dalam kedua tes itu sama, tetapi hal ini jarang ditemukan.

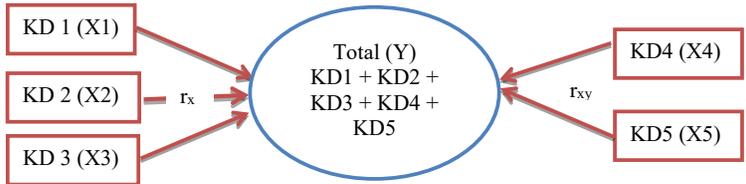
Tes seleksi yang paling sering digunakan dalam pengukuran validitas prediksi adalah tes kecerdasan atau tes potensi belajar yang tidak pernah dipelajari sebelumnya. Sedangkan kriteria yang digunakan adalah semua jenis kemampuan baik yang diukur melalui tes maupun non tes. Contoh: Pada saat penerimaan siswa baru, calon siswa diukur dengan *Scholastic Aptitude Test* (SAT) untuk mengukur kemampuan awal siswa. Hasil pengukuran kemampuan awal dengan SAT tersebut kemudian digunakan beberapa tahun kemudian setelah siswa menempuh pendidikan atau pada saat ujian kelulusan. Alat pengukur yang baik akan mempunyai validitas prediksi yang tinggi setelah hasil pengukuran dengan kedua alat pengukuran tersebut dikorelasikan. Apabila SAT mempunyai validitas prediktif yang tinggi, siswa yang mempunyai skor SAT tinggi secara substansial dapat bekerja lebih baik di sekolah. Apabila siswa dengan skor tinggi menunjukkan kesamaan dengan siswa dari kelompok rata-rata atau yang lebih rendah, maka SAT tersebut mempunyai validitas prediksi yang rendah.

4) **Construct validity**

Validitas konstruk (*construct validity*) digunakan untuk seperangkat alat ukur/tes yang memiliki indikator ganda. Teori dibangun berdasarkan indikator-indikator yang saling mendukung. Dalam konteks penilaian hasil belajar, indikator disepadankan dengan kisi-kisi soal. Satu indikator pada umumnya mewakili satu kompetensi dasar (KD). Alat ukur yang valid ditunjukkan oleh konsisten hasil penilaian meskipun indikator yang digunakan bervariasi. Analisis validitas konstruk dilakukan agar tidak terjadi *overlapping* pengukuran pada indikator yang sama. Untuk menghindari *overlapping* tersebut, setiap kisi-kisi perlu ditegaskan definisi konseptualnya secara jelas. Contoh: Materi pelayanan makan dan minum terdiri atas beberapa KD yang perlu dievaluasi ,yaitu

- 3.1. Menentukan penataan meja prasmanan (*buffet*) (X1)
- 3.2. Mengevaluasi layanan makan minum di kamar (X2)
- 3.3. Membedakan berbagaiminuman nonalcohol (X3)
- 3.4. Menganalisis minuman panas (X4)
- 3.5. Menganalisis minuman dingin (X5)

Isi soal tes valid jika mewakili semua KD yang diajarkan. Butir-butir soal yang mengukur setiap KD dinyatakan valid jika jumlah skor pada KD tersebut berkorelasi positif dengan skor seluruh KD. Dalam proses analisis, jumlah skor KD diberi simbol X sedangkan jumlah skor total (KD1 + KD2+ KD3 + KD4+ KD5) diberi simbol Y. Ilustrasi model pengukuran validitas konstruk terdapat pada gambar di bawah ini.



Gambar 5.4 Model Pengukuran Validitas Konstruk

Contoh analisis validitas konstruk soal Tata Hidang pada materi penataan meja makan untuk SMKTata Boga mengukur tiga KD, setiap KD terdiri atas 10 butir soal. Uji coba soal diikuti oleh 15 peserta. KD yang diuji validitas konstruknya adalah:

- 3.1. Memilih peralatan makan dan minum; alat hidang serta lenan sebagai X_1
- 3.2. Menganalisis penataan meja (*table set-up*) sebagai X_2
- 3.3. Menentukan pelayanan makan dan minum di restoran sebagai X_3

Analisis validitas konstruk menggunakan korelasi *product moment* dengan rumus sebagai berikut:

$$r_{xy} = \frac{N \sum XY - \sum X \sum Y}{\sqrt{\{ N \sum X^2 - (\sum X)^2 \} \{ N \sum Y^2 - (\sum Y)^2 \}}}$$

Keterangan

- $\sum X_1$ = total nilai per KD
- $\sum Y$ = total nilai semua KD
- N = jumlah peserta tes

Contoh analisis korelasi *product moment* menggunakan bantuan program excel dilakukan melalui langkah-langkah sebagai berikut:

- 1) Buat tabel persiapan dari data respon butir yang telah di beri skor 1 jika betul dan 0 jika salah.
- 2) Jumlahkan skor setiap KD, kemudian jejerkan jumlah skor per KD tersebut pada kolom sebelahnya kemudian beri simbol KD1(ΣX_1), KD2 (ΣX_2), dan KD3 (ΣX_3).
- 3) Tambahkan jumlah skor semua KD (KD1+ KD2 + KD3) pada kolom terakhir. Beri simbol total (ΣY),.
- 4) Ketik =CORREL (di bawah baris terakhir, secara otomatis program excel akan meminta kolom input data X (array1) dan kolom input data Y (array2). Contoh :



- 5) Setelah muncul dialog =CORREL(array1, array2), maka tempatkan kursor pada kolom X1, kemudian tarik kursor dari data nomor 1 sampai terakhir.
- 6) Beri tanda baca koma atau titik koma (, atau ;) sesuai dengan permintaan program.
- 7) Ulangi langkah ke 5 untuk array2 pada kolom data Y (total).
- 8) Beri tanda kurung tutup ...) kemudian tekan ENTER, hasil analisis korelasi sudah muncul pada sel tempat saudara menulis.

Tabel 5.2 Input Data Validitas Konstruk dari Program Excel

No	Skor Butir Soal KD1										Jumlah Skor Butir per KD			Total
	1	2	3	4	5	6	7	8	9	10	ΣX_1	ΣX_2	ΣX_3	
1	1	1	1	1	1	0	1	1	1	1	9	5	9	23
2	1	1	1	1	1	0	0	0	0	0	5	2	6	13
3	1	1	0	1	0	0	1	1	1	0	6	5	1	12

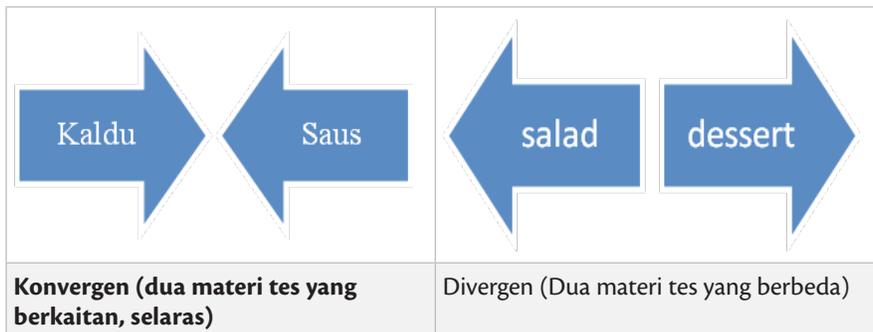
4	1	1	1	1	0	0	1	0	1	1	7	8	8	23
5	1	1	1	1	0	0	1	1	1	1	8	8	7	23
6	1	1	0	1	1	1	1	0	1	1	8	6	6	20
7	1	0	1	0	1	0	0	0	0	0	3	9	9	21
8	1	1	0	1	1	0	1	1	0	0	6	5	7	18
9	1	1	0	0	1	0	0	0	0	0	3	8	9	20
10	1	0	0	0	1	0	0	0	0	0	2	9	9	20
11	1	0	0	1	0	0	0	1	1	0	4	7	8	19
12	1	1	0	0	1	0	0	1	1	0	5	5	9	19
13	1	1	1	1	1	0	1	1	1	0	8	6	8	22
14	1	0	0	1	1	0	0	0	0	0	3	5	9	17
15	1	1	1	1	1	1	1	1	1	1	10	8	8	26
	15	11	7	11	11	2	8	8	9	5	87	96	113	296
	=CORREL(array1,array2)										r_{xiy}	r_{x2y}	r_{x3y}	
X		Array1				Koefisien korelasi				0.47	0.65	0.61		
Y		Array2				Interpretasi				Rendah	Sedang			

	Skor Butir Soal KD2										ΣX_2
	1	2	3	4	5	6	7	8	9	10	
1	1	0	1	1	0	1	0	0	0	1	5
2	1	0	0	0	0	0	1	0	0	0	2
3	1	0	1	1	0	1	1	0	0	0	5
4	1	1	1	1	1	1	0	1	0	1	8
5	0	1	1	1	1	1	0	1	1	1	8
6	0	1	1	1	1	1	0	0	0	1	6
7	0	1	1	1	1	1	1	1	1	1	9
8	1	0	1	0	1	1	0	0	0	1	5
9	0	1	1	1	0	1	1	1	1	1	8

10	0	1	1	1	1	1	1	1	1	1	9
11	1	1	1	1	1	1	0	1	0	0	7
12	0	0	0	1	1	1	1	0	0	1	5
13	1	0	1	0	1	1	0	1	0	1	6
14	1	0	0	1	0	1	1	0	1	0	5
15	1	0	0	1	1	1	1	1	1	1	8
	9	7	11	12	10	14	8	8	6	11	96
No	Skor Butir Soal KD3										ΣX_3
	1	2	3	4	5	6	7	8	9	10	
1	1	1	1	1	1	1	1	1	0	1	9
2	0	0	0	1	1	1	1	1	0	1	6
3	0	0	0	0	0	0	0	1	0	0	1
4	0	1	1	1	1	1	1	1	0	1	8
5	1	0	1	1	1	0	1	1	0	1	7
6	1	0	1	0	0	1	0	1	1	1	6
7	1	1	1	1	1	1	1	1	0	1	9
8	0	0	0	1	1	1	1	1	1	1	7
9	1	1	1	1	1	1	1	1	0	1	9
10	1	1	1	1	1	1	1	1	0	1	9
11	1	1	1	1	0	1	1	1	0	1	8
12	1	1	1	1	1	1	1	1	0	1	9
13	1	1	1	0	1	1	1	1	0	1	8
14	1	1	0	1	1	1	1	1	1	1	9
15	1	1	0	1	1	1	1	0	1	1	8
	11	10	10	12	12	13	13	14	4	14	113

Validitas konstruk dibedakan menjadi dua, yaitu validitas konvergen dan divergen atau diskriminan (*discriminant validity*). Validitas konvergen diterapkan ketika indikator-indikator yang digunakan dalam satu variabel/ materi konvergen atau saling berasosiasi antara indikator yang satu dengan

indikator yang lainnya. Validitas divergen dinamakan juga validitas diskriminan. Validitas divergen mempunyai peranan yang berlawanan dengan validitas konvergen yaitu indikator yang satu tidak saling tergantung dengan indikator yang lainnya. Ini berarti bahwa dalam satu variabel/materi dapat tersusun indikator yang saling tergantung dan berasosiasi positif atau konvergen tetapi dapat pula terdapat indikator-indikator yang berasosiasi negatif atau divergen. Validitas divergen ditemukan apabila dua konstruk A dan B sangat berbeda, dan hasil pengukuran A dan B tidak berasosiasi. Ilustrasi validitas materi konvergen dan divergen terdapat pada gambar berikut ini:



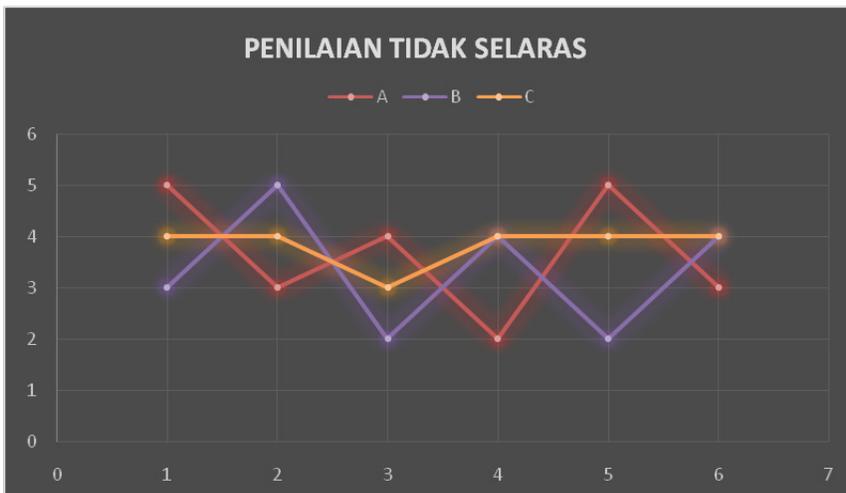
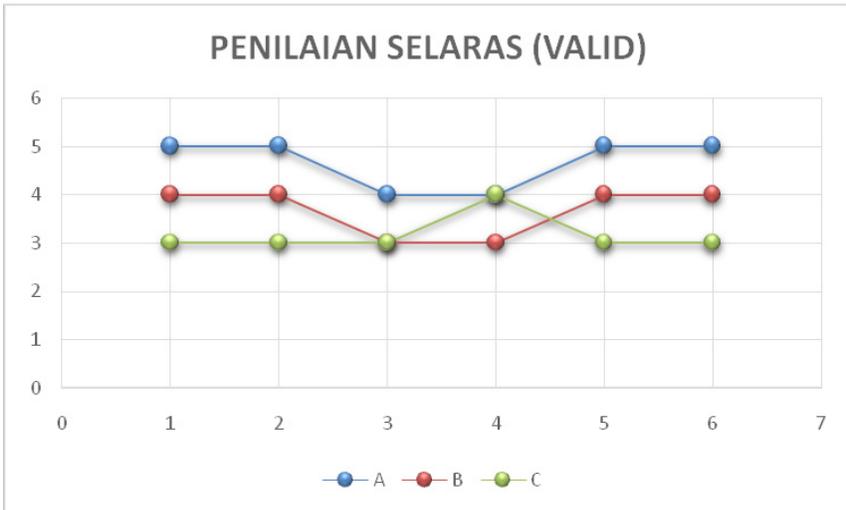
Gambar 5.5 Ilustrasi Validitas Konvergen dan Divergen

5) Validitas Penilaian Kinerja

Penilaian kinerja termasuk dalam penilaian autentik. Kinerja dinilai menggunakan lembar observasi. Penilaian kinerja minimal dilakukan oleh dua orang asesor (penilai). Pengendalian validitas hasil penilaian kinerja dilakukan melalui validitas *inter rater* atau *interobserver*. Jika hasil penilaian dari penilai (rater) pertama dan kedua sudah selaras, maka hasil penilaian sudah dinyatakan valid. Keselarasan nilai antara asesor (penilai) satu dan lainnya dapat dianalisis menggunakan korelasi produk moment.

Penilai profesional harus mengetahui dan dapat membedakan kinerja yang baik dan kurang baik supaya dapat memberi penilaian yang objektif. Jika masing-masing penilai menguasai bidang kerja yang dinilainya maka selisih nilai antar asesor tidak terlalu jauh, dengan kata lain selaras. Dalam penilaian kinerja, sering ditemukan hasil penilaian yang tidak selaras, oleh sebab itu sebelum Keputusan (lulus, tidak lulus) ditetapkan perlu dilakukan penyesuaian terlebih dahulu. Berikut ini terdapat contoh hasil penilaian

yang selaras dan tidak selaras. Penilaian selaras terdapat pada gambar atas, yaitu nilai antar asesor tidak bertentangan. Penilaian tidak selaras terdapat pada gambar bawah, yaitu nilai antar asesor saling bertentangan.



Gambar 5.6 Contoh Hasil Penilaian Interrater

Linn, Baker and Dunbar (1991), berpendapat bahwa penilaian berbasis kinerja valid jika dilakukan secara objektif (tidak berpihak) dan transparan. Untuk meningkatkan objektivitas hasil penilaian kinerja maka digunakan

rubrik yang memuat kumpulan kriteria untuk mengevaluasi kinerja. Nilai tinggi diberikan kepada siswa yang dapat memenuhi kriteria kerja yang baik. Validitas penilaian kinerja dalam kegiatan praktik mingguan dapat dilakukan dengan bantuan penilaian teman sebaya (*peer assessment*). Hasil praktik siswa dinilai oleh dua atau tiga siswa lain yang bukan anggota kelompoknya. Masing-masing siswa secara terpisah mengevaluasi pekerjaan temannya sesuai dengan rubrik penilaian yang sudah ditetapkan. Hasil penilaian teman sebaya ini kemudian dikorelasikan.

Terdapat tiga perpekstif pengukuran validitas penilaian otentik yaitu: *One-dimensional perspective, desegregated perspective, reciprocal perspective*. Perspektif satu terjadi jika nilai individu dikorelasikan dengan nilai lain (ekternal) yang telah diperoleh sebelumnya. Perspektif yang disagregat (terpisahkan) diperoleh jika penilai mempertimbangkan bahwa kemampuan individu bersifat multidimensi yang dipengaruhi oleh berbagai macam factor. Oleh sebab itu, faktor-faktor yang mempengaruhi hasil penilaian perlu diuji validitasnya. Contoh: kompetensi diukur dari komponen kognitif, afektif, dan psikomotor sehingga perlu diukur pengaruh antar komponen tersebut menggunakan rumus korelasi. Perspektif resiprokal yang menegaskan bahwa keberhasilan dipengaruhi oleh adanya faktor timbal balik antara sikap terhadap sistem evaluasi dan penggunaannya. Sistem evaluasi yang dapat dipercaya, adil, dan dapat dipertanggungjawabkan dapat menimbulkan sikap positif peserta ujian (Raimundo & Zulantay, 2007).

B. Reliabilitas

Reliabilitas mengandung pengertian ada ketetapan atau konsistensi yaitu apabila sebuah stimulus diulang, atau terulang di bawah kondisi yang hampir sama maka akan menghasilkan respon yang sama. Stimulus dapat berupa pertanyaan atau pernyataan sedangkan respon dapat berupa jawaban atau hasil pengukuran. Proses pengukuran yang hasilnya tidak stabil atau tidak konsisten akan dikatakan tidak reliabel. Contoh jika siswa mengerjakan dua soal tes paralel A dan B memperoleh hasil yang hampir sama maka tes A maupun tes B sudah reliabel untuk mengukur kemampuan siswa.

Reliabilitas dapat berarti keterikatan, ketergantungan, ketetapan atau keajegan hasil pengukuran. Contoh keajegan beberapa alat pengukur dapat dibuktikan melalui ilustrasi sebagai berikut: (a) Alat pengukur

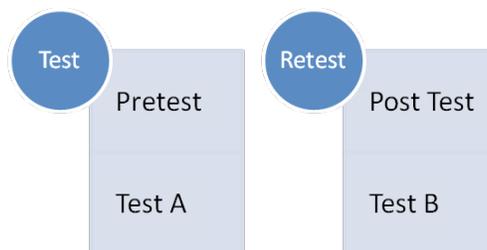
berat mempunyai skala yang reliabel apabila memberi bobot yang sama pada barang yang sama tetapi ditimbang pada waktu yang berbeda; (b) Spedometer kendaraan dapat mengukur dengan reliabel pada saat kendaraan berjalan cepat atau lambat, spedometer tersebut cukup konstan untuk mengukur jarak yang sudah ditempuh (Neuman, 2000); Neuman menjelaskan tiga tipe reliabilitas yaitu *stability reliability*, *representative reliability* dan *equivalence reliability* sebagai berikut:

1) Stability Reliability

Soal tes memenuhi syarat reliabel jika diujikan beberapa kali (*test-retest*) di bawah kondisi yang sama, hasilnya juga hampir sama atau **stabil**. Selain menggunakan metode test-retest, stabilitas juga dapat diukur dengan menggunakan *parallel test* atau tes alternatif yang memiliki bentuk sangat mirip. Tes retest dilaksanakan menggunakan butir soal yang sama. Pada parallel test ini, indikator yang dikembangkan dalam perangkat tes dirancang sama dan butir-butir instrumen yang digunakan juga memiliki kesamaan isi. Contoh butir soal untuk parallel tes:

Tes A	Tes B
Sayuran yang banyak mengandung pro vitamin A adalah	Sayuran yang banyak mengandung pro vitamin A adalah
<p>A. Wortel</p> <p>B. Bayam</p> <p>C. Taoge</p> <p>D. Labu siam</p> <p>E. Daun kol</p>	<p>A. Labu kuning</p> <p>B. Bayam</p> <p>C. Taoge</p> <p>D. Labu siam</p> <p>E. Daun kol</p>

Paralel tes jarang digunakan karena peneliti harus membuat dua soal tes yang setara isinya sehingga membutuhkan waktu lama. Contoh ilustrasi reliabilitas yang menggunakan konsep stabilitas terdapat pada gambar:

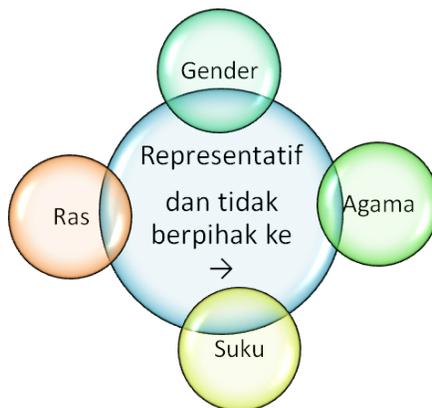


Gambar 5.7 Tipe Kestabilan dalam Mengukur Reliabilitas

Gambar menunjukkan bahwa tes dilakukan dalam dua waktu yang berbeda. Test–retest dapat menggunakan soal yang sama untuk pre test dan post test atau menggunakan dua perangkat tes yang berbeda (tes A dan tes B) tetapi indikator yang diukur sama (*parallel test*) dan butir soal dibuat setara.

2) Representative reliability

Reliabilitas representatif adalah reliabilitas silang antar kelompok. Reliabilitas representatif dibuktikan dengan menyebarkan instrumen pada beberapa kelompok siswa yang berbeda. Sebuah indikator mempunyai reliabilitas yang tinggi apabila memperoleh variasi hasil yang sama/setara ketika diterapkan pada kelas yang berbeda. Untuk memperoleh hasil pengukuran yang reliabel ini, tiap-tiap kelompok yang diukur sebaiknya representatif, mewakili seluruh karakteristik siswa yang beragam. Hasil analisis menentukan indikator mana yang reliabel dan tidak pada beberapa kelompok orang tertentu. Soal dinyatakan memiliki reliabilitas representatif jika tidak bias gender, suku, agama, ras, umur, dll. Pembuktian reliabilitas soal menggunakan uji beda, jika kelompok A dan kelompok B atau memiliki varians yang homogen dan tidak menunjukkan ada perbedaan.



Gambar 5.8 Tipe Representatif dalam Mengukur Reliabilitas

Gambar di atas mendeskripsikan reliabilitas representatif jika kumpulan soal mewakili semua kelompok yang berbeda. Soal tidak memihak kepada kelompok gender, suku, ras dan agama (SARA) tertentu sehingga salah satu dari mereka tidak merasa lebih diuntungkan. Untuk menyusun soal yang representatif ini gunakanlah permasalahan yang umum, tidak bersifat

rahasia yang hanya diketahui oleh kelompok tertentu saja, memiliki referensi yang jelas dan bisa dilacak. Contoh soal bias SARA:

Sumber bias	Soal yang bias
Gender	Susunlah menu diet untuk wanita yang sedang hamil dan menyusui Sumber zat gizi apa yang diprioritaskan pada menu tersebut
Budaya	Makna simbolis yang terdapat pada sesaji makanan pada budaya nyadran di DIY
Ras	Sebutkan hidangan yang harus ada pada tahun baru imlek :
Agama	Susun menu Aqiqah untuk kelahiran bayi laki-laki

3) *Equivalence Reliability*

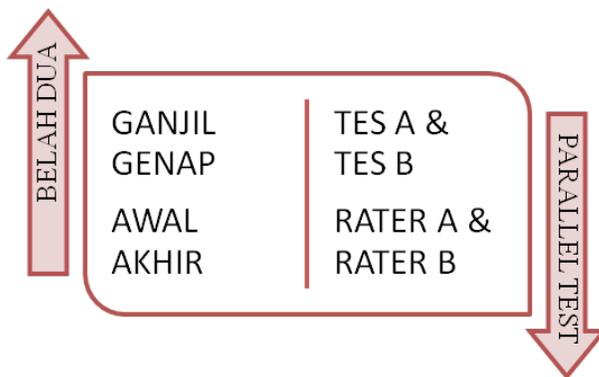
Reliabilitas ekuivalensi diterapkan ketika penilai menggunakan indikator ganda dalam mengukur kemampuan siswa. Maksud ekuivalen di sini adalah ada dua kelompok butir soal yang mempunyai nilai (ukuran, kondisi, arti, atau efek) yang sama, seimbang, sebanding, sepadan akan dibandingkan. Untuk memperoleh keseimbangan, butir pertanyaan disusun sedemikian rupa sehingga semua indikator memiliki bobot yang sama. Contoh kongkret misalnya sebuah tes terdiri atas 50 butir soal yang dikembangkan dari 5 indikator yang berbeda dan setiap indikator memiliki jumlah butir soal yang hampir sama. Reliabilitas terbukti apabila data yang diperoleh cukup konsisten untuk 5 indikator yang berbeda. Apabila sejumlah indikator yang berbeda berfungsi, pengukuran yang reliabel akan memberi hasil yang relatif sama untuk semua indikator.

Reliabilitas ekuivalensi dilakukan melalui metode *split half* (belah dua) dan *parallel test*. Pengukuran reliabilitas dengan metode *split half* menganut konsep *konsistensi internal*, yaitu guru cukup membuat satu perangkat soal dan satu kali pengukuran. Cara analisis reliabilitasnya adalah: butir-butir pertanyaan yang mengukur konstruk/indikator yang sama dibagi menjadi dua (ganjil-genap) atau (awal-akhir). Apabila analisis data menunjukkan hasil yang konsisten, kelompok butir yang telah dibagi dua tersebut memiliki korelasi tinggi ($>0,6$) maka instrumen tersebut dinyatakan reliabel. Paralel tes (tes A dan tes B) yang diujikan dalam waktu bersamaan juga menunjukkan validitas ekuivalensi karena soal-soal mengukur indikator sama dan memiliki tingkat kesulitan butir yang relatif sama pula.

Tabel 5.3 Pedoman untuk Memberikan Interpretasi Terhadap Koefisien Korelasi

Interval Koefisien	Tingkat Hubungan/korelasi	
0,00 – 0,199	Sangat lemah	Sangat rendah
0,20 – 0,399	Lemah	Rendah
0,40 – 0,599	Sedang	Sedang
0,60 – 0,799	Kuat	Tinggi
0,80 – 1,000	Sangat kuat	Sangat tinggi

Reliabilitas ekuivalensi diterapkan pada penilaian uji kompetensi, lomba keterampilan siswa (LKS), dan penilaian keterampilan lainnya menggunakan validitas (*interater dan intercoder*). Satu orang peserta dinilai atau diamati oleh beberapa orang asesor/rater. Penilaian oleh asesor kepada peserta uji kompetensi diberi rentang waktu (jam/hari). Reliabilitas ekuivalensi akan tercapai apabila data hasil penilaian/pengamatan beberapa asesor/rater pada rentang waktu tertentu cukup konsisten, selaras, sepadan sehingga tidak banyak selisihnya. Sebuah pengukuran reliabel apabila masing-masing penilai/pengamat memberi nilai yang hampir sama pada rentang waktu yang berbeda. Ilustrasi tipe ekuivalensi dalam pengukuran reliabilitas terdapat pada gambar berikut ini.



Gambar 5.9 Tipe Ekuivalensi dalam Mengukur Reliabilitas

4) Reliabilitas Penilaian Kinerja

Penilaian kinerja menggunakan pengamatan yang cara pembuktiannya berbeda karena hasil pengukurannya tidak memperoleh data kuantitatif. Reliabilitas berarti keterkaitan atau konsistensi. Penilaian kinerja dilakukan oleh 2-3 observer/rater menggunakan rubrik penilaian sama. Untuk mencapai konsistensi (tidak berubah-ubah) perilaku subjek diobservasi dari waktu ke waktu, peneliti melakukan observasi berulang-ulang seperti prinsip-prinsip stabilitas dalam pengukuran reliabilitas penilaian tes objektif.

Tabel 5.4 Rangkuman Cara-cara Pengukuran Reliabilitas

Konsep	Jenis	Contoh	
Stabilitas	Test-retest	Pre test	Post test
	Parallel tes	Tes A	Test B
Representatif	Keterwakilan kelompok	Pria	Wanita
		Islam	Non Islam
		Negeri	Swasta
Ekuivalensi	Split half	Ganjil Awal	Genap Akhir
	Parallel	Tes A	Tes B
	Interrater/observer	Rater A Rater 1	Rater B Rater 2

5) Upaya Meningkatkan Reliabilitas

Instrumen yang berkualitas dituntut memiliki reliabilitas yang tinggi. Apabila instrumen tersebut belum dapat memenuhi kriteria reliabilitas yang ditetapkan, maka ada kemungkinan proses pengukuran perlu diperbaiki. Ada 4 cara yang dapat dilakukan untuk memperbaiki reliabilitas instrumen, yaitu: (1) memperjelas konseptualisasi materi, (2) meningkatkan ketelitian pengukuran, (3) menggunakan indikator ganda, (4) menggunakan uji coba tes.

- a) Reliabilitas akan meningkat apabila ruang lingkup materi yang diujikan jelas batasannya. Tes dikembangkan berdasarkan beberapa indikator (kisi-kisi) yang telah ditentukan. Satu indikator terdiri atas

beberapa soal dan setiap soal hanya memiliki satu jawaban yang benar. Indikator yang baik akan saling mendukung dan tidak saling tumpang tindih (*overlapping*). Masing-masing indikator mempunyai ciri spesifik sehingga pertanyaan tidak ada yang diulang. Meskipun indikator sudah spesifik, pengembangan pertanyaan yang digunakan dalam pengukuran indikator juga tidak boleh berlawanan (+/-) dengan indikator lainnya.

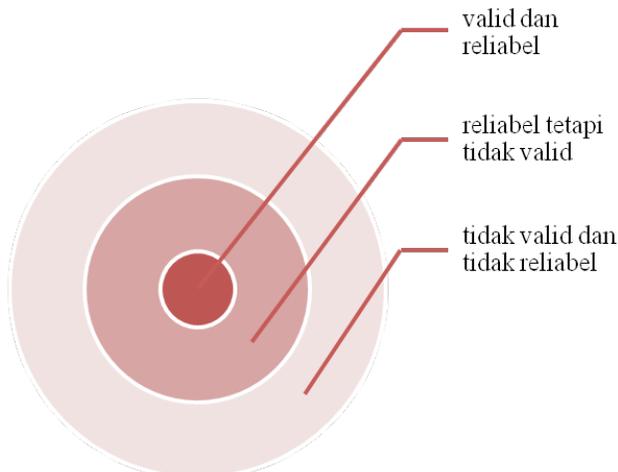
- b) Meningkatkan ketelitian pengukuran. Menulis soal, entri data dan analisis data harus dikakukan dengan teliti. Dari sisi konseptualnya, bacalah butir-butir pertanyaan dan pernyataan secara berulang-ulang dan cobalah untuk menjawab semua butir yang tertulis. Tanyakanlah pada diri sendiri, apakah semua pertanyaan/pernyataan yang tertulis tersebut mudah untuk dijawab, jelas kriterianya, dan jawaban yang dikehendaki juga cukup spesifik? Dari sisi entri data dan analisis data, telitilah data dengan seksama karena kesalahan pengukuran dan kesalahan *entry* (memasukkan) data ke dalam komputer dapat menyebabkan kesalahan hasil analisis data yang cukup fatal.
- c) Menggunakan indikator ganda. Soal sebaiknya ditulis untuk mengukur lebih dari satu indikator. Indikator ganda memungkinkan siswa untuk memahami isi dan konsep secara luas. Butir-butir soal lebih mudah dikembangkan apabila jumlah indikator yang digunakan cukup banyak. Jumlah butir yang lebih banyak cenderung memperoleh hasil pengukuran yang lebih stabil. Jumlah butir soal yang cukup banyak lebih mudah diatur, yaitu ketika respon butir kurang bagus (terlalu mudah atau terlalu sulit) maka butir soal tersebut dapat dibuang atau diganti tanpa mengurangi keluasan isi karena sudah terwakili oleh butir-butir lainnya.
- d) Menggunakan pretes, *pilot studies*/try out/uji coba dan replikasi. Reliabilitas dapat ditingkatkan dengan menggunakan dua kali pengukuran. Pada pengukuran pertama, peneliti dapat mencoba soal test tersebut sebelum menerapkan pada situasi yang sebenarnya. Karakteristik jawaban hasil uji coba soal test kemudian dianalisis dan diperbaiki kekurangannya. Butir butir soal yang baik adalah yang memiliki karakteristik jawaban bervariasi, tingkat kesulitan sedang (30%-70%) dan daya pembeda butir tinggi ($>0,3$). Apabila respon butir/jawaban sama, maka butir tersebut perlu dicek kembali untuk diperbaiki, diganti atau dibuang.

Tabel 5.4 Rangkuman Tipe-Tipe Pengukuran Reliabilitas dan Validitas

<i>Reliability (dependable measure)</i>	<i>Validity (true measure)</i>
<ul style="list-style-type: none"> • Stabilitas dari waktu ke waktu atau antar dua alat pengukuran (<i>test - retest</i> atau <i>parallel tes</i>) • Representatif antar sub populasi Diujikan pada sub populasi yang berbeda hasil tetap konsisten • Ekuivalen antar indikator (konsistensi internal, <i>split half</i>) 	<ul style="list-style-type: none"> • <i>Face</i> - tampang fisik instrumen menurut pertimbangan orang lain
	<ul style="list-style-type: none"> • <i>Content</i>- keseluruhan isi instrumen mewakili indikator yang diukur
	<ul style="list-style-type: none"> • <i>Criteria</i> – berasosiasi dengan alat ukur eksternal <ul style="list-style-type: none"> - <i>Concurrent</i>- berasosiasi dengan pengukuran sebelumnya - <i>Predictive</i>- berasosiasi dengan pengukuran berikutnya
	<ul style="list-style-type: none"> • <i>Construct</i>- konsistensi pengukuran antar indikator <ul style="list-style-type: none"> - <i>convergent</i> –serupa satu dengan yang lain - <i>discriminant</i> - berbeda satu dengan yang lain.

6) Hubungan antara Reliabilitas dan Validitas

Hubungan antara reliabilitas dan validitas dapat diilustrasikan pada gambar bantalan target anak panah berikut ini.



Gambar 5.10 Hubungan antara Validitas dan Reliabilitas

Keterangan gambar:

- A. Soal valid dan reliabel karena tepat berada di tengah (*center*) dan ketika pengukuran diulang beberapa kali, hasilnya tetap sama (reliabel) yaitu anak panah jatuh di titik yang berdekatan.
- B. Soal reliabel tetapi tidak valid karena selalu jatuh pada titik yang berdekatan tetapi tidak tepat mengukur materi yang diajarkan.
- C. Soal tidak valid dan tidak reliabel karena jatuh di luar materi yang seharusnya diukur, dan hasilnya selalu tidak konsisten.

7) Bias dalam Penilaian Pendidikan

Penilaian dapat menghasilkan beberapa keputusan, namun keputusan yang diambil belum tentu objektif atau bias atau tidak menunjukkan kondisi yang sebenarnya. Nitko (Nitko & Brookhart, 2011) mengidentifikasi beberapa penyebab bias yaitu perbedaan: rerata, fungsi butir (*differential item functioning*), kesalahan interpretasi, jenis kelamin dan ras, validitas, isi dan pengalaman, model statistik yang digunakan untuk pengambilan keputusan, kriteria pengukuran, suasana dan kondisi asesmen.

Beberapa contoh penyebab bias yang telah diidentifikasi tadi dapat dijelaskan sebagai berikut: (1) Hasil analisis data nilai rerata kelas A lebih rendah dari kelas B pada ujian Tata Hidang. Materi pembelajaran antara kelas A dan B sama, prosedur asesmen yang digunakan juga sama. Hal ini belum tentu membuktikan rerata kemampuan kelas A lebih rendah dari kelas B jika: (a) guru yang mengajar berbeda, (b) metode pembelajaran berbeda, (c) fasilitas belajar yang dimiliki peserta didik berbeda; (2) Bias perbedaan fungsi butir (*differential item functioning*) dapat dikaji jika dua individu yang memiliki kemampuan sama diuji dengan materi sama tetapi menunjukkan hasil yang berbeda. Hal ini dapat disebabkan oleh rasa senang/tidak senang terhadap butir soal/tugas yang diujikan. Siswa yang menyukai selada buah akan mendapat nilai lebih bagus jika dia mendapat tugas membuat salad; (3) Siswa laki-laki memiliki perbedaan kemampuan dengan siswa perempuan dalam beberapa jenis tes. Hal ini dapat menyebabkan kesalahan interpretasi terhadap skor. Bias hasil penilaian juga dapat bersumber dari perbedaan ras, budaya, dan jenis kelamin (*facial bias*). Siswa yang pernah tinggal di luar negeri jika di tes kemampuan bahasa asing akan lebih bagus karena sudah terbiasa menggunakan bahasa tersebut daripada siswa yang belum pernah ke luar negeri.

Asesmen seleksi dapat menghasilkan keputusan prediksi yang bias terhadap kesuksesan pekerjaan atau sekolah. Skor kriteria prediksi dapat bagus pada sekelompok orang tetapi belum tentu bagus pada kelompok lain. Berdasarkan beberapa hasil pengamatan, banyak skor tes seleksi yang tidak dapat memprediksi kesuksesan akademik atau pekerjaan. Hasil ujian nasional tahun 2007 yang digunakan untuk pengambilan keputusan kelulusan dan seleksi masuk ke jenjang SMP memiliki prediksi rendah ($<0,5$) terhadap prestasi akademik siswa setelah di SMP. Hasil analisis kasus ditemukan ada beberapa penyebab daya prediksi nilai ujian nasional rendah yaitu: siswa yang tidak jujur, soal terlalu mudah sehingga tidak bisa membedakan siswa yang pandai dan kurang pandai, kemampuan yang diuji berbeda dengan kemampuan yang diprediksi, siswa hanya belajar pada materi yang diujikan saja (Mulyatiningsih, (2008)

Suasana dan kondisi ujian yang kurang nyaman dapat menyebabkan hasil tes menjadi bias. Siswa yang memiliki tingkat kecemasan tinggi saat menghadapi ujian, kondisi fisik siswa kurang sehat dan lingkungan sekolah yang bising dapat menyebabkan hasil tes kurang maksimal.

DAFTAR PUSTAKA

- Aiken, L. R. (1991). *Psychological testing and assessment*. Boston: Allyn and Bacon.
- Aiken, L. R. (1996). *Tests Psicológicos y Evaluación*. México: Prentice Hall.
- Alkin, M. C., & Woolley, D. C. (1969, 10 8-11). A Model for Educational Evaluation,. *PLEDGE Conference*, hal. 2.
- American Culinary Federation. (2013, 3 14). *Assessment Blue Print, Culinary Arts Certification*. Diambil kembali dari https://www.nocti.org/PDFs/JobReady/4336_Culinary_Arts_Level-2.pdf
- Anderson, L. W., Krathwohl, D. R., & et all. (2001). *A taxonomy for learning, teaching and assessing: A revision of Bloom's taxonomy of educational objectives* . New York : Longman.
- Ani Faridah, d. (2008). *Patiseri Jilid 2 untuk SMK* . Jakarta: Direktorat Pembinaan Sekolah Menengah Kejuruan, Departemen Pendidikan Nasional,.
- Anonim. (2005, 12 12). *Performance Measurement and Evaluation: .:* Diambil kembali dari General Accounting Office (GAO). United States: <https://www.gao.gov/products/GAO-11-646SP>
- Atkinson, S. (2015, 7 9). *Graduate Competencies, Employability and Educational Taxonomies: Critique of Intended Learning Outcomes*. Diambil kembali dari Practice and Evidence of the Scholarship of Teaching and Learning in Higher Education: <http://community.dur.ac.uk/pestlhe.learning/index.php/pestlhe/article/view/194/281>
- Ayala, R. J., Plake., B. S., & Impara, J. C. (2001). The impact of omitted responses on the accuracy of ability estimation in item response theory. *Journal of Educational Measurement, Vol. 38, No. 3*, 213 234.

- Bachman, L. (1990). *Fundamental Considerations in Language Testing*. . Oxford: Oxford University Press.
- Bielinski, J., & Davison, M. L., M. L. (2001). A sex difference by item difficulty interaction in multiple choice mathematics items administered to national probability samples . *Journal of Educational Measurement, Vol. 38, No. 1*, 51-77.
- Bloom, B. S. (1976). *Human characteristic and school learning*. New York: McGraw-Hill book Company.
- Burden, P. R., & Byrd, D. M. (1998). *Methods for Effective Teaching*. Needham Height: Allyn & Bacon.
- DiMartino, J., & Castaneda, A. (2007). Assessing applied skills. *Educational Leadership, 60(7)*, 38-42.
- Direktorat PSMK. (2017). *Panduan penilaian hasil belajar pada Sekolah Menengah Kejuruan*. Jakarta: Direktorat Jenderal Pendidikan dan Dasar dan Menengah, Kementerian Pendidikan dan Kebudayaan.
- Eby, J. W. (1998). *Reflective Planning, Teaching, and Evaluation: K-12*. . United States of America: : Prentice-Hall.
- Gierl, M. J., Bisanz , G. L., & Boughton, K. A. (2003). Identifying content and cognitive skills that produce gender differences in mathematics: A demonstration of the multidimensionality-based DIF analysis paradigm. *Journal of Educational Measurement., 40, No. 4* , 281-306.
- Griffin, P., & Peter, N. (1991). *Educational assessment and reporting*. Sidney:: Harcourt Brace Javanovich Publisher.
- Hamilton, L. S. (2003). Studying Large Scale Reforms of Instructional Practice., An example from mathematics and science. . *Educational Evaluation and Policy Analysis, 25, No. 1*,, 1-29.
- Herrington, J., & Herrington, A. (1998). How university students respond to a model of authentic assessment. *Higher Education Research and Development, 17 (3)*, 305-322.
- Hinkle, D. E., Wiersma, W., & Jurs, S. G. (1979). *Applied statistics for behavioral sciences*. Boston: Houghton Mifflin Company.
- Iskandar, S. M. (2014). Pendekatan Keterampilan Metakognitif dalam Pembelajaran Sains di Kelas. *ERUDIO, , Vol. 2, No. 2, Desember 2014*, 13-20.

- Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment, A manageable and cooperative process*. Boston.: Allyn and Bacon.
- Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment, A manageable and cooperative process*. Boston: Allyn and Bacon.
- Kizlik, B. (2012, 99). *Measurement, Assessment, and Evaluation in Education*. Diambil kembali dari <http://www.adprima.com/measurement.html>
- Krathwohl, D., Bloom, B.S, B. S., & Masia, B. (1964). *Taxonomy of Educational Objectives: The Classification of Educational Goals, Handbook 2: Affective Domain*. . New York: McKay.
- Linn, R. L. (1989). *Educational measurement (3rd ed)*. New York: American Council on Education and Macmillan.
- Linn, R. L., Baker, E., & Dunbar, S. (1991). Complex, performance-based assessment: Expectations and validation criteria. *Educational Researcher*, , 20 (8), , 15-21.
- Lunz, M. E., & Bergstrom, . B. (1994). An empirical study of computerized adaptive test administration conditions. *Journal of Educational Measurement, Vol. 31, No 3*, 251 263.
- Madaus, G. F., & Stufflebeam, D L, D. I. (1985). *Systematic Evaluation, evaluation in education and human services*:. Massachusetts: Chestnut Hill.
- Mardapi, D. (2003). *Konstruksi Tes dan analisis butir. Bahan lokakarya Metodologi Interaksi Pembelajaran*. . Surakarta: : Universita Muhammadiyah Surakarta.
- McIntire, S. A., & Miller, I. A. (2000). *Foundation of psychological testing*. . Boston.: McGraw-Hill.
- Mehrens, W. A. (1973). *Measurement and evaluation in education and psychology*. . New York: Holt: Rinehart and Winston, Inc.
- Mehrens, W. A., & Lehman, I. (1973). *Measurement and evaluation in education and psychology*. . New York: Holt : Rinehart and Winston, Inc.
- Montero, D. H., Yulian, M. C., & Yen, W. M. (2003). Multiple measures: Alternative design analysis models. *Educational Measurement: Issues and Practice*, , 7 – 12.
- Moshman, & Schraw. (1995). Metacognitive Theories. *Educational Psychology Review, Vol 7, No 4*, 13.

- Muchtadi, T. R. (2011). *Prinsip, Proses dan Teknologi Pangan*. Bandung: Alfabeta.
- Mulyatiningsih, E. ((2008). Model evaluasi keberlanjutan SD ke SMP dalam rangka wajib belajar 9 tahun. *Jurnal Penelitian Kebijakan Pendidikan*.
- Naga, D. S. (1992). *Pengantar teori skor pada pengukuran pendidikan*. Jakarta : Gunadarma.
- Neuman, W. L. (2000). *Social research methods, qualitative and quantitative approaches 4th ed* . Boston : Allyn & Bacon.
- Nicola, S. E. (2004). (De)grading the standardized test: Can standardized testing evaluate school? . *Education in Canada*, 44(Iss 3), 37. Dipetik 5 5, 2005, dari <http://proquest.umi.com/pqdweb>
- Nitko, A. J., & Brookhart, S. M. (2011). *Educational assessment of student. 6th*. Boston : Pearson Education, Inc.
- Nitko, A. L., & Brookhart, S. M. (2011). *Educational assessment of students 6th edition* . Boston: Pearson.
- Noor, A. F. (2013). *Modul 1-4: Pengertian, hakekat, dan teori belajar dan pembelajaran*: Palangkaraya: Universitas Muhammadiyah Palangkaraya.
- Ogunniyi, M. B. (1984). *Educational Measurement and Evaluation*. Longman : Nig Mc. Ibadan.
- Okpalla P. M., e. a. (1999). *Measurement and Evaluation in Education*. Benin City: Stiching – Horden Publishers (nig.) Ltd.
- Olfos, R., & Zulantay, H. (2007). Reliability and validity of authentic assessment in a web based course. *Educational Technology & Society*, 10(4), , 156-173.
- Oshima, T. C. (1994). The effect of speediness on parameter estimation in item response theory. *Journal of Educational Measurement*.Vol 31, No.3, 200-219.
- Overton, T. (2011, May 11). *Assessing Learners with Special Needs: 6 th. Ed* . Diambil kembali dari <http://www.enotes.com/ref/q-and-a/distinction-between-assessment-evaluation-201131>
- Pitkin, A. K., & Vispoel, W. P. (2001). Differences between self adapted and computerized adaptive test: A meta analysis. *Journal of Educational Measurement.* , Vol. 38, No. 3,, 235 241.

- Raimundo, & Zulantay, H. (2007). Reliability and Validity of Authentic Assessment in a Web Based Course. *Educational Technology & Society*, 10 (4), 156-173.
- Roderick, M., Jacob, B. A., & Bryk, A. S. (2002). The impact of high stakes testing in Chicago on student achievement in promotional gate grades. *Educational Evaluation and Policy Analysis*, Vol. 24, No. 4, 333-357.
- Roid, G. H., & Haladyna, T. M. (1982). *A technology for test-item writing*. New York : Harcourt Brace Jovanovich.
- Seels, & Glasgow. (1990). *Exercises in instructional design*. . Columbus OH: Merrill Publishing Company.
- Simpson, J. S. (1996). *The classification of educational objectives, psychomotor domain*. Urbana, IL: University of Illinois: Office of Education Project No. 5-85-104.
- Suryabrata, S. (1981). *Pengukuran dan penilaian pendidikan, (Kumpulan makalah penataran bimbingan dan konseling untuk tenaga pengajar di PT se Indonesia, edt*. Jakarta: Dirjen Pendidikan Tinggi.
- Teitelbaum, P. (2003). The influence of high school graduation requirement policy in mathematic and science on student course taking patterns and achievement. *Educational Evaluation and Policy Analysis*, 25, No. 1., 3-157.
- Umar, Y. (1996). *Bahan penataran pengujian pendidikan*. Jakarta: Depdikbud.
- Wang, T. K. (2001). Evaluating comparability in computerized adaptive testing: Issues, Criteria and an example. *Journal of Educational Measurement*, Vol. 38, No. 1, 19-49.
- Widana, I. W. (2017). *Modul penyusunan soal higher order thinking skills (HOTS)*. . Jakarta: Direktorat Pembinaan SMA, Direktorat Jenderal Pendidikan Dasar dan Menengah Departemen Pendidikan dan Kebudayaan.
- Willingham, W. W., Pollack, B. M., & Lewis, C. (2002). Grades and test scores: accounting for observed differences. *Journal of Educational Measurement*, 39, No. 1, 1-37.
- Wollack, B. A., Cohen, A. S., & Wells, C. A. (2003). A method for maintaining scale stability in the presence of test speediness. *Journal of Educational Measurement*, Vol. 40, No. 4, 307-330.



ASESMEN PEMBELAJARAN KEJURUAN

BIDANG PARIWISATA & TATA BOGA



Buku ini ditulis untuk membantu pendidik (dosen dan guru) dalam melaksanakan tanggung jawab profesionalnya (professional responsibility) kepada peserta didik. Salah satu tanggung jawab profesi pendidik adalah memberi penilaian yang objektif kepada peserta didik. Jika penilaian dilakukan secara objektif diharapkan dapat memberi motivasi kepada peserta didik untuk memperoleh prestasi/nilai yang sebaik-baiknya dan selalu meningkatkan hasil belajarnya. Dalam buku ini dijelaskan tentang berbagai macam alat penilaian, cara penulisan, telaah soal secara teoritis maupun empiris.



ISBN : 978-602-5566-69-1



9 786025 556669 1

UNY Press

Jl. Gejayan, Gg. Alamanda, Komplek Fakultas Teknik UNY
Kampus UNY Karangmalang Yogyakarta 55281

Telp: 0274 - 589346

E-Mail: unypress.yogyakarta@gmail.com

Anggota Ikatan Penerbit Indonesia (IKAPI)

Anggota Asosiasi Penerbit Perguruan Tinggi Indonesia (APPTI)