

LAPORAN AKHIR HIBAH BERSAING



PENGEMBANGAN ALAT ASESMEN HASIL BELAJAR TATA HIDANG BERBANTUAN KOMPUTER

Tahun ke 2 dari rencana 2 tahun

Ketua:

**Prihastuti Ekawatiningsih, M. Pd
NIDN. 0028047504**

Anggota:

**Endang Mulyatiningsih NIDN. 0011016306
Sigit Yatmono NIDN. 0025017301**

**Dibiayai oleh DIPA Direktorat Penelitian Pengabdian kepada Masyarakat
Nomor DIPA – 023.04.1.673453/2015, tanggal 14 November 2014, DIPA
revisi 01 tanggal 03 Maret 2015. Skim: Penelitian Hibah Bersaing
Tahun Anggaran 2015 Nomor: 062/SP2H/PL/DIT.LITABMAS/II/2015
Tanggal 5 Februari 2015**

**UNIVERSITAS NEGERI YOGYAKARTA
OKTOBER, 2015**

HALAMAN PENGESAHAN

Judul : PENGEMBANGAN ALAT ASESMEN HASIL BELAJAR TATA HIDANG BERBANTUAN KOMPUTER

Peneliti/Pelaksana
Nama Lengkap : PRIHASTUTI EKAWATININGSIH
Perguruan Tinggi : Universitas Negeri Yogyakarta
NIDN : 0028047504
Jabatan Fungsional : Lektor Kepala
Program Studi : Tata Rias Dan Kecantikan
Nomor HP : 08122752922
Alamat surel (e-mail) : prihastuti@uny.ac.id

Anggota (1)
Nama Lengkap : ENDANG MULYATININGSIH
NIDN : 0011016306
Perguruan Tinggi : Universitas Negeri Yogyakarta


Anggota (2)
Nama Lengkap : SIGIT YATMONO M.T.
NIDN : 0025017301
Perguruan Tinggi : Universitas Negeri Yogyakarta
Institusi Mitra (jika ada) : -
Nama Institusi Mitra : -
Alamat : -
Penanggung Jawab : -
Tahun Pelaksanaan : Tahun ke 2 dari rencana 2 tahun
Biaya Tahun Berjalan : Rp 67.000.000,00
Biaya Keseluruhan : Rp 117.000.000,00

Mengetahui,
Dekan Fakultas Teknik



(Dr. M. Bruri Triyono)
NIP/NIK 195602161986031003

Yogyakarta, 14 - 10 - 2015
Ketua,



(PRIHASTUTI EKAWATININGSIH)
NIP/NIK 197504281999032002

Menyetujui,
Ketua LPPM UNY



(Prof. Dr. Anik Ghufron)
NIP/NIK 196211111988031001

RINGKASAN

Dunia pendidikan menghadapi banyak tantangan akibat perkembangan teknologi informasi berbasis komputer yang sangat pesat. Di masa depan, semakin banyak dibutuhkan perangkat pembelajaran yang dioperasikan dengan komputer. Untuk menjawab tantangan tersebut, pada penelitian ini dikembangkan perangkat pembelajaran yaitu alat asesmen hasil belajar Tata Hidang yang dioperasikan menggunakan komputer. Kegiatan penelitian tahun pertama telah berhasil mengembangkan bank soal ujian Tata Hidang dan merancang program computerized test menggunakan *software wondershare*. Penelitian tahun kedua bertujuan untuk mengimplementasikan penyelenggaraan ujian Tata Hidang berbantuan komputer dan menganalisis validitas *concurrent* hasil ujian menggunakan komputer dengan hasil ujian *pencil and paper test*.

Penelitian ini merupakan penelitian dan pengembangan (R&D) yang menggunakan model ADIE (*analysis, design, implementation, and evaluation*). Pada tahun pertama, penelitian telah sampai tahap *analysis* dan *design*. Penelitian tahun kedua difokuskan pada kegiatan *implementation, and evaluation*. Subjek uji coba hasil rancangan soal ujian Tata Hidang adalah 100 orang mahasiswa yang mengambil mata kuliah Tata Hidang. Metode pengumpulan data menggunakan dokumentasi respon butir soal ujian Tata Hidang. Analisis data respon butir soal ujian menggunakan Analisis Iteman sedangkan validitas *concurrent* menggunakan analisis korelasi *product moment*.

Hasil penelitian menunjukkan bahwa untuk pengembangan alat penilaian masih ada beberapa revisi antara lain: dengan penambahan suara pada kompetensi I untuk pembacaan soalnya. Disamping itu ada beberapa tampilan yang diperbaiki berkaitan dengan lay out program. Program komputer yang digunakan untuk ujian Tata Hidang secara *on line* atau *computerized test* adalah *software Wondershare*. Pada saat merancang program telah disusun langkah-langkah pembuatan untuk programmer dan langkah-langkah pengoperasian pengguna. Telaah kuantitatif menunjukkan bahwa 100 butir soal *multiple choice* sudah diujikan kepada mahasiswa dengan bantuan program komputer. Telaah kualitas butir soal ujian Tata Hidang secara kuantitatif telah membuktikan 73% soal berada pada tingkat kesulitan sedang, 52% butir soal memiliki daya pembeda sangat baik dan 28% soal memiliki daya pembeda yang baik. Nilai reliabilitas soal sebesar 0,945 telah memenuhi kriteria soal yang baik, dengan angka validitas konkuren 0,167.

Kata kunci: *asesmen hasil belajar, tata hidang, computerized test*

PRAKATA

Puji Syukur kami panjatkan kehadirat Allah SWT yang telah memberikan rahmat dan hidayah-Nya, sehingga laporan tahunan Penelitian Hibah Bersaing yang berjudul Pengembangan Alat Asesmen Hasil Belajar Tata Hidang Berbantuan Komputer dapat diselesaikan.

Tim peneliti merasa laporan penelitian ini tidak akan terwujud tanpa bantuan pihak-pihak terkait. Dalam kesempatan ini peneliti memberikan penghargaan dan terima kasih yang setinggi-tingginya kepada Kementerian Riset, Teknologi dan Pendidikan Tinggi, Direktur Riset dan Pengabdian Masyarakat, Rektor Universitas Negeri Yogyakarta, Dekan Fakultas Teknik, Para praktisi dari Industri Boga yang telah bersedia untuk melakukan FGD secara mendalam, serta rekan-rekan sejawat yang telah terlibat baik secara langsung maupun tidak langsung dalam kegiatan penelitian ini.

Akhirnya semoga laporan hasil penelitian ini bisa memberikan manfaat khususnya bagi kegiatan pembelajarannya, sehingga dapat meningkatkan kualitas pembelajaran di Program Studi Pendidikan Teknik Boga khususnya dan dunia pendidikan pada umumnya.

Yogyakarta, 31 Oktober 2015
Tim Penyusun

Prihastuti Ekawatiningsih, M.Pd.
Dr. Endang Mulyatiningsih
Sigit Yatmono, M.T.

DAFTAR ISI

HALAMAN SAMPUL	i
HALAMAN PENGESAHAN	ii
RINGKASAN	iii
PRAKATA	iv
DAFTAR ISI	v
DAFTAR TABEL	vi
DAFTAR GAMBAR	vii
DAFTAR LAMPIRAN	viii
BAB 1. PENDAHULUAN	1
A. Latar Belakang Masalah	1
B. Rumusan Masalah	3
C. Urgensi Penelitian	3
D. Luaran Produk	4
BAB 2. TINJAUAN PUSTAKA	5
A. State of The Art	5
B. Asesmen Hasil Belajar	6
C. Pengembangan Alat Asesmen	9
D. Kualitas Butir Tes	10
E. Penafsiran Hasil Tes	13
F. Kompetensi Tata Hidang	14
G. Peta Jalan (Roadmap) Penelitian	16
BAB 3. TUJUAN DAN MANFAAT PENELITIAN	17
A. Tujuan Penelitian	17
B. Kegunaan Hasil Penelitian	17
BAB 4. METODE PENELITIAN	18
A. Jenis Penelitian	18
B. Prosedur Penelitian	18
C. Target Luaran Per Tahun	20
D. Sumber Data Penelitian	20
E. Metode Pengumpulan Data	21
F. Instrumen Penelitian	21
G. Metode Analisis Data	22
BAB 5. HASIL PENELITIAN DAN PEMBAHASAN	24
A. Hasil Analisis Penelitin	24
B. Pembahasan	31
BAB 6. KESIMPULAN DAN SARAN	33
A. Kesimpulan	33
B. Saran	34
DAFTAR PUSTAKA	35
LAMPIRAN	36

DAFTAR TABEL

	Halaman
Tabel 1	Sumber Data Penelitian 20
Tabel 2	Kisi-Kisi Soal Ujian Tata Hidang 24
Tabel 3	Indeks Tingkat Kesulitan Butir 25
Tabel 4	Kriteria Indeks Daya Pembeda Butir 25
Tabel 5	Interpretasi Hasil Analisis Korelasi Butir Soal 26

DAFTAR GAMBAR

	Halaman
Gambar 1	Interrelasi Mengajar dan Belajar, Evaluasi dan Asemen..... 7
Gambar 2	Mekanisme Analisis Soal..... 10
Gambar 3	Validitas Kriteria..... 12
Gambar 4	Diagram Alir Proses Penelitian Tahun Pertama..... 18
Gambar 5	Diagram Alir Proses Penelitian Tahun Kedua..... 19
Gambar 6	<i>Fishbone</i> Kegiatan Penelitian..... 23
Gambar 7	Tampilan Awal untuk Masuk sebagai User..... 28
Gambar 8	Tampilan Awal untuk Masuk sebagai Peserta Ujian..... 29
Gambar 9	Tampilan Hasil Ujian..... 29
Gambar 10	Tampilan Review Kunci Jawaban..... 30
Gambar 11	Alur Pelaksanaan Peserta Ujian Menggunakan Perangkat Komputer..... 30

DAFTAR LAMPIRAN

Lampiran 1.	Contoh Kisi-Kisi dan Soal Tes Kognitif Tata Hidang....	38
Lampiran 2.	Hasil Analisis Iteman Soal Tata Hidang On line.....	41
Lampiran 3.	Hasil Analisis Uji Validitas.....	57
Lampiran 4.	Personalia Tim Peneliti dan Kualifikasinya.....	61
Lampiran 5.	Publikasi (Draft Buku Evaluasi Pembelajaran).....	62
Lampiran 6.	Publikasi (Artikel Jurnal).....	110

BAB I

PENDAHULUAN

A. Latar Belakang Masalah

Hasil belajar diukur dan dinilai dengan menggunakan berbagai alat asesmen (penilaian). Hasil belajar dinyatakan bagus jika hasil asesmen menunjukkan nilai yang tinggi. Namun hasil asesmen yang bagus belum tentu menunjukkan kemampuan belajar yang bagus karena banyak alat asesmen dan asesornya kurang objektif dalam memberikan penilaian. Untuk mengatasi hambatan tersebut, pada saat ini mulai banyak dikembangkan alat-alat asesmen hasil belajar berbantuan komputer. Dengan menggunakan alat ini, hasil belajar diuji dengan seperangkat tes yang dapat ditampilkan oleh komputer, peserta tes mengerjakan soal tes langsung di komputer dan komputer juga langsung memberikan skor. Hasil tes dapat dilihat oleh peserta tes pada saat itu juga.

Hasil belajar Tata Hidang dinilai dari tiga ranah tujuan pembelajaran yaitu pengetahuan, sikap dan keterampilan. Penilaian hasil belajar Tata Hidang dilakukan dua kali yaitu melalui ujian tulis dan ujian praktik. Penilaian hasil belajar ranah pengetahuan menggunakan ujian tulis dengan metode *paper and pencil test* sedangkan penilaian hasil belajar ranah keterampilan menggunakan uji kompetensi praktik. Penilaian hasil belajar Tata Hidang dengan menggunakan cara tersebut sangat melelahkan dosen sehingga pada masa yang akan datang perlu dikembangkan alat asesmen baru yang berbasis pada teknologi elektronik berbantuan komputer (*E-assessment*).

Penilaian hasil belajar menggunakan komputer sudah banyak dikembangkan pada bidang ilmu lain. Salah satu inovasi penilaian hasil belajar yang dikembangkan menggunakan teknologi komputer adalah *Computerized Adaptive Testing (CAT)*. Pada CAT, soal tes dibuat dalam beberapa level kemampuan yang disesuaikan dengan kemampuan peserta tes. Peserta tes yang memiliki kemampuan tinggi mendapat peluang untuk melanjutkan ke beberapa butir tes yang memiliki kesulitan tinggi. Menurut hasil penelitian, CAT dapat menghasilkan tingkat validitas dan reliabilitas yang sama dengan tes konvensional yang menggunakan *paper and pencil test* (Reynold, C. R., 2010). Alat penilaian hasil belajar keterampilan berbantuan

komputer juga telah banyak dikembangkan. Beberapa contoh alat penilaian keterampilan berbantuan komputer yang sudah digunakan misalnya pada simulasi ujian untuk mendapat SIM (surat ijin mengemudi), simulasi ujian keterampilan pilot. Tanpa bantuan komputer, keterampilan tersebut sulit dilatihkan karena beresiko tinggi terhadap keselamatan manusia. Hal ini sejalan dengan pengembangan tes untuk CPNS dan bahkan untuk Ujian Nasional mulai tahun 2014 akan dilakukan secara on line. Dengan demikian pengembangan alat penilaian ini menjadi pedoman atau panduan yang dapat diimplementasikan untuk pengembangan tes on line mata kuliah lain.

Alat asesmen/penilaian hasil belajar Tata Hidang berbantuan komputer diharapkan dapat membantu sebagian tugas dosen. Dengan alat penilaian berbantuan komputer tersebut, hasil penilaian akan semakin obyektif. Dosen yang masih ragu-ragu pada saat memberi nilai dapat membandingkan hasil penilaian yang telah diberikan dengan hasil penilaian dari komputer. Selain itu, alat penilaian berbantuan komputer juga dapat digunakan sebagai alat untuk penilaian diri sendiri (*self assessment*) bagi mahasiswa. Mahasiswa dapat menggunakan simulasi alat asesmen/penilaian berbantuan komputer untuk menguji kemampuannya.

Penyelenggaraan tes berbantuan komputer memiliki beberapa keunggulan antara lain: (1) dapat mengurangi biaya cetak soal tes; (2) mengurangi waktu dan tenaga dosen untuk mengoreksi jawaban peserta tes karena komputer langsung dapat memberikan skor tes; 3) penyelenggaraan tes lebih efisien karena semua perangkat tes dapat di *download* dari internet sehingga dapat dikerjakan dari jarak jauh. Meskipun demikian, ada sedikit kelemahan dari penyelenggaraan tes menggunakan komputer yaitu harus tersedia fasilitas komputer dan jaringan internet jika tes akan diselenggarakan secara *on- line*. Dalam penilaian *on line*, perlu dirancang sistem yang tidak memberi peluang pengisian jawaban kuis dan tugas dilakukan oleh orang lain atau bukan siswa yang belajar materi kuliah Tata Hidang terdiri dari materi kuliah teori dan praktik. Materi kuliah teori terdiri dari sub materi yang berisi pengenalan macam-macam perlengkapan peralatan Tata Hidang beserta fungsi dan cara penyiapannya, etika makan, dan struktur organisasi restoran, dan lain-lain. Materi pelajaran praktik Tata Hidang terdiri dari sistem pelayanan makan secara

American Service, Russian Service dan *Buffet Service*. Materi praktis mengandung unsur Standar Operasional Prosedur yang jelas. Materi kuliah praktik maupun teori Tata Hidang keduanya berpotensi dikembangkan menjadi bank soal yang penyimpanan dan pengoperasionalnya dilakukan menggunakan komputer.

B. Rumusan Masalah Penelitian

Penelitian dirancang selama dua tahun. Masalah penelitian yang diatasi pada masing-masing tahun adalah sebagai berikut:

Tahun pertama:

1. Apa saja ruang lingkup materi Tata Hidang yang layak dikembangkan menjadi kisi-kisi soal tes pilihan ganda?
2. Bagaimanakah hasil telaah kualitas butir soal tes secara kualitatif dan kuantitatif yang diuji coba dengan teknik *paper and pencil test*?
3. Bagaimanakah rancangan program komputer yang layak digunakan untuk ujian Tata Hidang secara *on line* atau *computerized test*?

Tahun kedua:

1. Bagaimanakah persepsi peserta tes terhadap soal tes Tata Hidang secara *on line* menggunakan perangkat komputer?
2. Bagaimanakah kualitas butir soal tes Tata Hidang bentuk pilihan ganda yang diuji cobakan secara *on line*?
3. Bagaimanakah validitas dan reliabilitas soal tes Tata Hidang yang diuji cobakan secara *on line*?
4. Apakah hasil tes Tata Hidang secara manual (*paper and pencil test*) memiliki validitas konkuren dengan tes menggunakan perangkat komputer?

C. Urgensi Penelitian

Pengembangan alat asesmen/penilaian berbantuan komputer penting dilakukan untuk memanfaatkan teknologi informasi yang semakin canggih. Dengan alat asesmen/penilaian berbantuan komputer ini, model pembelajaran dapat diinovasi menjadi *blended learning*. Pada masa yang akan datang, *blended*

learningmodel's diprediksi akan menjadi model pembelajaran yang paling dibutuhkan oleh masyarakat modern. *Blended learning* yaitu perpaduan antara fasilitas belajar *online*, tatap muka dan memberi pengalaman belajar langsung melalui proses penemuan atau penelitian.

Alat asesmen/penilaian berbantuan komputer penting dikembangkan untuk memfasilitasi mahasiswa yang mengambil kuliah jarak jauh. Pada saat ini, jarak tempuh kuliah mungkin masih terjangkau oleh mahasiswa sehingga proses pembelajaran dapat dilakukan dengan tatap muka. Di masa depan, dengan kepadatan arus lalu lintas, mahasiswa akan lebih nyaman jika belajar dan ujian dari rumah. Dengan memfasilitasi pembelajaran jarak jauh, masyarakat yang berada di daerah terpencil juga mendapat peluang untuk mengakses pendidikan. Alat asesmen/penilaian berbantuan komputer ini menjadi alat asesmen alternatif yang layak untuk digunakan.

Alat asesmen menggunakan bantuan komputer untuk menilai hasil belajar masih relatif baru karena sampai saat ini belum banyak peneliti yang mengembangkannya. Alat asesmen berbantuan komputer yang sudah banyak dikembangkan adalah alat asesmen yang digunakan pada tes seleksi, TOEFL, *e-learning*, dsb. Alat tes seleksi berbantuan komputer memiliki kelebihan karena peserta dapat mengikuti tes dari mana saja, tanpa batasan jarak. Hasil tes berbantuan komputer sangat objektif dan langsung dapat diketahui oleh peserta tes saat itu juga.

D. Luaran yang di targetkan

1. Tahun ke-1: Bank soal ujian Tata Hidang
2. Tahun ke-2: Software Alat Penilaian Kognitif Program *Wondershare Quiz Creator* dan Modul/buku: "Evaluasi Pembelajaran Kejuruan"

BAB II

TINJAUAN PUSTAKA

A. State of the Art

Alat asesmen menggunakan bantuan komputer sudah mulai dikembangkan di beberapa negara. Berikut diperoleh beberapa hasil penelitian penggunaan alat tes berbantuan komputer.

1. Clauser, et.al (2000) pernah mengembangkan *software* evaluasi kinerja di bidang kedokteran. Kompetensi yang diuji adalah keterampilan menangani pasien. Soal tes berupa permasalahan yang dihadapi pasien. Peserta ujian merespon dengan berbagai tindakan penanganan medis. Komputer secara otomatis memberi skor berdasarkan transaksi tindakan yang telah dilakukan peserta ujian. Berdasarkan hasil analisis ternyata skor diperoleh dari penilai ahli (manusia) mempunyai korelasi yang cukup tinggi dengan skor yang dibuat oleh komputer.
2. Sebuah penelitian yang dilakukan oleh Lunz, & Bergstrom (1994: 251-263) tentang evaluasi penyelenggaraan CAT (*computerized adaptive test*). Hasil penelitian menunjukkan bahwa pengkondisian CAT berpengaruh pada hasil tes. Setelah dicoba dengan empat kondisi 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. Secara berturut-turut hasil tes akan lebih baik jika peserta tes diberi kesempatan memilih soal yang bisa terlebih dulu (*skip*), kemudian mengoreksi kembali jawabannya (*review*), dan menunda untuk menjawab (*defer*). Hasil yang paling buruk adalah jika peserta tes dibiarkan tanpa kondisi.
3. Hasil penelitian lain yang dilakukan oleh Oshima (1994: 200-219) menemukan kondisi tes kecepatan dapat mempengaruhi estimasi butir dan parameter kemampuan. Hasil mengindikasikan bahwa di bawah kondisi kecepatan, ada korelasi tinggi antara jawaban benar dan parameter estimasi kemampuan.
4. Piktin and Vispoel (2001) memperoleh temuan bahwa SAT (*Scholastic Aptitude Test*) memiliki reliabilitas lebih rendah dibanding CAT karena seleksi butir CAT

menggunakan algoritma yang dapat menurunkan kesalahan pengukuran secara maksimal. Sekitar 43% penyelenggaraan SAT membutuhkan waktu yang lebih lama daripada CAT pada saat tes dikerjakan bersamaan.

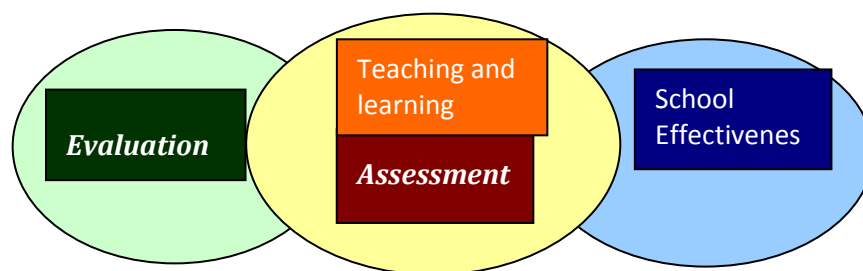
5. CISCO mengantisipasi perubahan di masa depan dengan mengembangkan berbagai macam produk yang berbasis jaringan komputer, salah satunya adalah *networking academy*. Semua komponen program *Networking Academy* dilakukan oleh komputer seperti *instructor-led classroom sessions, web-based course content, interactive learning activities and tools, online assessments, hands-on labs with real equipment, and innovative simulations*.

Beberapa hasil penelitian yang terdahulu menunjukkan bahwa tes yang menggunakan bantuan komputer lebih baik dan efisien digunakan. Selama ini, perangkat tes berbantuan komputer lebih banyak pada tes kognitif sedangkan tes keterampilan (*skills*) masih relatif sedikit. Hal ini disebabkan karena tidak semua materi tes keterampilan dapat disimulasikan dengan komputer. Oleh sebab itu, dalam pengembangan tes keterampilan menggunakan komputer perlu diidentifikasi secara matang. Kriteria unjuk kerja yang benar dan salah juga perlu dirumuskan dengan tegas supaya komputer dapat membuat skor yang benar. Pengukuran hasil belajar kognitif oleh komputer dapat dilakukan jika tes berbentuk pilihan ganda. Pada tes ini harus dipastikan hanya ada satu jawaban yang benar supaya komputer dapat memberi skor yang benar.

B. Asesmen Hasil Belajar

Istilah pengukuran (*measurement*), asesmen (*assessment*), dan evaluasi (*evaluation*) menjadi satu rangkaian kegiatan yang tidak dapat dipisahkan. Huit (2001: 1) menjelaskan bahwa *assessment* mengacu pada pengumpulan data untuk memahami sebuah isu, objek secara lebih baik. *Measurement* adalah proses mengkuantitatifkan data *assessment*, dan *evaluation* menekankan pada perbandingan antara data yang ada dengan standar yang telah ditetapkan. Penelitian menggunakan data *assessment, measurement* dan *evaluation* untuk tujuan deskripsi, prediksi, dan pengendalian supaya fenomena yang terjadi dapat dipahami dengan lebih baik lagi.

Griffin (1992: 5) memberi contoh penggunaan istilah *assessment*, *measurement* (pengukuran) 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. Asesmen dapat pula berarti pengumpulan informasi tentang kualitas atau kuantitas perubahan dalam diri siswa, kelompok, guru dan tenaga administrasi yang sedang diamati, diukur dan dinilai (Johnson, 2002: 1). *Interrelasi* antara asesmen, proses belajar mengajar, evaluasi dan efektivitas sekolah dapat dilihat pada Gambar 1.



Gambar 1: *Interrelasi* Mengajar dan Belajar, Evaluasi dan Asesmen Sumber: Conner (1991: 10)

Gambar 1 mengisyaratkan bahwa kegiatan belajar mengajar, asesmen, evaluasi dan efektivitas sekolah merupakan kegiatan yang saling terkait. Kegiatan belajar mengajar dan asesmen merupakan basis data evaluasi yang digunakan untuk mengetahui efektivitas sekolah. Tanpa ada kegiatan belajar mengajar yang dapat diukur dan dinilai maka evaluasi tidak dapat dilakukan. Evaluasi ini penting untuk membuat keputusan tentang efektivitas sekolah atau dalam lingkup yang lebih sempit yaitu keputusan tentang efektivitas pembelajaran.

Pengukuran dan asesmen dalam pendidikan mempunyai tiga fungsi yaitu: (1) melayani kebutuhan psikologis siswa 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 diperlukan untuk membantu guru mengetahui tingkat pencapaian tujuan pembelajaran dan membantu guru mengetahui posisi siswa secara individual (Burden & Byrd, 1998: 332).

Mehrens (1973: 378) mendefinisikan macam-macam tes berdasarkan tujuannya. Tes hasil belajar bertujuan untuk mengevaluasi dampak belajar, evaluasi mengajar, evaluasi kurikulum, *diagnosis*, kelulusan, dan motivasi. Tes bimbingan (*guidance*) bertujuan untuk mengadakan bimbingan kejuruan, pendidikan, dan personal. Tes sekolah bertujuan untuk seleksi, klasifikasi, penempatan, perencanaan dan evaluasi kurikulum, evaluasi guru, dan penyediaan data/informasi untuk penelitian. Semua jenis tes dapat digunakan untuk tujuan bimbingan kejuruan dan penelitian.

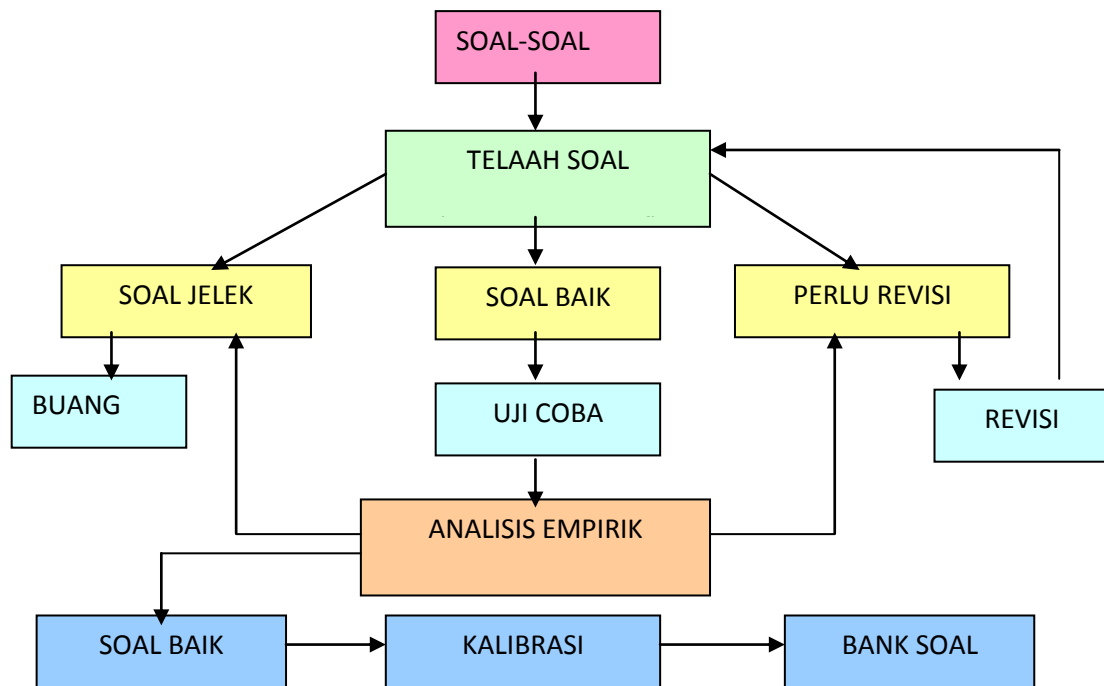
Pembuatan perangkat tes baku perlu memperhatikan karakteristik khusus tes seperti konstruk/domain/isi yang akan diukur, bentuk pertanyaan, bentuk penyelenggaraan dan bentuk penyekoran tes. Konstruk tes dikembangkan dari komponen materi atau lebih dikenal dengan kisi-kisi tes. Hasil tes yang dapat diberi skor objektif adalah tes dalam bentuk pilihan ganda/MC (*multiple choice*), menjodohkan atau benar-salah (*true-false*). Rodriguez (2003: 163-184) menemukan korelasi tinggi antara tes dengan bentuk *constructed – response* (jawaban singkat) dan *multiple-choice* apabila dimensi yang diukur dan panjang tes sama. Lukhele (1994: 234-250) menemukan *constructed – response (CR)* yang baik dapat mengukur kemampuan siswa secara lebih mendalam tetapi sulit diberi skor yang objektif oleh penilai. Butir CR sering digunakan untuk mengukur beberapa proses kognitif yang hasilnya lebih akurat. Butir MC yang terlalu mudah atau terlalu sulit dapat menghasilkan estimasi kemampuan siswa yang kurang tepat. Jika butir soal dengan bentuk jawaban MC terlalu sulit dapat memberi kemungkinan kepada siswa untuk menjawab pertanyaan hanya dengan cara menebak-nebak (*guessing*). Oleh sebab itu, tes dengan bentuk jawaban MC yang baik adalah tes yang memiliki tingkat kesulitan sedang.

C. Pengembangan Alat Asesmen (Tes)

Pengembangan tes dilakukan melalui proses yang cukup panjang. Tes yang baik mencakup pengetahuan, keterampilan, kemampuan dan karakteristik yang lain (KSAOs = *knowledge, skill, abilities and other characteristics*) yang dituntut untuk ditunjukkan. McIntire (2000) menetapkan 10 langkah pengembangan tes yang harus dilalui dalam proses pengembangan tes 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*
- 9) *Developing norms*
- 10) *Complete test manual*

Puslitbang Sisjian (1996: 12) menerbitkan sebuah buku panduan untuk melengkapi kegiatan pengembangan tes dengan petunjuk analisis soal yang dapat dilihat pada Gambar 2. Diagram alir pada Gambar 2 mengindikasikan soal yang berkualitas baik memerlukan prosedur pengembangan yang panjang. Soal-soal yang akan diujikan harus melewati telaah kualitatif dari ahlinya. Setelah soal diuji coba, kemudian dilanjutkan dengan kegiatan telaah soal melalui analisis kuantitatif yang berisi kegiatan pembuktian validitas, reliabilitas dan analisis respon butir. Soal-soal yang sudah teruji kualitasnya dapat langsung dianalisis sesuai dengan kebutuhan dan soal-soal yang kurang baik diperlakukan melalui dua cara yaitu dibuang atau direvisi. Soal-soal yang sudah mengalami pengujian berkali-kali dan sudah memenuhi kriteria kualitas soal yang baik pada umumnya dapat dibakukan untuk memperkaya bank soal.



Gambar 2. Mekanisme Analisis Soal

Sumber: Yahya Umar, Puslitbang Sisjian (1996: 35)

D. Kualitas Butir Tes

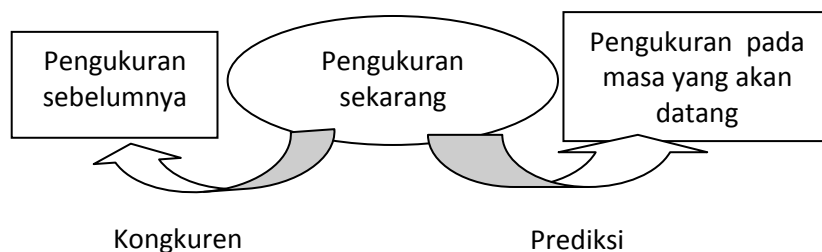
Analisis kualitas butir tes secara kuantitatif dilakukan dengan menganalisis respon (jawaban) butir tes dari peserta tes. Substansi yang dianalisis meliputi tingkat kesukaran tes, daya pembeda tes, reliabilitas, dan pengukuran kesalahan baku (*standard error measurement*). Analisis berdasarkan teori respon butir dapat diterapkan hanya pada butir yang mempunyai karakteristik jawaban *local independent* atau *unidimensi* (Hambleton, 1991: 13). Dalam teori tes klasik, kesulitan (*difficulty*) masing-masing butir dapat dihitung dengan menemukan persentase peserta tes yang menjawab butir dengan benar. Kebanyakan pengembang tes mencari daerah kesulitan pada tingkat rata-rata sekitar 0,5. Pengembang tes juga membuat sebuah indeks daya pembeda (*discrimination*) berdasarkan hasil analisis korelasi point biserial.

Output analisis respon butir dapat digunakan untuk membantu proses pembuatan keputusan. Contoh: matrik korelasi inter-item memberikan beberapa informasi konsistensi internal tes. Masing-masing butir yang mempunyai korelasi tinggi dengan setiap butir lain menunjukkan butir tersebut mengukur konstruk yang sama. Dalam teori respon butir, kinerja masing-masing butir berhubungan dengan kemampuan peserta tes pada konstruk yang diukur. Kesalahan estimasi dalam menginterpretasikan hasil pengukuran dapat terjadi apabila tes yang dikembangkan mengandung bias. Bias butir terjadi apabila sebuah butir lebih mudah untuk satu kelompok dan sulit bagi kelompok lainnya. Bias butir dapat ditelusuri dengan cara membandingkan skor antar beberapa kelompok atau membandingkan kinerja kelompok dengan kriteria eksternal.

Proses validasi tes dapat dilakukan dengan beberapa cara tergantung pada tujuannya. Validitas isi digunakan untuk memvalidasi tes prestasi dan validitas konstruk digunakan untuk menjelaskan kesesuaian konstruk/indikator pada teori yang sedang dianalisis. Messick (1989) yang dikutip dari Kane (2006: 19) menganjurkan bahwa pembuktian validitas isi tidak dapat menggunakan data empiris, karena skor tes tidak menunjukkan kesesuaian isi tes dengan materi yang telah diajarkan. Anastasi (2003: 86) menjelaskan bahwa pada dasarnya validitas isi adalah pengujian sistematis isi tes untuk menentukan apakah tes tersebut cukup representatif mengukur domain perilaku yang hendak diteliti. Sebuah tes hasil belajar dapat memenuhi persyaratan validitas isi, jika butir-butir soal yang dikembangkan mewakili seluruh materi yang diajarkan.

Setelah validitas isi terpenuhi, soal tes kemudian diukur berdasarkan validitas konstruk dan validitas kriterianya. Validitas konstruk digunakan untuk pengukuran indikator ganda. Pengukuran dinyatakan valid apabila indikator-indikator yang bervariasi memiliki korelasi yang tinggi dengan skor gabungan dari semua indikator. Validitas konstruk menuntut ada batasan-batasan konseptual yang spesifik. Validitas konstruk dikatakan konvergen apabila indikator ganda saling berasosiasi antara satu dengan yang lain. Apabila pengukuran tidak konvergen, maka indikator yang digunakan untuk mengkonstruk variabel tidak dapat digabungkan menjadi satu dalam pengukuran variabel tersebut (Neuman, 2003: 168). Validitas kriteria ditemukan jika

skor tes berkorelasi dengan kriteria lain di luar materi tes yang diujikan. Ada dua jenis validitas kriteria yaitu validitas prediksi dan validitas *concurrent*. Skor tes memiliki validitas prediksi yang tinggi jika skor peserta tes saat ini memiliki korelasi yang tinggi terhadap skor peserta tes pada pengukuran kemampuan berikutnya (*longitudinal*). Skor tes memiliki validitas *concurrent*, jika skor peserta tes yang diperoleh saat ini konsisten atau berkorelasi tinggi dengan skor peserta tes yang sudah diperoleh sebelumnya. Untuk memperoleh data pengukuran validitas *concurrent* dapat dilakukan dengan metode pengambilan data *cross sectional*. Konsep validitas kriteria diilustrasikan pada gambar 3.



Gambar 3: Validitas Kriteria

Validitas kriteria merupakan validitas yang selalu dikaitkan dengan kriteria eksternal yang dijadikan dasar pegujian skor tes. Ada dua validitas kriteria yaitu validitas prediktif dan validitas kongkuren(*concurrent*).Skor tes memiliki validitas prediksi yang tinggi jika skor peserta tes saat ini memiliki korelasi yang tinggi terhadap skor tes pada pengukuran kemampuan berikutnya (*longitudinal*). Skor tes memiliki validitas *concurrent*, jika skor peserta tes yang diperoleh saat ini konsisten atau berkorelasi tinggi dengan skor tes yang sudah diperoleh sebelumnya. Untuk memperoleh data pengukuran validitas *concurrent* dapat dilakukan dengan metode pengambilan data *cross sectional*. Koefisien korelasi antara skor tes dan skor kriteria menunjukkan sejauhmana kesesuaian antara hasil ukur tes dengan hasil ukur tes lain yang sudah teruji kualitasnya.

Reliabilitas

Istilah reliabel dapat diartikan tetap atau konstan. Reliabilitas mengukur kemampuan instrumen untuk menghasilkan data yang mendekati sama bila instrumen tersebut digunakan berulang-ulang pada objek yang sama dan dengan cara

yang sama. Analisis reliabilitas selalu dikaitkan dengan konsistensi pengukuran, yaitu bagaimana hasil pengukuran tetap (konstan) dan konsisten dari satu pengukuran ke pengukuran yang lain.

Pengujian reliabilitas suatu instrumen dapat dikerjakan secara internal dan eksternal. Pengujian reliabilitas secara internal (*internal consistency*) berkaitan dengan analisis konsistensi butir-butir yang ada dalam instrumen dengan cara membagi satu set butir-butir pertanyaan menjadi dua bagian yang sama, bisa dengan cara membagi butir awal dan akhir atau ganjil dan genap. Pengujian reliabilitas secara eksternal dapat dilakukan melalui analisis tes ulang (*test-retest*) berkaitan dengan stabilitas tes (*stability*), tes paralel (*parallel test*). *Test-retest* diterapkan dengan cara melakukan pengujian pada kelompok orang yang sama, dengan instrumen yang sama tetapi waktu yang berbeda. Apabila jawaban peserta test relatif sama, maka instrumen dinyatakan reliabel. Tes paralel dilakukan dengan dua perangkat instrumen yang menggunakan indikator sama tetapi susunannya setara yang diujikan pada orang yang sama. Instrumen dinyatakan reliabel apabila menunjukkan hasil pengukuran yang sama, meskipun perangkat tesnya berbeda atau paralel.

Dengan ketiga metode tersebut, yaitu metode tes ulang, tes paralel dan konsistensi internal, akan menghasilkan taksiran koefisien reliabilitas yang berbeda. Koefisien reliabilitas yang sebenarnya adalah sulit untuk dapat diamati sehingga yang diperoleh hanyalah koefisien reliabilitas taksiran. Tes hasil belajar dinyatakan reliabel apabila koefisien reliabilitasnya (α) > 7 (Nunnally: 1987)

E. Penafsiran Hasil Tes

Skor baris (raw score) hasil tes merefleksikan kinerja siswa atau mencerminkan nilai yang diperoleh siswa pada saat menyelesaikan tes. Untuk memahami kinerja siswa maka skor tes perlu diinterpretasikan. Cara yang sederhana untuk menginterpretasikan hasil tes adalah dengan membuat skor baris (a raw score). Menurut Reynold (2010) *a raw score is simply the number of items scored or coded in a specific manner such as correct/incorrect, true/false, and soon*. Pengertian tersebut dapat diartikan sebagai berikut: skor baris adalah jumlah skor butir atau

kode yang secara khusus digunakan untuk menetapkan jumlah respon butir benar/salah. Contoh: apabila siswa menjawab 70 butir soal dengan benar dari 100 butir soal yang diujikan maka siswa akan memperoleh skor baku sebesar 70. Apabila siswa menjawab 10 butir soal dengan benar dari 40 butir soal yang diujikan, maka 25% jawaban siswa benar atau jika diberi nilai menggunakan skor baku (10 – 100) maka dia akan memperoleh skor 25.

F. Kompetensi Tata Hidang

Kompetensi merupakan sasaran evaluasi hasil belajar. Kompetensi diartikan sebagai 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 berkompeten 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.

Kompetensi yang ingin dicapai pada mata kuliah Tata Hidang adalah mahasiswa mampu mengajarkan cara melayani tamu yang benar dan mampu menjadi tamu yang beretika. Tata Hidang memiliki banyak tata cara yang sudah diatur dengan standard operational procedure (SOP). Tata cara yang dipelajari dalam mata kuliah Tata Hidang adalah:

- 1) Tata cara dalam *food and baverages service* (pelayanan makanan dan minuman) yang meliputi: (a) *Greeting the guest*; (b) *escuting and sitting the guest*; (c) *pouring ice water*; (d) *serving bread and butter*; (e) *presenting the menu/taking order*; (f) *presenting the wine list*; (g) *adjustment*; (h) *serving the food*; (i) *serving the wine*; (j) *clear up*; (k) *crumbing down*; (l) *presenting coffee or tea*; (m) *presenting the bill*; (n) *bid forewell*; (o) *table setting*
- 2) SOP *After Operation* (restoran tutup) antara lain: inventaris peralatan, *sauce*, dan *condiment*, mencatat pendapatan, menghitung pengeluaran, membuat laporan tertulis, mematikan *electricity* yang tidak diperlukan dan menutup pintu restoran

seperti semula. Pramusaji harus melakukan: (a) menyimpan menu, *wine list*, *beverages list* dalam rak; (b) menyimpan *sauce* di atas *side stand* dengan rapi; (c) menyingkirkan sisa-sisa makanan dari area restoran; (d) merapikan meja, kursi, serta *service cart* yang masih berantakan; (e) menghubungi *houseman* untuk membersihkan karpet dan lantainya; (f) melaporkan kepada supervisor apabila terjadi complaint, saran serta pujian dari tamu; (g) mengupayakan perbaikan di hari yang akan datang; (h) memeriksa kembali laci, pintu apakah sudah terkunci dengan baik, sebelum meninggalkan restoran.

- 3) Pengetahuan menu yang akan dihidangkan karena menu menentukan jenis alat, tata cara menghidangkan atau tata cara makan. Susunan menu klasik terdiri dari 13 macam makanan. Menu modern disusun lebih sederhana dari menu klasik yaitu terdiri dari 3–5 hidangan. Dalam sistem pelayanan makanan Formal, menu dihidangkan satu persatu sesuai dengan giliran makan mulai dari makanan pembuka sampai makanan penutup.
- 4) Etika Makan dan *Table Manner*. Dalam perjamuan makan, pelayan dan tamu dituntut memiliki sopan santun yang berlaku dalam suatu perjamuan makan. Pada saat di tempat jamuan makan, ada beberapa etiket yang perlu diketahui yaitu: (1) perempuan dipersilahkan duduk terlebih dahulu (*lady first*) dan masuk ke tempat duduk dari sebelah kiri kursi. Posisi tubuh saat duduk tetap tegak dengan posisi tangan menggantung dan pergelangan tangan diletakkan dimeja; (2) buka dan letakkan napkin di atas pangkuan; (3) gunakan alat makan mulai dari alat yang terletak disebelah luar; (4) menyuap makanan sedikit demi sedikit; (5) mengunyah makanan dengan mulut tertutup; (6) suap makanan mendekati mulut bukan sebaliknya; (7) makanan ditelan terlebih dahulu sebelum berbicara; (8) jangan mengajak berbicara dengan tamu yang sedang menyuap; (9) pada waktu bicara jangan menunjuk sesuatu dengan menggunakan alat makan; (10) minuman tidak boleh digunakan untuk berkumur; (11) jika bermaksud menolak tambahan minuman yang disajikan waiter, letakkan jari telunjuk di atas bibir gelas atau cangkir; (12) jika disela-sela waktu makan tamu akan minum, bersihkan mulut dengan napkin terlebih dahulu agar gelas tidak kotor; (13) potong makanan dari garpu dengan pisau bukan dengan tangan; (14) jika ingin

membersihkan makanan yang terselip di gigi, tutuplah mulut dengan serbet makan; (15) jika ada tulang atau benda keras lainnya ikut termakan, tutup mulut dan ambil dengan jari atau sendok dan letakkan di piring; (16) jika menemukan benda asing atau rambut di dalam makanan, tetapkan diam agar tidak mengganggu tamu lain.

G. Peta Jalan Penelitian (*road-map*)

1. Kegiatan yang telah dilaksanakan

Peneliti telah mengembangkan media pembelajaran interaktif. Judul penelitian dan hasil yang telah dicapai: “Pengembangan Media Pembelajaran Interaktif Tata Hidang (2003) telah digunakan untuk media pembelajaran di program studi Pendidikan Teknik Boga. Kegiatan penelitian dilakukan berkolaborasi dengan beberapa dosen dan mahasiswa dari program studi Teknik Informatika. Peneliti membuat desain konseptualnya kemudian dibantu oleh programmer komputer. Anggota tim peneliti juga memiliki *track record* dalam mengembangkan media pembelajaran interaktif dengan perangkat tes elektronik di dalamnya.

2. Kegiatan yang saat ini sedang dikerjakan

Penelitian “Pengembangan Alat Asesmen Tata Hidang Berbantuan Komputer (2014) yang sedang dilaksanakan masih relevan dengan rekam jejak peneliti yaitu studi S2 pada program studi Penelitian dan Evaluasi Pendidikan (PEP), mata kuliah yang diampu dan beberapa penelitian yang mendukung terciptanya alat bantu pembelajaran komputer, khususnya untuk Mata Kuliah Tata Hidang.

3. Kegiatan yang akan dilaksanakan

Setelah sukses mengembangkan alat asesmen pembelajaran Tata Hidang, rencana penelitian adalah pengembangan alat asesmen Tata Hidang berbantuan komputer. Dengan bantuan media, mahasiswa lebih mudah mengingat kembali tata cara melayani maupun tata cara makan yang merupakan materi dasar kompetensi Tata Hidang. Disamping itu juga pada tahap implementasi, akan dilakukan uji implementasi (uji coba) dalam skala luas berkaitan dengan penggunaan (Alat Penilaian berbantuan Komputer kepada Mahasiswa yang mengambil Mata Kuliah Tata Hidang).

BAB III

TUJUAN DAN MANFAAT PENELITIAN

A. Tujuan Penelitian

Konsisten dengan masalah penelitian yang diatasi, penelitian ini bertujuan untuk:

Tahun pertama:

1. Menganalisis ruang lingkup materi Tata Hidang yang layak dikembangkan menjadi kisi-kisi soal tes pilihan ganda.
2. Menelaah kualitas butir soal tes Tata Hidang secara kualitatif dan kuantitatif yang diujikan melalui teknik *paper and pencil test*.
3. Merancang program komputer yang layak digunakan untuk ujian Tata Hidang *on line*.

Tahun kedua:

1. Mengetahui persepsi peserta tes terhadap uji coba ujian Tata Hidang secara *on line* menggunakan perangkat komputer.
2. Mengetahui kualitas butir soal tes Tata Hidang bentuk pilihan ganda yang diuji cobakan secara *on-line*.
3. Mengetahui validitas dan reliabilitas soal ujian Tata Hidang diuji cobakan secara *on-line*.
4. Menguji validitas konkuren hasil tes Tata Hidang secara manual dengan hasil tes menggunakan perangkat komputer.

B. Kegunaan Hasil Penelitian

Hasil penelitian diharapkan dapat digunakan bagi:

1. Dosen Tata Hidang dapat menggunakan hasil penelitian ini sebagai alat bantu dalam mengevaluasi hasil belajar. Dosen mata kuliah lain dapat mengembangkan alat asesmen serupa untuk mengevaluasi hasil belajar pada mata kuliah yang diajarkannya.
2. Mahasiswa dapat melatih kemampuan dengan mengerjakan kuis-kuis latihan dan simulasi uji kompetensi dengan program komputer.

BAB IV METODE PENELITIAN

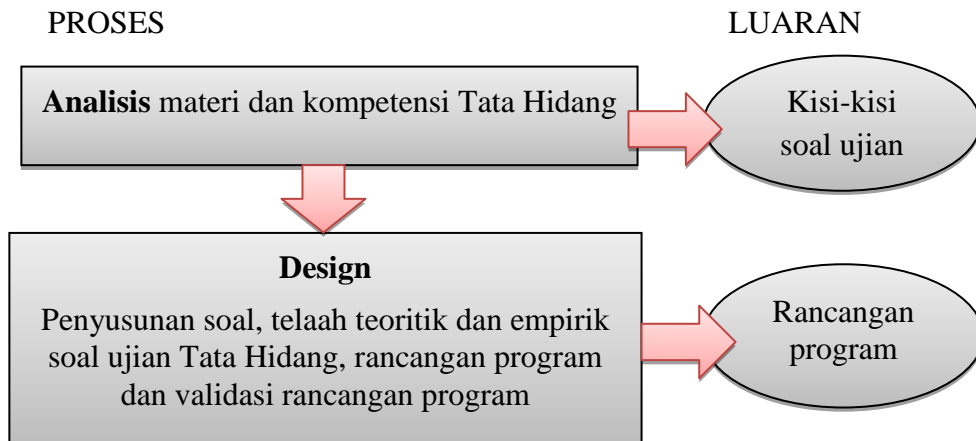
A. Jenis Penelitian

Pengembangan alat asesmen berbantuan komputer dilakukan dengan metode penelitian dan pengembangan. Model pengembangan mengacu pada model ADIE yang merupakan akronim dari *analysis, design, implementation, evaluation* (Dick and Carey: 1996). Kegiatan **penelitian tahun pertama** telah dilakukan pada tahap analisis dan desain. Kegiatan **penelitian tahun kedua** dilakukan pada tahap implementasi dan evaluasi. Kegiatan penelitian terjadi pada saat implementasi yaitu dengan mengambil data respon butir soal ujian Tata Hidang. Kegiatan **pengembangan tahun kedua** dilakukan melalui pengujian dan revisi *software wondershare* sampai software tersebut layak digunakan.

B. Prosedur Penelitian

Tahap-tahap penelitian dan pengembangan yang lebih lengkap dilustrasikan pada gambar 4 dan keterangan gambar berikut ini:

1. Tahun pertama:



Gambar 4. Diagram Alir Proses Penelitian Tahun Pertama

Keterangan:

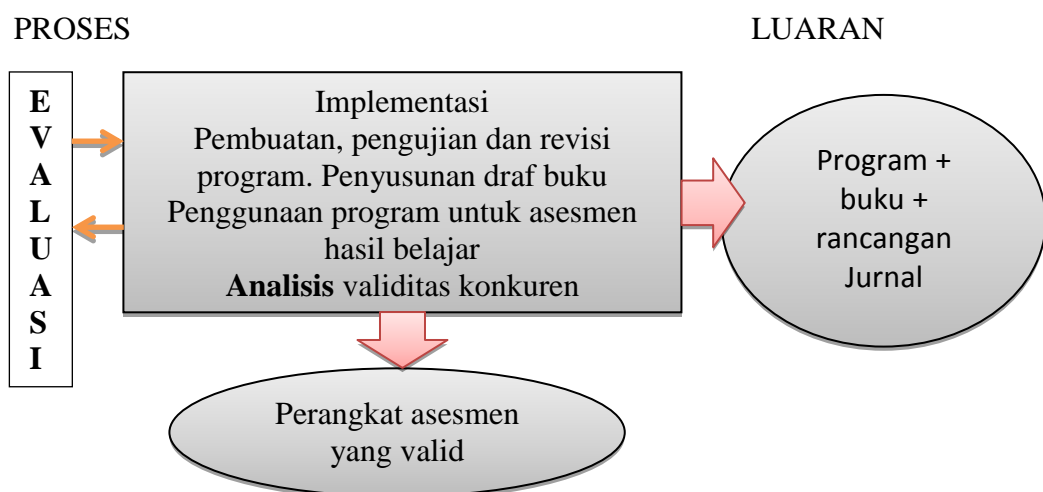
a. Analysis

Pada tahap awal penelitian dilakukan analisis materi dan indikator pencapaian kompetensi dasar Tata Hidang. Ruang lingkup materi Tata Hidang terdapat pada silabus dan bahan ajar. Berdasarkan hasil analisis materi dan kompetensi kemudian dipilih materi Tata Hidang yang dapat diuji dengan tes bentuk pilihan ganda.

b. Design

Setelah berhasil diidentifikasi materi dan kompetensi Tata Hidang, pada tahap perancangan program ini dilakukan: (1) penyusunan soal ujian Tata Hidang; (2) validasi isi rancangan soal ujian Tata Hidang; (3) merancang software wondershare; (4) validasi rancangan program.

2. Tahun kedua



Gambar 5. Diagram Alir Proses Penelitian

Keterangan:

c. Implementation

Implementasi dimulai dari membuat program sesuai rancangan yang telah ditulis, menguji coba dan implementasi program dalam ujian semester. Program diujicoba sebelum digunakan untuk ujian Tata Hidang yang diselenggarakan dengan bantuan komputer. Hasil uji coba digunakan untuk memperbaiki content soal maupun programnya. Implementasi alat asesmen (soal ujian) berbantuan komputer dilakukan bertepatan dengan waktu ujian semester.

d. Evaluation

Evaluasi dilakukan bersamaan dengan tahap implementasi. Materi yang dievaluasi meliputi kelayakan program dan isi program. Uji kelayakan program dilakukan terhadap tampilan gambar dan kemudahan akses. Uji kelayakan isi dilakukan terhadap tingkat kesulitan, daya pembeda, pengecoh, validitas isi dan konstruk, reliabilitas soal dan validitas konkuren.

C. Target Luaran PerTahun

Mengingat panjangnya proses penelitian, maka kegiatan dibagi menjadi dua periode pencapaian target luaran.

1. **Pada tahun pertama**, kegiatan dilakukan untuk menyusun bank soal, perancangan program, dan validasi rancangan alat asesmen hasil belajar Tata Hidang.
2. **Pada tahun kedua**, dilakukan kegiatan pembuatan, pengujian program, implementasi program, analisis validitas konkuren hasil penilaian komputer dengan nilai Tata Hidang Konvensional. Pada tahun kedua ini sudah mulai disusun buku ajar "**Evaluasi Pembelajaran Kejuruan**".

D. Sumberdata Penelitian

Tabel 1. Sumber Data Penelitian

Kegiatan	Sumberdata	Jumlah	Metode
TAHUN PERTAMA			
Analisis materi	Silabus dan bahan ajar	8 KD	Dokumentasi
Validasi rancangan	Ahli asesmen, materi, programmer komputer, desain grafis	15 orang	Delphi dan FGD
TAHUN KEDUA			
Uji coba program	Mahasiswa Pendidikan Teknik Boga dan Teknik Boga	100 orang	Tes, wawancara, observasi dan dokumentasi
Implementasi program	Mahasiswa Pendidikan Teknik Boga dan Teknik Boga	100 orang	

Rangkuman proses pengambilan data pada tabel 1 menunjukkan ada berbagai sumber data yang dilibatkan dalam penelitian ini. **Pada tahun kedua**, kegiatan penelitian adalah uji coba dan implementasi program, sumber data atau subjek penelitian yang terlibat adalah mahasiswa Pendidikan Teknik Boga dan Teknik

Boga. Subjek uji coba program adalah mahasiswa yang sudah lulus ujian Tata Hidang dan implementasi program adalah mahasiswa yang sedang ujian semester mata kuliah Tata Hidang. Subjek penelitian minimal 100 orang yang terdiri dari mahasiswa S1 dan D3 Teknik Boga tahun akademik 2014/2015.

E. Metode Pengumpulan Data

Pada Tabel 1, tahun kedua ada tiga metode pengumpulan data yang digunakan yaitu tes, wawancara, observasi dan dokumentasi. Metode tes dilakukan sekaligus untuk memperoleh data dokumen respon butir peserta tes yang diperlukan untuk menguji kualitas tes secara kuantitatif. Observasi dan wawancara dilakukan untuk mengamati dan menggali respon peserta tes pada saat mengerjakan ujian dengan komputer.

F. Instrumen Penelitian

Penelitian ini merupakan penelitian pengembangan alat asesment. Instrumen dalam penelitian ini adalah alat asesmen (soal ujian) Tata Hidang yang dikembangkan dan lembar validasi soal ujian. Penyusunan kisi-kisi soal ujian merupakan bagian dari proses pengembangan soal. Alat asesmen (soal ujian) dikendalikan kualitasnya dengan cara mengecek validitas dan reliabilitas tes. Proses validasi instrumen dilakukan dengan beberapa cara yaitu validitas isi, validitas butir dan validitas kriteria. Anastasi (2003: 86) menjelaskan bahwa pada dasarnya validitas isi adalah pengujian sistematis isi tes untuk menentukan apakah tes tersebut cukup representatif mengukur domain perilaku yang hendak diteliti. Setelah validitas isi terpenuhi, soal tes kemudian diukur berdasarkan validitas butir dan validitas kriterianya. Validitas butir dilakukan dengan cara mengkorelasikan skor butir dengan skor total. Validitas kriteria dilakukan untuk membandingkan skor tes yang menggunakan komputer dan skor tes manual menggunakan paper and pencil test.

Reliabilitas tes digunakan untuk mengetahui kestabilan alat ukur/tes ketika tes tersebut digunakan untuk mengukur hasil belajar mahasiswa. Reliabilitas memiliki dua pengertian yaitu konsistensi internal dan konsistensi eksternal. Konsistensi internal berkaitan dengan keandalan butir soal dalam seperangkat tes cukup

konsisten untuk mengukur kemampuan siswa. Konsistensi eksternal bermakna pada kemampuan siswa yang cukup konsisten meskipun diukur dengan alat ukur dan waktu pengukuran yang berbeda. Pembuktian reliabilitas tes dalam penelitian ini menggunakan pendekatan konsistensi internal karena soal hanya diujikan satu kali. Koefisien reliabilitas (Alpha) dihitung dengan rumus Cronbach's Alpha, yaitu:

$$\alpha = \frac{K}{(K-1)} \left(1 - \frac{\sum \sigma_i^2}{\sigma_x^2} \right)$$

Keterangan:

K = jumlah butir dalam tes

σ_i^2 = varian butir soal i

σ_x^2 = varian tes total

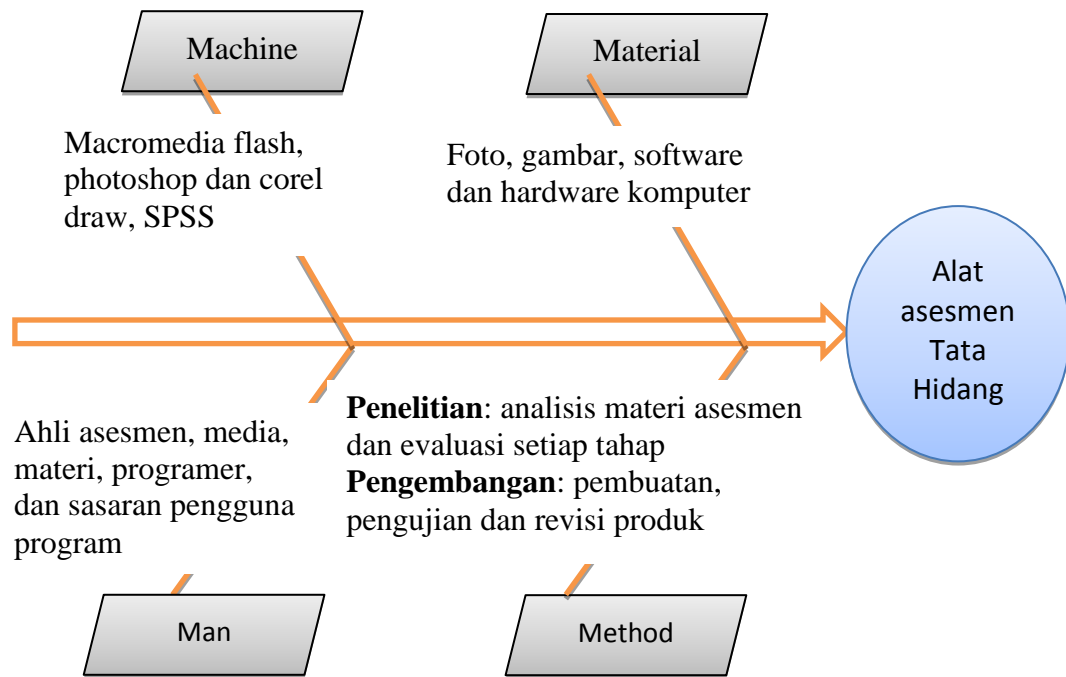
Analisis kualitas tes juga dilakukan dengan teori respon butir. Materi yang dapat dianalisis antara lain tingkat kesukaran tes, daya pembeda tes, reliabilitas, dan pengukuran kesalahan baku (*standard error measurement*). Analisis berdasarkan teori respon butir dapat diterapkan hanya pada butir yang mempunyai karakteristik jawaban *local independent* atau *unidimensi* (Hambleton, 1991: 13).

Dalam teori tes klasik, kesulitan (*difficulty*) masing-masing butir dihitung dengan menemukan persentase peserta tes yang menjawab butir dengan benar. Kebanyakan pengembang tes mencari daerah kesulitan pada tingkat rata-rata sekitar 0,5. Pengembang tes juga membuat sebuah indeks daya pembeda (*discrimination*). Statistik bekerja dengan cara membandingkan kinerja masing-masing butir yang membuat seseorang memiliki skor sangat tinggi dengan seseorang yang memiliki skor sangat rendah (Puslitbang Sisjian, 1996: 39-40). Pembuktian reliabilitas tes menggunakan program IteMan secara otomatis dikeluarkan dari hasil analisis sehingga peneliti tinggal menginterpretasikannya saja. Kriteria soal/variabel dinyatakan reliabel apabila mempunyai koefisien Alpha > 0,70 (McIntire, 2000: 122).

G. Metode Analisis Data

Metode analisis data dilakukan dua tahap yaitu telaah butir secara empiris dan pengujian validitas konkuren. Telaah butir secara empiris dilakukan dengan bantuan Program IteMan untuk menganalisis tingkat kesulitan, daya pembeda, pengecoh, reliabilitas soal ujian. Analisis validitas butir, konstruk dan konkuren

dilakukan dengan bantuan program SPSS. Metode analisis data yang digunakan adalah analisis korelasi *Product Moment* dari Karl Pearson dan korelasi biserial dari Spearman. Rangkuman metode penelitian dalam *fishbone diagram* tertera pada gambar 6, berikut ini:



Gambar 6. *Fishbone* Kegiatan Penelitian

BAB IV HASIL DAN PEMBAHASAN

A. Hasil Penelitian

1. Ruang lingkup kisi-kisi materi Tata Hidang

Bank soal Tata Hidang yang telah berhasil disusun dan dikembangkan berisi dari 100 butir soal pilihan ganda (MC/*multiple choice*). Soal ditulis dalam booklet. Kisi-kisi soal Tata Hidang dapat disimak pada tabel berikut ini:

Tabel 2. Kisi-kisi Soal Ujian Tata Hidang

No	Materi Dasar Kompetensi	Jumlah butir	Bentuk soal
1.	Ruang Lingkup Tata Hidang	10	MC
2.	<i>Personalia Food and Beverage Service</i>	15	MC
3.	Pengetahuan Menu Tata Hidang	15	MC
4.	Perlengkapan Tata Hidang	15	MC
5.	Persiapan Area Restoran	10	MC
6.	Etika dalam Jamuan Makan	10	MC
7.	Sistem Pelayanan Makanan	10	MC
	Jumlah	100	

2. Hasil telaah kuantitatif butir soal tes secara kuantitatif (hasil ujian dengan komputer)

Setelah melalui penelaahan (validasi) isi butir soal secara kualitatif, soal ujian Tata Hidang kemudian digunakan untuk ujian akhir semester. Respon jawaban ujian dianalisis secara kuantitatif menggunakan program analisis butir untuk mengetahui kualitas butir tes dari aspek tingkat kesulitan butir, daya pembeda butir, validitas butir dan reliabilitasnya. Hasil analisis butir secara kuantitatif dilaporkan berikut ini:

Tabel 3. Indeks Tingkat Kesulitan Butir (p)

Kategori	Proporsi benar	Nomor Butir	Jumlah
Mudah	$p > 0,7$	24, 25, 31, 52, 62	5
Sedang	$0,3 \leq p \leq 0,7$	1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 26, 27, 28, 29, 30, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 54, 55, 56, 58, 60, 63, 65, 66, 70, 74, 76, 77, 80, 81, 82, 83, 84, 85, 88, 89, 90, 91, 92, 93, 94, 95	73
Sulit	$p < 0,3$	51, 53, 57, 59, 61, 64, 69, 71, 72, 73, 75, 78, 79, 86, 87, 96, 97, 98, 99, 100	20
Salah		67, 68	2

Kriteria soal yang berkualitas baik adalah jika memiliki indek tingkat kesulitan butir pada kategori sedang ($0,3 \leq p \leq 0,7$). Hasil telaah butir secara kuantitatif menunjukkan soal ujian Tata Hidang sudah menunjukkan kualitas yang baik dengan proporsi 73% butir soal berada pada tingkat kesulitan sedang.

Selain indeks kesulitan butir, soal yang berkualitas baik harus memenuhi kriteria daya pembeda butir yang baik. Butir soal tes yang baik harus dapat membedakan kemampuan peserta tes diantara peserta tes lainnya. Soal tes yang sulit hanya dapat dijawab benar oleh peserta tes yang pandai. Hasil analisis daya pembeda butir dapat disimak pada tabel berikut ini:

Tabel 4. Kriteria Indeks Daya Pembeda Butir (D/r_{pb})

Kategori	Indeks pembeda	Nomor Butir	Jumlah
Sangat baik	$D \geq 0,4$	2, 4, 8, 10, 14, 16, 17, 20, 23, 28, 29, 30, 31, 36, 37, 39, 45, 46, 48, 49, 55, 56, 58, 59, 61, 62, 64, 65, 66, 69, 70, 71, 72, 73, 74, 76, 77, 79, 80, 81, 82, 84, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, 100	52
Baik, tanpa revisi	$0,3 \leq D \leq 0,39$	1, 5, 6, 7, 12, 13, 18, 19, 22, 24, 26, 27, 32, 33, 34, 35, 38, 41, 42, 43, 47, 52, 54, 60, 66, 70, 83, 85, 88, 89, 92	28
Perbatasan atau perlu revisi	$0,2 \leq D \leq 0,29$	3, 9, 11, 15, 21, 25, 35, 40, 63, 86, 87	11
Dibuang atau diganti	$D \leq 0,19$	21, 25, 35, 40, 44, 50, 51, 53, 57, 64	9

Hasil analisis menunjukkan terdapat 11% soal perlu direvisi dan 9% soal perlu diganti. Setelah dicermati lebih mendalam, soal yang harus direvisi antara lain karena jawaban ujian memiliki pengecoh dan banyak peserta tes yang terkecoh memilih jawaban tersebut. Hasil analisis ini ditindak lanjuti dengan mereview soal ujian butir demi butir. Dilihat berdasarkan rerata skor ang diperoleh pada tes dengan pepper & pencil sebesar 7,01 sedangkan rerata skor pada perolehan tes dengan komputer sebesar 8,02. terjadi kenaikan sebesar 1,01 terhadap rerata perolehan peserta tes. Hal ini menunjukkan alat esesmen yang digunakan cukup efektif dapat meningkatkan penguasaan kompetensi peserta tes, lebih menarik dan tidak membosankan.

a. Validitas

Pengujian validitas butir dilakukan dengan metode analisis korelasi biserial. Koefisien korelasi hasil analisis iteman diklasifikasikan menjadi beberapa kategori. Menurut Hinkle (1979: 85), kriteria yang digunakan untuk menginterpretasikan koefisien korelasi adalah sebagai berikut:

Tabel 5. Interpretasi Hasil Analisis Korelasi Butir Tes

Range $r \pm$	Interpretasi	Nomor butir	Jumlah
0,9 – 1	Korelasi sangat tinggi, positif atau negatif	61, 68 , 69, 79, 82, 84	6
0,7 – 0,9	Korelasi tinggi, positif atau negatif	2, 10, 16, 39, 55, 56, 59, 72, 73, 74, 76, 77, 91, 94, 96,	15
0,5 – 0,7	Korelasi sedang, positif atau negatif	1, 4, 8, 14, 17, 20, 23, 24, 26, 28, 29, 30, 31, 36, 37, 41, 45, 46, 48, 49, 58, 62, 65, 66, 71, 80, 81, 87, 88, 90, 93, 95, 97, 98. 100	35
0,3 – 0,5	Korelasi rendah, positif atau negatif	3, 5, 6, 7, 12, 13, 18, 19, 22, 27, 32, 33, 34, 35, 38, 42, 43, 47, 52, 54, 60, 70, 83, 85, 86, 89, 92,	27
0,0 – 0,3	Ada korelasikecil positif atau negatif	9, 11, 15, 21, 25, 40, 44, 50, 51, 53, 57, 63, 64, 67, 75, 78, 99	17

Berdasarkan klasifikasi koefisien korelasi yang terdapat pada tabel 5 terdapat 17 butir masih memiliki validitas yang rendah. Hasil analisis validitas butir ini kemudian di *cross cek* dengan validitas konstruk yaitu dengan cara mengkorelasikan kumpulan skor setiap kompetensi dasar dengan skor total. Berdasarkan hasil analisis validitas konstruks diperoleh soal sudah dinyatakan valid.

b. Reliabilitas

Reliabilitas soal tes Tata Hidang dianalisis menggunakan pendekatan konsistensi internal karena soal hanya diujikan satu kali. Koefisien reliabilitas (Alpha) dihitung dengan rumus Alpha Cronbach's. Hasil analisis reliabilitas soal menggunakan program iteman memperoleh koefisien Alpha sebesar 0,945. Kriteria soal yang baik adalah memiliki koefisien Alpha $>0,6$ (Nunally, 1978) oleh sebab itu soal telah memiliki kriteria reliabilitas yang ditetapkan.

3. Pengembangan dan Implementasi *Computerized Test*

Pengembangan alat asesmen berbantuan program komputer yang menjadi media penyelenggaraan tes. Program komputer dikembangkan menggunakan *software Wondershare*. Soal Tata Hidang tipe jawaban pilihan ganda yang telah memenuhi kriteria baik kemudian diisikan ke dalam *software Wondershare*. Langkah-langkah yang dilakukan untuk mengikan soal ke dalam *software Wondershare* adalah sebagai berikut:

- a) Membuat *User ID field* dan *password field* beserta tombol untuk melanjutkan proses.
- b) Membuat *NIM field*, *First name field* dan *Last Name field* beserta tombol untuk melanjutkan proses.
- c) Memasukkan soal dan jawaban dengan bentuk pilihan ganda.
- d) Merekam suara *dubbing* untuk media pembacaan soal.
- e) Memasukkan gambar/media lain untuk media penjelas soal/jawaban.
- f) Mengidentifikasi dan mengelompokkan soal berdasarkan Kompetensi Dasar dan tingkat kesulitan.
- g) Menerbitkan soal ke dalam bentuk *Flash*.
- h) Instalasi *Flash Player* untuk memutar soal yang telah diterbitkan ke dalam bentuk *Flash*.

Soal ujian yang sudah dimasukkan ke dalam *software Wondershare* telah siap dioperasikan. Ujian Tata Hidang dengan *software Wondershare* diujicobakan pada tahun kedua ini. Fasilitas soal ujian Tata Hidang berbantuan komputer telah dilengkapi dengan *Audio Dubbing* (diisi suara) pembacaan soal untuk soal pada urutan Kompetensi Dasar 1. Soal ujian Tata Hidang hanya bisa diakses oleh mahasiswa yang telah memiliki *User ID* dan *password*. Tampilan Soal ujian Tata Hidang dapat disimak pada gambar berikut ini:



Gambar 7. Tampilan awal untuk untuk masuk sebagai user

Tampilan pertama sebelum peserta tes/ujian memulai mengerjakan soal-soal Tata Hidang, peserta harus mengisi username dan password. Soal ujian hanya bisa diakses oleh mahasiswa yang telah memasukkan identitas dan passwordnya. Langkah berikutnya dapat disimak pada gambar berikut ini:



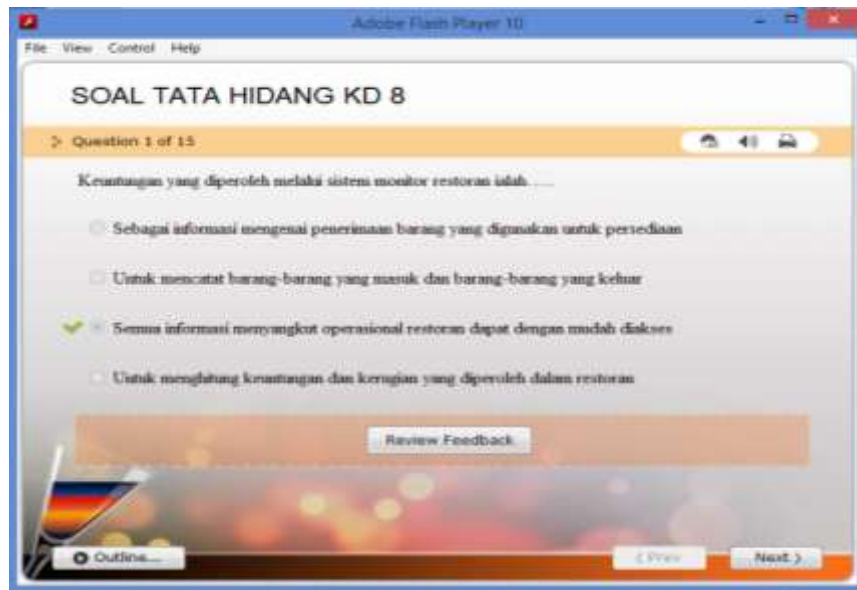
Gambar 8. Tampilan awal untuk untuk masuk sebagai peserta ujian

Mahasiswa yang akan mengikuti ujian Tata Hidang harus mengisi nama dan NIM untuk melanjutkan ke langkah berikutnya. Kompetensi dasar yang mau dikerjakan oleh peserta ujian dapat dipilih. Hasil ujian langsung dapat dilihat oleh peserta ujian setelah jawaban dikumpulkan (*Submit*) seperti tertera pada gambar berikut ini.



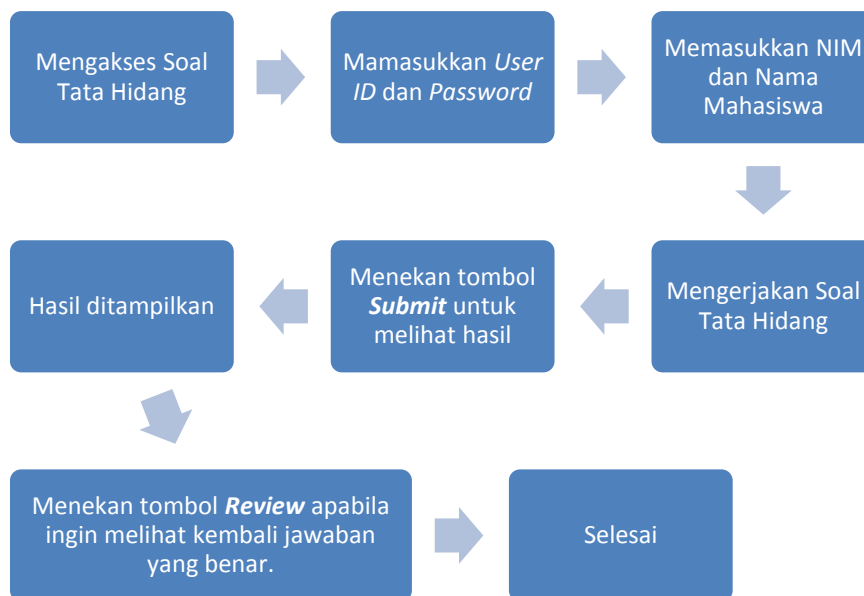
Gambar 9. Tampilan Hasil Ujian

Skor potong (*cut score*) yang merupakan batas lulus atau tidak lulus telah ditetapkan sebelumnya. Mahasiswa dapat melihat keputusan hasil ujian dengan membandingkan skor yang diperoleh dari skor potong yang telah ditetapkan. Peserta dapat mengoreksi kembali jawaban yang telah dipilih dengan kunci jawaban seperti terlihat pada gambar berikut ini:



Gambar 10. Tampilan Review Kunci Jawaban

Pada tampilan layar review tampak: (1) tombol *Author info* untuk melihat informasi pembuat soal; (2) tombol *Audio* untuk menghidupkan/mematikan efek suara dan pengaturan volume; dan (3) tombol *Print* untuk mencetak soal. Alur penggunaan program ujian yang dioperasikan menggunakan komputer terlihat pada gambar berikut:



Gambar 11. Alur Pelaksanaan Ujian Menggunakan Perangkat Komputer

4. Draft Buku Evaluasi Pembelajaran

Draft Buku Evaluasi Pembelajaran dibuat dalam 5 Bab mengupas tentang: a. Konsep Pengukuran, Penilaian dan Evaluasi; b. Peranan Penilaian dalam Pembelajaran; c. Analisis Kompetensi dan Tujuan Pembelajaran; d. Pengembangan Tes Baku; e. Pengendalian Kualitas Alat Penilaian. Buku dengan jumlah halaman 100 lembar.

B. Pembahasan

1. Analisis Kebutuhan

Mata kuliah Tata Hidang terdiri dari 1 SKS teori dan 2 SKS praktik. Materi yang dapat disusun menjadi butir soal pilihan ganda adalah materi untuk pelajaran teori yang berisi banyak pengetahuan. Materi memiliki banyak poin-poin penting yang perlu diketahui atau dihafalkan sehingga memberi kemungkinan untuk disusun pertanyaan-pertanyaan yang memilikibanyak alternatif jawaban pendek. Butir soal *multiple choice* (MC) hanya mengukur aspek pengetahuan (*knowledge*) saja. Soal MC tidak menuntut jawaban yang memberisolusi atas sebuah permasalahan, atau jawaban yang diuraikan.. Dengan karakteristik seperti itu, maka dalam analisis materi yang dikembangkan menjadi butir soal mata kuliah Tata Hidang adalah materi yang mewakili kompetensi dasar: pengetahuan menu, perlengkapan Tata Hidang, persiapan area restoran, etika makan, sistem pelayanan dan transaksi keuangan. Materi ini memiliki banyak point-point yang perlu dipahami oleh mahasiswa.

Butir soal yang telah berhasil disusun menjadi soal pilihan ganda sebanyak 193. Level kognisi yang ditulis meliputi level ingatan, pemahaman, dan aplikasi. Butir soal yang diujikan hanya 100, sehingga sisanya menjadi butir soal cadangan yang bisa diujikan secara paralel dengan dua paket soal yang berbeda. Jika semua butir soal dimasukkan ke dalam bank soal ujian di program komputer, maka bentuk penyelenggaraan ujian dapat lebih bervariasi misalnya menggunakan adaptive tes yaitu denganpenjenjangan tingkat kesulitan, dan sebagainya. Pada *adaptive test*, peserta diberi soal setelah dapat melewati butir soal yang mudah. Jumlah butir yang banyak dapat dikeluarkan secara acak sehingga tidak akan saling terjadi saling contoh antar peserta tes.

2. Telaah Soal

Soal yang akan didokumentasi di dalam bank soal harus telah memiliki kualitas butir yang baik. Untuk menjamin supaya soal yang masuk di bank soal telah memenuhi kriteria kualitas yang baik, maka soal harus melewati beberapa tahap penelaahan (*review*). Telaah soal telah dilakukan secara teoritik dan empiris. Telaah teoritik dilakukan untuk memvalidasi isi dan kualitas butir soal dari unsur materi, konstruksi, dan bahasa. Telaah empiris dilakukan untuk menguji kualitas butir soal dari unsur tingkat kesulitan, daya pembeda, validitas konstruk dan reliabilitas soal.

Hasil telaah soal teoritik menggunakan teknik Delphi memperoleh banyak masukan untuk memperbaiki susunan butir soal. Hasil telaah inipun telah ditindaklanjuti dengan menyusun kembali butir-butir soal yang perlu diperbaiki. Hasil telaah empiris masih menemukan butir-butir yang kurang bagus dari sisi tingkat kesulitan, daya pembeda, dan validitas butirnya. Pada penelitian tahun pertama, hasil analisis empiris belum ditindaklanjuti untuk memperbaiki konstruksi soal. Perbaikan soal akan dilakukan pada awal tahun penelitian kedua dan diuji coba menggunakan komputer pada saat ujian akhir semester genap 2014/2015 yaitu bulan Januari 2015. Revisi terakhir dilakukan pada semester genap, yaitu bulan Juli 2015.

3. Rancangan Pengembangan Program

Ada beberapa macam *software* yang dapat membantu penayangan soal ujian menggunakan komputer. Dalam penelitian ini, program yang digunakan untuk menayangkan soal ujian Tata Hidang dengan komputer adalah *software Wondershare*. Software ini memiliki beberapa kelebihan yaitu: (1) tampilan menarik dan bisa dimodifikasi; (2) pengoperasian sederhana sehingga mudah dilakukan oleh siapa saja; (3) biaya maintenance murah; (4) mudah diakses; (5) kapasitas ruang yang diperlukan untuk penyimpanan cukup kecil yaitu hanya memerlukan 1 KB .

BAB VI

KESIMPULAN DAN SARAN

A. Kesimpulan

1. Ruang lingkup materi Tata Hidang yang layak dikembangkan menjadi kisi-kisi soal tes pilihan ganda adalah: (1) *food and beverage service*; (2) pengetahuan menu Tata Hidang; (3) perlengkapan Tata Hidang; (4) persiapan area restoran; (5) etika dalam jamuan makan; (6) sistem pelayanan makanan; dan (7) transaksi keuangan; dan (8) dan ruang lingkup Tata Hidang.
2. Telaah kualitas butir soal ujian Tata Hidang secara kualitatif telah ditindaklanjuti untuk memperbaiki butir soal yang masih kurang baik yaitu: (1) belum mempunyai satu jawaban yang paling benar; (2) pilihan jawaban belum homogen dan logis ditinjau dari sisi materi; (3) rumusan soal dan rumusan jawaban tidak hanya merupakan pernyataan yang diperlukan saja; (4) soal menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia. Telaah kualitas butir soal tes secara kuantitatif telah membuktikan 77% soal berada pada tingkat kesulitan sedang, 48% butir soal memiliki daya pembeda sangat baik dan 30% soal memiliki daya pembeda yang baik. Nilai reliabilitas soal sebesar 0,945 telah memenuhi kriteria soal yang baik dan validitas butir soal on line adalah 0,167.
3. Program komputer yang digunakan untuk ujian Tata Hidang secara *on line* atau *computerized test* adalah *software Wondershare*. Pada saat merancang program telah disusun langkah-langkah pembuatan untuk programmer dan langkah-langkah pengoperasian pengguna. Program memiliki pengaman sehingga hanya pengguna yang diberi *password* saja yang dapat mengakses soal. Setelah mengerjakan soal, pengguna dapat langsung melihat skor tes dan kunci jawaban jika diperlukan. Program diterbitkan dalam berbagai bentuk (*Adobe flash, Exe, Microsoft Word, Mikrosoft Excel*) sehingga mudah diinstall di komputer.

B. Saran

1. Kepada tim pengajar mata kuliah Tata Hidang disarankan untuk menganalisis kembali kompetensi dasar lain yang dapat ditambahkan ke dalam bank soal pilihan ganda. Kepada Jurusan Pendidikan Teknik Boga dan Busana, disarankan untuk memperbanyak matakuliah yang penyelenggaraan ujian semesternya menggunakan komputer.
2. Kepada pengguna soal ujian Tata Hidang disarankan agar mempelajari materi secara lebih komprehensif supaya tidak banyak melakukan guessing (coba-coba) pada saat menjawab soal-soal ujian dan tidak terjebak untuk memilih jawaban pengecoh.
3. Kepada ahli perancang program, agar terus mengembangkan program-program yang sama dengan tampilan yang lebih bagus, hasil yang akurat, cepat dan mudah digunakan. Kegiatan penelitian berikutnya dapat dilanjutkan dengan mengidentifikasi kompetensi dasar keterampilan Tata Hidang yang dapat disimulasikan dengan animasi dan dikembangkan menjadi tes kinerja berbantuan komputer.

DAFTAR PUSTAKA

- Burden, P. R., & Byrd, D. M. (1998). *Methods for effective teaching*. Boston: Allyn and Bacon.
- Cisco Public Information. (2012). *A cisco networking academy point of view advancing assessment with technology*, diperoleh dari www.cisco.com/go/trademarks tanggal 18 Mei 2012
- Clauser, B. E., Harik, P. & Clyman, S. G., (2000). The generalizability of scores for a performance assessment scored with a computer automated scoring system. *Journal of Educational Measurement*. Fall 2000, Vol. 37, No. 3, pp. 245-261
- Conner, C. (1991). *Assessment and testing the primary school*. Philadelphia: The Falmer Press
- Dick, W., & Carey, L. (1996). *The Systematic Design of Instruction* (4th Ed.). New York: Haper Collins College Publishers.
- Griffin, P., & Peter, N. (1991). *Educational assessment and reporting*. Sidney: Harcourt Brace Javanovich Publisher (Hambleton, 1991: 13).
- Huitt, W. (2001). Assessment, measurement and evaluation: Undergraduate version. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved (17 Nopember 2007), from <http://chiron.valdosta.edu/whuitt/edpsyc/edpmsv1.html>
- Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment, A manageable and cooperative process*. Boston : Allyn and Bacon
- Kane, M. T. (2006). Validation. In Brennan, R. B. (Ed). *Educational measurement* (4th ed) pp 17 – 64. American Council on Education Praeger.
- Lunz, M. E., & Bergstrom, B. A. (1994). An empirical study of computerized adaptive test administration conditions. *Journal of Educational Measurement*. Fall 1994, Vol. 31, No. 3, pp. 251-263.
- McIntire, S. A., & Miller, L. A. (2000). *Foundation of psychological testing*. Boston: McGraw-Hill.
- Mehrens, W. A., & Lehman, I. (1973). *Measurement and evaluation in education and psychology*. New York: Holt, Rinehart and Winston, Inc.
- Messick, S. (1989). Validity. In Robert L. Linn (Ed). *Educational measurement* (3rded) pp. 13-103. New York: American Council on Education and Macmillan
- Neuman, W. L. (2003). *Social research methods, qualitative and quantitative approaches* (5th). Boston: Pearson Education Inc
- Nunnaly, J. C. (1978). *Psychometric theory*, (2nd ed.). New Delhi: McGraw-Hill Publishing Company Limited

- Oshima, T. C. (1994). The effect of speediness on parameter estimation in item response theory. *Journal of Educational Measurement*. Fall 1994, Vol. 31, No. 3, pp. 200-219.
- Pitkin, A. K. and Vispoel, W. P. (2001). Differences between self-adapted and computerized adaptive test: A meta-analysis. *Journal of Educational Measurement*. Fall 2001, Vol. 38, No. 3, pp. 235-241
- Reynolds, C. R., Livingston, R. B. & Willson, V. (2010). *Measurement and assessment in education 2nd edition*: New Jersey: Pearson Prentice Hall
- Rodiguez, M. C. (2003). Construct equivalence of multiple-choice and constructed-response items: A random effects synthesis of correlation. *Journal of Educational Measurement*, Summer 2003, Vol. 40, No. 2, pp. 163-184
- Sumadi Suryabrata. (1981). *Pengukuran dan penilaian pendidikan*, (Kumpulan makalah penataran bimbingan dan konseling untuk tenaga pengajar di PT se Indonesia, edt). Jakarta: Dirjen Pendidikan (Suryabrata, 1981).
- Yahya Umar. (1996). *Bahan penataran pengujian pendidikan*. Jakarta: Puslitbangsisjian-Depdikbud

LAMPIRAN

Lampiran 1. Instrumen Penelitian

Contoh Kisi-Kisi dan Soal Tes Kognitif Tata Hidang

Kisi-kisi Soal (KD V) Persiapan Area Restoran

No.	Indikator	Aspek Kognitif	Soal	Kunci Jawaban
1.	Membersihkan restoran	C2	Pekerjaan meluruskan atau menempatkan meja dan kursi pada tempat yang serasi dilakukan oleh seseorang dengan pengawasan dari a. <i>Waitress</i> b. <i>Housekeeping Department</i> c. <i>Food & Beverage Departement</i> d. <i>Front Office Department</i>	B
2	<i>Polishing</i> Peralatan makan dan minum	C1	Tujuan <i>polishing</i> adalah ... a. Peralatan dapat tahan lama b. Peralatan bebas dari bakteri c. Peralatan bebas dari debu d. Peralatan bebas dari kotoran	B
3	<i>Polishing</i> Peralatan makan dan minum	C1	Kegiatan yang dilakukan untuk menghilangkan kotoran atau debu yang ada di peralatan-peralatan makan dan minum disebut a. <i>Table Setting</i> b. <i>Polishing</i> c. <i>Side station</i> d. <i>Waiters point</i>	B
4	<i>Polishing</i> Peralatan makan dan minum	C2	Tujuan utama penggunaan uap panas pada saat <i>polishing</i> adalah a. Menghilangkan bau amis b. Mematikan bakteri c. Membersihkan lemak d. Menghilangkan kotoran	A
5	Mengatur meja (table setting)	C1	Rangkaian kegiatan menutup meja menggunakan peralatan seperti <i>chinaware</i> , <i>silverware</i> dan linen disebut a. <i>Table Setting</i> b. <i>Polishing</i> c. <i>Side station</i> d. <i>Waiters point</i>	A

No.	Indikator	Aspek Kognitif	Soal	Kunci Jawaban
6	Mempersiapkan meja persediaan (<i>side station</i>)	C3	Jenis alat yang termasuk <i>assorted cutlery</i> adalah a. <i>Ashtray and bread and butter plate</i> b. <i>Champagne glasses and bread butter plate</i> c. <i>Champagne glasses and service spoon</i> d. <i>Service spoon and dessert knife</i>	D
7	Mempersiapkan meja persediaan (<i>side station</i>)	C3	Di bawah ini yang termasuk perlengkapan linen adalah a. <i>Flower vase</i> b. <i>Table number</i> c. <i>Moulton</i> d. <i>Ashtray</i>	C
8	Mengatur Meja (<i>Table setting</i>)	C3	Berikut ini adalah urutan <i>table set-up</i> secara acak 1. <i>Menyiapkan silverware</i> 2. <i>Memasang taplak meja</i> 3. <i>Memasang moulton</i> 4. <i>Meletakkan show plate</i> 5. <i>Meletakkan table accessories</i> 6. <i>Meletakkan cutlery dan silverware sesuai aturan</i> Urutan <i>table set-up</i> yang benar adalah a. 1 – 2 – 3 – 4 – 5 – 6 b. 1 – 2 – 3 – 5 – 4 – 6 c. 1 – 3 – 2 – 5 – 4 – 6 d. 1 – 3 – 2 – 4 – 5 – 6	C
9	Mengatur Meja (<i>Table setting</i>)	C3	Dalam <i>table setting Bread and butter plate</i> diletakkan pada a. Sisi kanan <i>dinner knife</i> b. Di atas <i>dinner knife</i> c. Sisi kiri <i>dinner fork</i> d. Di atas <i>show plate</i>	C
10	Jenis Table Setting	C2	<i>Table setting</i> yang terdiri dari <i>B&B Plate, B&B Knife, Dinner fork, show plate, napkin, dinner knife, water goblet</i> termasuk jenis a. <i>American table setting</i> b. <i>Standart table setting</i> c. <i>Elaborate table setting</i> d. <i>Basic table setting</i>	D

No.	Indikator	Aspek Kognitif	Soal	Kunci Jawaban
11	Tahap-tahap Pelayanan Makanan	C1	Suatu prosedur pelaksanaan yang bertujuan untuk memudahkan pekerjaan karyawan dan menciptakan suasana yang membuat para tamu menjadi nyaman disebut ... a. GMP b. SNI c. ISO d. SOP	D
12	Tahap-tahap Pelayanan Makanan	C1	Seseorang yang bertugas untuk menyambut dan menerima tamu pada saat memasuki area restoran adalah ... a. <i>Captain</i> b. <i>Sommelier</i> c. <i>Waitres</i> d. <i>Hostess</i>	D
13	Tahap-tahap Pelayanan Makanan	C1	Presenting the wine list dilakukan oleh a. <i>Captain</i> b. <i>Sommelier</i> c. <i>Waitres</i> d. <i>Hostess</i>	B
14	Tahap-tahap Pelayanan Makanan	C3	Peralatan-peralatan yang di <i>clear up</i> hidangan <i>dessert</i> disajikan adalah a. <i>Flower vase, B&B plate and wine glass</i> b. <i>Ashtray, B&B plate and goblet</i> c. <i>Ashtray, B&B plate and wine glass</i> d. <i>Ashtray, B&B plate and tea cup</i>	C
15	Tahap-tahap Pelayanan Makanan	C2	Proses table setting yang dilaksanakan di sebelah kiri tamu adalah a. Penyajian hidangan dengan <i>platter</i> b. Penyajian minuman c. Proses clear-up d. Penyajian soup	A

Lampiran 2. Hasil Analisis Butir Soal

MicroCAT (tm) Testing System
 Copyright (c) 1982, 1984, 1986, 1988 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 1

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
1	0-1	0.367	0.508	0.397	A	0.167	0.178	0.119	
					B	0.233	-0.298	-0.216	
					C	0.233	-0.472	-0.341	
					D	0.367	0.508	0.397	*
					Other	0.000	-9.000	-9.000	
2	0-2	0.667	0.751	0.579	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.667	0.751	0.579	*
					D	0.333	-0.751	-0.579	
					Other	0.000	-9.000	-9.000	
3	0-3	0.633	0.342	0.267	A	0.000	-9.000	-9.000	
					B	0.633	0.342	0.267	*
					C	0.100	-0.442	-0.258	
					D	0.267	-0.156	-0.116	
					Other	0.000	-9.000	-9.000	
4	0-4	0.400	0.542	0.427	A	0.000	-9.000	-9.000	
					B	0.400	0.542	0.427	*
					C	0.167	-0.239	-0.160	
					D	0.433	-0.380	-0.302	
					Other	0.000	-9.000	-9.000	
5	0-5	0.600	0.476	0.376	A	0.000	-9.000	-9.000	
					B	0.600	0.476	0.376	*
					C	0.167	-0.173	-0.116	
					D	0.233	-0.460	-0.333	
					Other	0.000	-9.000	-9.000	
6	0-6	0.467	0.382	0.305	A	0.200	-0.090	-0.063	
					B	0.133	-0.259	-0.164	
					C	0.200	-0.253	-0.177	
					D	0.467	0.382	0.305	*
					Other	0.000	-9.000	-9.000	
7	0-7	0.400	0.480	0.379	A	0.233	-0.095	-0.069	
					B	0.233	-0.466	-0.337	
					C	0.133	-0.064	-0.041	
					D	0.400	0.480	0.379	*
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 2

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics					Key
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.		
8	0-8	0.433	0.517	0.411	A	0.433	0.517	0.411	*	
					B	0.367	-0.357	-0.279		
					C	0.133	-0.293	-0.186		
					D	0.067	-0.046	-0.024		
					Other	0.000	-9.000	-9.000		
9	0-9	0.367	0.275	0.215	A	0.000	-9.000	-9.000	*	
					B	0.367	0.275	0.215		
					C	0.300	-0.185	-0.140		
					D	0.333	-0.107	-0.083		
					Other	0.000	-9.000	-9.000		
10	0-10	0.567	0.784	0.623	A	0.000	-9.000	-9.000	*	
					B	0.567	0.784	0.623		
					C	0.333	-0.656	-0.506		
					D	0.100	-0.400	-0.234		
					Other	0.000	-9.000	-9.000		
11	0-11	0.533	0.298	0.238	A	0.533	0.298	0.238	*	
					B	0.167	-0.071	-0.047		
					C	0.200	-0.142	-0.100		
					D	0.100	-0.348	-0.204		
					Other	0.000	-9.000	-9.000		
12	0-12	0.433	0.392	0.311	A	0.433	0.392	0.311	*	
					B	0.233	-0.173	-0.125		
					C	0.233	-0.292	-0.212		
					D	0.100	-0.067	-0.039		
					Other	0.000	-9.000	-9.000		
13	0-13	0.633	0.498	0.389	A	0.633	0.498	0.389	*	
					B	0.067	0.223	0.115		
					C	0.133	-0.667	-0.422		
					D	0.167	-0.290	-0.195		
					Other	0.000	-9.000	-9.000		
14	0-14	0.567	0.659	0.523	A	0.167	-0.488	-0.327		
					B	0.100	-0.577	-0.338		
					C	0.167	-0.144	-0.097		
					D	0.567	0.659	0.523	*	
					Other	0.000	-9.000	-9.000		

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 3

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
15	0-15	0.500	0.298	0.238	A	0.000	-9.000	-9.000	
					B	0.500	0.298	0.238	*
					C	0.233	-0.119	-0.086	
					D	0.267	-0.251	-0.186	
					Other	0.000	-9.000	-9.000	
16	0-16	0.533	0.717	0.571	A	0.200	-0.469	-0.328	
					B	0.067	-0.386	-0.200	
					C	0.200	-0.371	-0.260	
					D	0.533	0.717	0.571	*
					Other	0.000	-9.000	-9.000	
17	0-17	0.567	0.659	0.523	A	0.567	0.659	0.523	*
					B	0.133	-0.166	-0.105	
					C	0.267	-0.568	-0.422	
					D	0.033	-0.496	-0.205	
					Other	0.000	-9.000	-9.000	
18	0-18	0.500	0.472	0.377	A	0.500	0.472	0.377	*
					B	0.233	-0.573	-0.415	
					C	0.167	-0.210	-0.141	
					D	0.100	0.225	0.132	
					Other	0.000	-9.000	-9.000	
19	0-19	0.400	0.471	0.371	A	0.000	-9.000	-9.000	
					B	0.400	0.471	0.371	*
					C	0.367	-0.095	-0.074	
					D	0.233	-0.478	-0.346	
					Other	0.000	-9.000	-9.000	
20	0-20	0.467	0.516	0.411	A	0.233	-0.388	-0.281	
					B	0.167	-0.107	-0.072	
					C	0.133	-0.276	-0.175	
					D	0.467	0.516	0.411	*
					Other	0.000	-9.000	-9.000	
21	0-21	0.467	0.037	0.030	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.467	0.037	0.030	*
					D	0.533	-0.037	-0.030	
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 4

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
22	0-22	0.500	0.481	0.384	A	0.000	-9.000	-9.000	
					B	0.500	0.481	0.384	*
					C	0.200	-0.489	-0.342	
					D	0.300	-0.159	-0.121	
					Other	0.000	-9.000	-9.000	
23	0-23	0.567	0.566	0.449	A	0.567	0.566	0.449	*
					B	0.133	-0.497	-0.315	
					C	0.100	-0.119	-0.069	
					D	0.200	-0.338	-0.237	
					Other	0.000	-9.000	-9.000	
24	0-24	0.733	0.529	0.393	A	0.000	-9.000	-9.000	
					B	0.733	0.529	0.393	*
					C	0.133	-0.599	-0.379	
					D	0.133	-0.208	-0.132	
					Other	0.000	-9.000	-9.000	
25	0-25	0.733	0.134	0.100	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.733	0.134	0.100	*
					D	0.267	-0.134	-0.100	
					Other	0.000	-9.000	-9.000	
26	0-26	0.400	0.504	0.397	A	0.400	0.504	0.397	*
					B	0.333	-0.464	-0.358	
					C	0.133	0.097	0.062	
					D	0.133	-0.217	-0.137	
					Other	0.000	-9.000	-9.000	
27	0-27	0.567	0.389	0.309	A	0.567	0.389	0.309	*
					B	0.100	-0.160	-0.094	
					C	0.200	-0.149	-0.104	
					D	0.133	-0.387	-0.245	
					Other	0.000	-9.000	-9.000	
28	0-28	0.467	0.626	0.499	A	0.467	0.626	0.499	*
					B	0.267	-0.501	-0.372	
					C	0.267	-0.256	-0.191	
					D	0.000	-9.000	-9.000	
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 5

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
29	0-29	0.533	0.634	0.505	A	0.000	-9.000	-9.000	
					B	0.533	0.634	0.505	*
					C	0.233	-0.537	-0.389	
					D	0.233	-0.286	-0.207	
					Other	0.000	-9.000	-9.000	
30	0-30	0.700	0.616	0.468	A	0.700	0.616	0.468	*
					B	0.100	-0.400	-0.234	
					C	0.067	-0.442	-0.229	
					D	0.133	-0.404	-0.256	
					Other	0.000	-9.000	-9.000	
31	0-31	0.800	0.717	0.502	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.800	0.717	0.502	*
					D	0.200	-0.717	-0.502	
					Other	0.000	-9.000	-9.000	
32	0-32	0.567	0.496	0.394	A	0.000	-9.000	-9.000	
					B	0.567	0.496	0.394	*
					C	0.233	-0.149	-0.108	
					D	0.200	-0.534	-0.374	
					Other	0.000	-9.000	-9.000	
33	0-33	0.467	0.474	0.378	A	0.467	0.474	0.378	*
					B	0.233	-0.788	-0.570	
					C	0.133	0.267	0.169	
					D	0.167	-0.020	-0.013	
					Other	0.000	-9.000	-9.000	
34	0-34	0.433	0.466	0.370	A	0.433	0.466	0.370	*
					B	0.200	-0.253	-0.177	
					C	0.200	-0.090	-0.063	
					D	0.167	-0.349	-0.234	
					Other	0.000	-9.000	-9.000	
35	0-35	0.533	-0.318	-0.253	A	0.200	0.380	0.266	?
					B	0.133	-0.149	-0.094	
					C	0.133	0.242	0.153	
					D	0.533	-0.318	-0.253	*
					Other	0.000	-9.000	-9.000	
CHECK THE KEY									
D was specified, A works better									

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 6

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics							
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key			
36	0-36	0.567	0.505	0.401	A	0.000	-9.000	-9.000				
					B	0.000	-9.000	-9.000				
					C	0.567	0.505	0.401	*			
					D	0.433	-0.505	-0.401				
					Other	0.000	-9.000	-9.000				
37	0-37	0.433	0.508	0.403	A	0.000	-9.000	-9.000				
					B	0.000	-9.000	-9.000				
					C	0.433	0.508	0.403	*			
					D	0.567	-0.508	-0.403				
					Other	0.000	-9.000	-9.000				
38	0-38	0.367	0.445	0.347	A	0.200	-0.182	-0.127				
					B	0.167	0.105	0.070				
					C	0.267	-0.434	-0.323				
					D	0.367	0.445	0.347	*			
					Other	0.000	-9.000	-9.000				
39	0-39	0.433	0.740	0.588	A	0.000	-9.000	-9.000				
					B	0.433	0.740	0.588	*			
					C	0.300	-0.522	-0.396				
					D	0.267	-0.334	-0.248				
					Other	0.000	-9.000	-9.000				
40	0-40	0.633	0.163	0.127	A	0.000	-9.000	-9.000				
					B	0.633	0.163	0.127	*			
					CHECK THE KEY			C	0.133	0.293	0.185	?
					B was specified, C works better			D	0.233	-0.406	-0.294	
					Other	0.000	-9.000	-9.000				
41	0-41	0.667	0.500	0.385	A	0.000	-9.000	-9.000				
					B	0.667	0.500	0.385	*			
					C	0.200	-0.469	-0.328				
					D	0.133	-0.234	-0.148				
					Other	0.000	-9.000	-9.000				
42	0-42	0.633	0.415	0.324	A	0.000	-9.000	-9.000				
					B	0.000	-9.000	-9.000				
					C	0.633	0.415	0.324	*			
					D	0.367	-0.415	-0.324				
					Other	0.000	-9.000	-9.000				

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 7

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
43	0-43	0.500	0.477	0.380	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.500	0.477	0.380	*
					D	0.500	-0.477	-0.380	
					Other	0.000	-9.000	-9.000	
44	0-44	0.533	0.229	0.183	A	0.533	0.229	0.183	*
					B	0.167	-0.349	-0.234	
					C	0.233	-0.071	-0.052	
					D	0.067	0.138	0.071	
					Other	0.000	-9.000	-9.000	
45	0-45	0.633	0.527	0.412	A	0.133	-0.030	-0.019	
					B	0.100	-0.317	-0.185	
					C	0.133	-0.633	-0.401	
					D	0.633	0.527	0.412	*
					Other	0.000	-9.000	-9.000	
46	0-46	0.633	0.629	0.491	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.633	0.629	0.491	*
					D	0.367	-0.629	-0.491	
					Other	0.000	-9.000	-9.000	
47	0-47	0.667	0.439	0.339	A	0.667	0.439	0.339	*
					B	0.133	-0.089	-0.057	
					C	0.167	-0.517	-0.347	
					D	0.033	-0.151	-0.062	
					Other	0.000	-9.000	-9.000	
48	0-48	0.167	0.603	0.404	A	0.167	0.603	0.404	*
					B	0.300	-0.448	-0.340	
					C	0.267	0.050	0.037	
					D	0.267	-0.034	-0.025	
					Other	0.000	-9.000	-9.000	
49	0-49	0.667	0.656	0.506	A	0.167	-0.554	-0.371	
					B	0.167	-0.400	-0.268	
					C	0.000	-9.000	-9.000	
					D	0.667	0.656	0.506	*
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 8

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
50	0-50	0.500	0.133	0.106	A	0.000	-9.000	-9.000	
					B	0.500	0.133	0.106	*
					C	0.300	-0.206	-0.156	
					D	0.200	0.067	0.047	
					Other	0.000	-9.000	-9.000	
51	0-51	0.267	-0.095	-0.071	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.267	-0.095	-0.071	*
					D	0.733	0.095	0.071	?
					Other	0.000	-9.000	-9.000	
CHECK THE KEY C was specified, D works better									
52	0-52	0.733	0.434	0.323	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.733	0.434	0.323	*
					D	0.267	-0.434	-0.323	
					Other	0.000	-9.000	-9.000	
53	0-53	0.267	0.078	0.058	A	0.267	0.078	0.058	*
					B	0.233	0.335	0.242	?
					C	0.333	-0.158	-0.122	
					D	0.167	-0.283	-0.190	
					Other	0.000	-9.000	-9.000	
CHECK THE KEY A was specified, B works better									
54	0-54	0.667	0.404	0.312	A	0.667	0.404	0.312	*
					B	0.133	-0.183	-0.116	
					C	0.100	-0.369	-0.216	
					D	0.100	-0.244	-0.143	
					Other	0.000	-9.000	-9.000	
55	0-55	0.467	0.718	0.572	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.467	0.718	0.572	*
					D	0.533	-0.718	-0.572	
					Other	0.000	-9.000	-9.000	
56	0-56	0.433	0.978	0.776	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.433	0.978	0.776	*
					D	0.567	-0.978	-0.776	
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 9

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
57	0-57	0.133	-0.089	-0.057	A	0.133	-0.089	-0.057	*
					B	0.267	-0.245	-0.182	
					C	0.433	0.127	0.101	
					D	0.167	0.200	0.134	?
					Other	0.000	-9.000	-9.000	
CHECK THE KEY A was specified, D works better									
58	0-58	0.300	0.572	0.434	A	0.000	-9.000	-9.000	
					B	0.300	0.572	0.434	*
					C	0.267	-0.468	-0.348	
					D	0.433	-0.115	-0.091	
					Other	0.000	-9.000	-9.000	
59	0-59	0.167	0.866	0.581	A	0.367	0.003	0.002	
					B	0.233	-0.095	-0.069	
					C	0.233	-0.615	-0.445	
					D	0.167	0.866	0.581	*
					Other	0.000	-9.000	-9.000	
60	0-60	0.633	0.439	0.343	A	0.100	-0.088	-0.051	
					B	0.200	-0.293	-0.205	
					C	0.067	-0.527	-0.273	
					D	0.633	0.439	0.343	*
					Other	0.000	-9.000	-9.000	
61	0-61	0.267	1.000	0.773	A	0.000	-9.000	-9.000	
					B	0.267	1.000	0.773	*
					C	0.500	-0.527	-0.421	
					D	0.233	-0.430	-0.311	
					Other	0.000	-9.000	-9.000	
62	0-62	0.767	0.669	0.484	A	0.000	-9.000	-9.000	
					B	0.767	0.669	0.484	*
					C	0.100	-0.734	-0.429	
					D	0.133	-0.353	-0.223	
					Other	0.000	-9.000	-9.000	
63	0-63	0.500	0.298	0.238	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.500	0.298	0.238	*
					D	0.500	-0.298	-0.238	
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 10

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
64	0-64	0.033	-0.151	-0.062	A	0.000	-9.000	-9.000					
					B	0.000	-9.000	-9.000					
					CHECK THE KEY				C	0.033	-0.151	-0.062	*
					C was specified, D works better				D	0.967	0.151	0.062	?
					Other	0.000	-9.000	-9.000					
65	0-65	0.567	0.547	0.434	A	0.000	-9.000	-9.000					
					B	0.000	-9.000	-9.000					
					C	0.567	0.547	0.434	*				
					D	0.433	-0.547	-0.434					
					Other	0.000	-9.000	-9.000					
66	0-66	0.633	0.503	0.393	A	0.000	-9.000	-9.000					
					B	0.633	0.503	0.393	*				
					C	0.167	-0.334	-0.224					
					D	0.200	-0.378	-0.264					
					Other	0.000	-9.000	-9.000					
67	0-67	0.167	-0.012	-0.008	A	0.233	-0.292	-0.212					
					B	0.333	0.416	0.321	?				
					CHECK THE KEY				C	0.267	-0.178	-0.133	
					D was specified, B works better				D	0.167	-0.012	-0.008	*
					Other	0.000	-9.000	-9.000					
68	0-68	0.000	-9.000	-9.000	A	0.000	-9.000	-9.000					
					B	0.000	-9.000	-9.000	*				
					CHECK THE KEY				C	0.433	-0.157	-0.124	
					B was specified, D works better				D	0.567	0.157	0.124	?
					Other	0.000	-9.000	-9.000					
69	0-69	0.200	0.988	0.691	A	0.000	-9.000	-9.000					
					B	0.200	0.988	0.691	*				
					C	0.400	-0.315	-0.249					
					D	0.400	-0.400	-0.316					
					Other	0.000	-9.000	-9.000					
70	0-70	0.400	0.390	0.308	A	0.400	0.390	0.308	*				
					B	0.233	-0.233	-0.168					
					C	0.167	-0.283	-0.190					
					D	0.200	-0.031	-0.022					
					Other	0.000	-9.000	-9.000					

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 11

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics								
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key				
71	0-71	0.267	0.606	0.450	A	0.000	-9.000	-9.000					
					B	0.000	-9.000	-9.000					
					C	0.267	0.606	0.450	*				
					D	0.733	-0.606	-0.450					
					Other	0.000	-9.000	-9.000					
72	0-72	0.233	0.819	0.593	A	0.233	0.819	0.593	*				
					B	0.333	-0.112	-0.087					
					C	0.200	-0.332	-0.232					
					D	0.233	-0.382	-0.276					
					Other	0.000	-9.000	-9.000					
73	0-73	0.200	0.988	0.691	A	0.433	-0.510	-0.405					
					B	0.100	-0.244	-0.143					
					C	0.267	-0.101	-0.075					
					D	0.200	0.988	0.691	*				
					Other	0.000	-9.000	-9.000					
74	0-74	0.333	0.818	0.631	A	0.000	-9.000	-9.000					
					B	0.333	0.818	0.631	*				
					C	0.267	-0.490	-0.364					
					D	0.400	-0.353	-0.278					
					Other	0.000	-9.000	-9.000					
75	0-75	0.033	-0.003	-0.001	A	0.033	-0.003	-0.001	*				
					B	0.300	0.094	0.071	?				
					CHECK THE KEY				C	0.400	-0.055	-0.043	
					A was specified, B works better				D	0.267	-0.034	-0.025	
					Other	0.000	-9.000	-9.000					
76	0-76	0.400	0.807	0.636	A	0.133	-0.361	-0.229					
					B	0.300	-0.464	-0.352					
					C	0.167	-0.290	-0.195					
					D	0.400	0.807	0.636	*				
					Other	0.000	-9.000	-9.000					
77	0-77	0.400	0.807	0.636	A	0.000	-9.000	-9.000					
					B	0.000	-9.000	-9.000					
					C	0.400	0.807	0.636	*				
					D	0.600	-0.807	-0.636					
					Other	0.000	-9.000	-9.000					

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 12

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
78	0-78	0.167	0.120	0.080	A	0.233	-0.352	-0.255	
					B	0.333	-0.012	-0.009	
					C	0.267	0.250	0.186	?
					D	0.167	0.120	0.080	*
					Other	0.000	-9.000	-9.000	
CHECK THE KEY D was specified, C works better									
79	0-79	0.200	0.988	0.691	A	0.267	-0.290	-0.215	
					B	0.333	-0.238	-0.184	
					C	0.200	-0.338	-0.237	
					D	0.200	0.988	0.691	*
					Other	0.000	-9.000	-9.000	
80	0-80	0.367	0.663	0.518	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.367	0.663	0.518	*
					D	0.633	-0.663	-0.518	
					Other	0.000	-9.000	-9.000	
81	0-81	0.400	0.641	0.505	A	0.200	-0.260	-0.182	
					B	0.167	-0.246	-0.165	
					C	0.233	-0.370	-0.268	
					D	0.400	0.641	0.505	*
					Other	0.000	-9.000	-9.000	
82	0-82	0.300	1.000	0.825	A	0.267	-0.535	-0.397	
					B	0.133	-0.378	-0.240	
					C	0.300	-0.348	-0.264	
					D	0.300	1.000	0.825	*
					Other	0.000	-9.000	-9.000	
83	0-83	0.500	0.385	0.307	A	0.500	0.385	0.307	*
					B	0.133	-0.353	-0.223	
					C	0.200	-0.123	-0.086	
					D	0.167	-0.173	-0.116	
					Other	0.000	-9.000	-9.000	
84	0-84	0.300	1.000	0.825	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.300	1.000	0.825	*
					D	0.700	-1.000	-0.825	
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 13

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
85	0-85	0.600	0.476	0.376	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.600	0.476	0.376	*
					D	0.400	-0.476	-0.376	
					Other	0.000	-9.000	-9.000	
86	0-86	0.033	0.489	0.202	A	0.000	-9.000	-9.000	
					B	0.033	0.489	0.202	*
					C	0.467	-0.013	-0.011	
					D	0.500	-0.078	-0.062	
					Other	0.000	-9.000	-9.000	
87	0-87	0.033	-0.619	-0.256	A	0.000	-9.000	-9.000	
					B	0.033	-0.619	-0.256	*
					C	0.600	0.571	0.450	?
					D	0.367	-0.464	-0.362	
					Other	0.000	-9.000	-9.000	
CHECK THE KEY B was specified, C works better									
88	0-88	0.333	0.506	0.391	A	0.000	-9.000	-9.000	
					B	0.333	0.506	0.391	*
					C	0.400	-0.453	-0.357	
					D	0.267	-0.028	-0.021	
					Other	0.000	-9.000	-9.000	
89	0-89	0.533	0.469	0.373	A	0.000	-9.000	-9.000	
					B	0.533	0.469	0.373	*
					C	0.200	-0.169	-0.118	
					D	0.267	-0.423	-0.315	
					Other	0.000	-9.000	-9.000	
90	0-90	0.400	0.688	0.543	A	0.400	0.688	0.543	*
					B	0.100	-0.140	-0.082	
					C	0.233	-0.358	-0.259	
					D	0.267	-0.401	-0.298	
					Other	0.000	-9.000	-9.000	
91	0-91	0.333	0.818	0.631	A	0.000	-9.000	-9.000	
					B	0.333	0.818	0.631	*
					C	0.400	-0.211	-0.166	
					D	0.267	-0.657	-0.488	
					Other	0.000	-9.000	-9.000	

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
92	0-92	0.433	0.406	0.322	A	0.000	-9.000	-9.000	
					B	0.433	0.406	0.322	*
					C	0.267	-0.112	-0.083	
					D	0.300	-0.353	-0.268	
					Other	0.000	-9.000	-9.000	
93	0-93	0.400	0.551	0.435	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.400	0.551	0.435	*
					D	0.600	-0.551	-0.435	
					Other	0.000	-9.000	-9.000	
94	0-94	0.433	0.736	0.584	A	0.167	-0.400	-0.268	
					B	0.200	-0.443	-0.310	
					C	0.200	-0.234	-0.164	
					D	0.433	0.736	0.584	*
					Other	0.000	-9.000	-9.000	
95	0-95	0.333	0.552	0.425	A	0.333	0.552	0.425	*
					B	0.233	0.030	0.022	
					C	0.167	-0.371	-0.249	
					D	0.267	-0.357	-0.265	
					Other	0.000	-9.000	-9.000	
96	0-96	0.467	0.783	0.624	A	0.000	-9.000	-9.000	
					B	0.467	0.783	0.624	*
					C	0.300	-0.495	-0.376	
					D	0.233	-0.454	-0.328	
					Other	0.000	-9.000	-9.000	
97	0-97	0.333	0.647	0.499	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.333	0.647	0.499	*
					D	0.667	-0.647	-0.499	
					Other	0.000	-9.000	-9.000	
98	0-98	0.333	0.567	0.437	A	0.300	-0.353	-0.268	
					B	0.200	-0.182	-0.127	
					C	0.167	-0.129	-0.087	
					D	0.333	0.567	0.437	*
					Other	0.000	-9.000	-9.000	

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 15

Seq. No.	Scale -Item	Item Statistics			Alternative Statistics				
		Prop. Correct	Biser.	Point Biser.	Alt.	Prop. Endorsing	Biser.	Point Biser.	Key
99	0-99	0.367	0.255	0.199	A	0.367	0.255	0.199	*
					B	0.167	-0.334	-0.224	
					C	0.233	-0.077	-0.056	
					D	0.233	0.036	0.026	
					Other	0.000	-9.000	-9.000	
100	0-100	0.533	0.570	0.454	A	0.000	-9.000	-9.000	
					B	0.000	-9.000	-9.000	
					C	0.533	0.570	0.454	*
					D	0.467	-0.570	-0.454	
					Other	0.000	-9.000	-9.000	

MicroCAT (tm) Testing System
Copyright (c) 1982, 1984, 1986, 1988 by Assessment Systems Corporation

Item and Test Analysis Program -- ITEMAN (tm) Version 3.00

Item analysis for data from file priha2.dat

Page 16

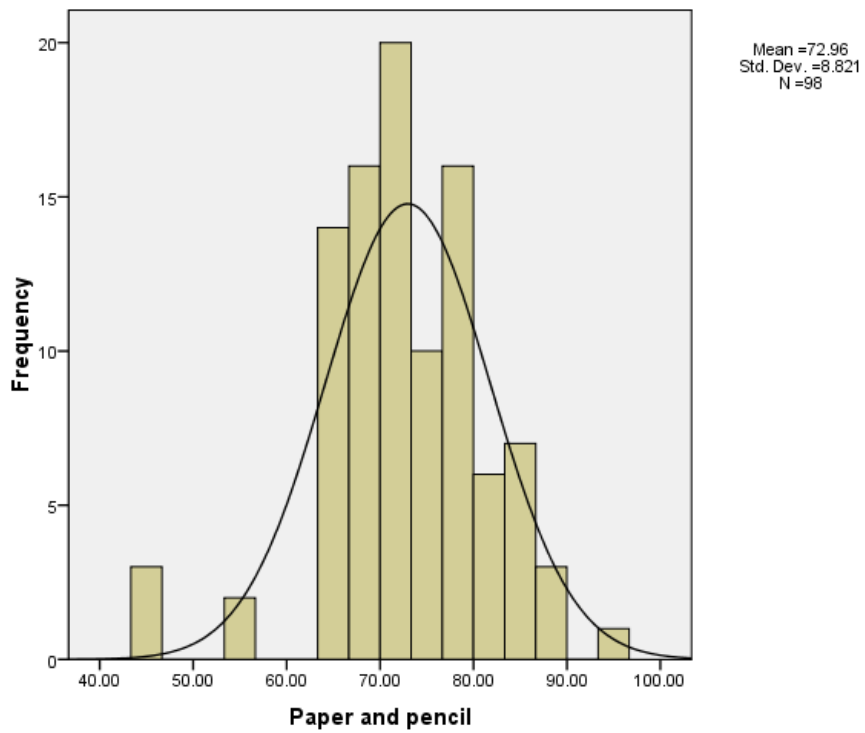
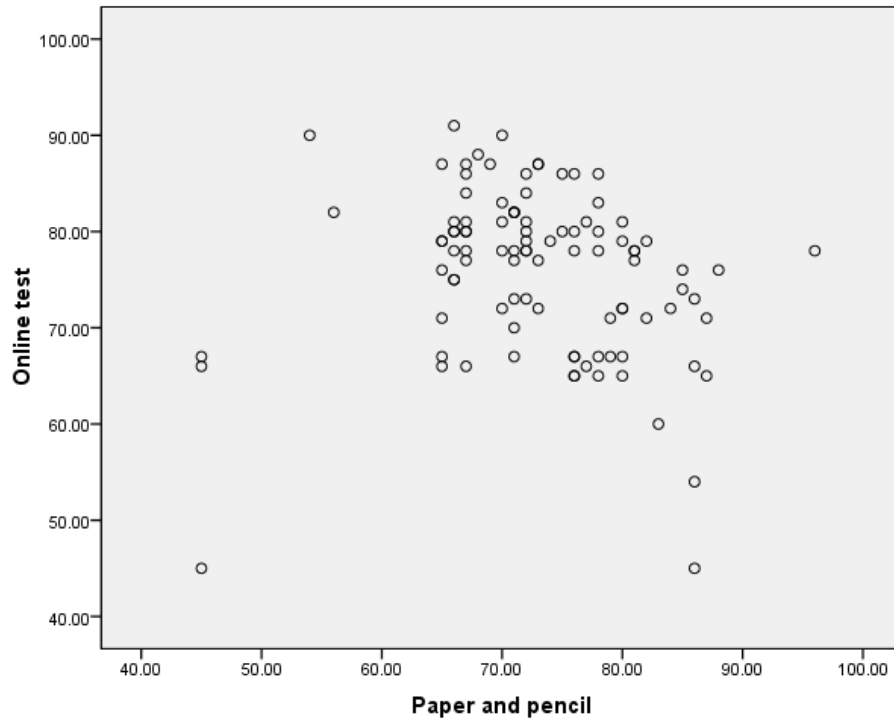
There were 30 examinees in the data file.

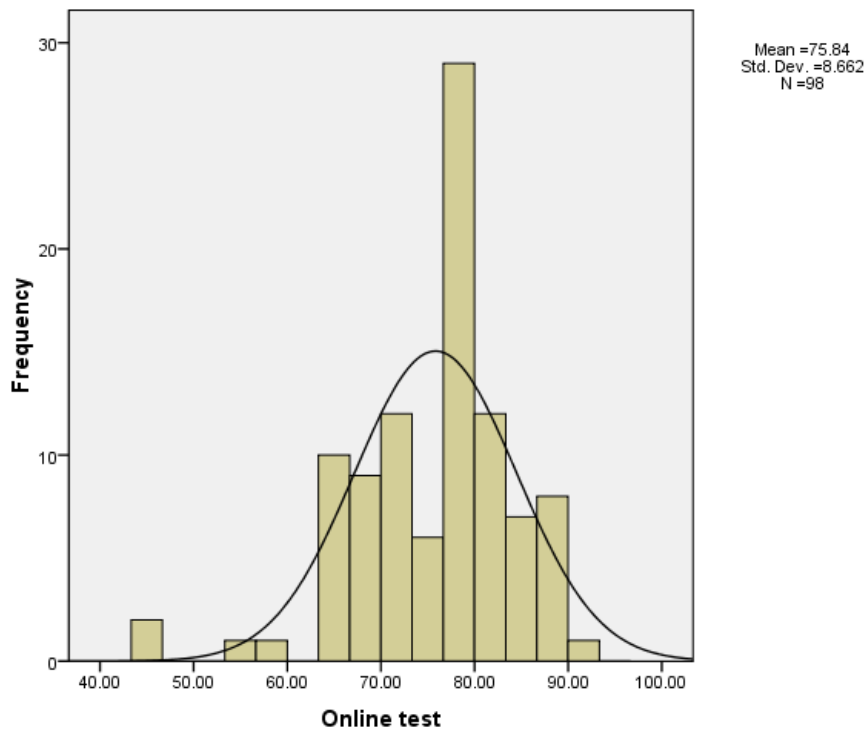
Scale Statistics

Scale:	0

N of Items	100
N of Examinees	30
Mean	44.133
Variance	332.249
Std. Dev.	18.228
Skew	0.295
Kurtosis	-1.316
Minimum	19.000
Maximum	77.000
Median	38.000
Alpha	0.945
SEM	4.287
Mean P	0.441
Mean Item-Tot.	0.388
Mean Biserial	0.499

Lampiran 2. Hasil Analisis Uji Validitas





Model Summary

Model	R	R Square	Adjusted R Square	Std. Error of the Estimate
1	.141 ^a	.020	.010	8.61991

a. Predictors: (Constant), Paper and pencil

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	144.308	1	144.308	1.942	.167 ^a
	Residual	7133.080	96	74.303		
	Total	7277.388	97			

a. Predictors: (Constant), Paper and pencil

b. Dependent Variable: Online test

ANOVA^b

Model		Sum of Squares	df	Mean Square	F	Sig.
1	Regression	144.308	1	144.308	1.942	.167 ^a
	Residual	7133.080	96	74.303		
	Total	7277.388	97			

a. Predictors: (Constant), Paper and pencil

b. Dependent Variable: Online test

Coefficients^a

Model		Unstandardized Coefficients		Standardized Coefficients	t	Sig.
		B	Std. Error	Beta		
		1	(Constant)	85.925		
	Paper and pencil	-.138	.099	-.141	-1.394	.167

a. Dependent Variable: Online test

Correlations

		Paper and pencil	Online test
Paper and pencil	Pearson Correlation	1	-.141
	Sig. (2-tailed)		.167
	N	98	98
Online test	Pearson Correlation	-.141	1
	Sig. (2-tailed)	.167	
	N	98	98

Statistics

		Paper and pencil	Online test
N	Valid	98	98
	Missing	0	0
Mean		72.9592	75.8367
Median		72.0000	78.0000
Mode		67.00	78.00
Std. Deviation		8.82115	8.66168
Variance		77.813	75.025
Skewness		-.657	-1.021
Std. Error of Skewness		.244	.244
Kurtosis		1.876	1.935
Std. Error of Kurtosis		.483	.483
Range		51.00	46.00
Minimum		45.00	45.00
Maximum		96.00	91.00

Lampiran 4. Personalia Tim Peneliti dan Kualifikasinya

Susunan Organisasi Tim Peneliti dan Kualifikasi

No	Nama/NIDN	Jabatan	Alaokasi waktu	Uraian tugas
1.	Prihastuti Ekawatiningsih	Ketua	10 jam/mg	Penyusunan soal, rancangan program, koordinasi pelaksanaan kegiatan, merancang buku Asesmen Pembelajaran Kejuruan (Ahli Evaluasi Pembelajaran dan Materi)
2.	Endang Mulyatiningsih	Anggota 1	8 jam/mg	Review soal, validasi produk, membantu pelaksanaan FGD, analisis data, merancang Jurnal Ilmiah (Ahli Penelitian Evaluasi Pendidikan)
3.	Sigit Yatmono	Anggota 2	8 jam/mg	Validasi Program (Alat Asesmen), persiapan program untuk keperluan implementasi, Menyiapkan Manual Program Wondershare untuk keperluan ujian on line Tata Hidang (Ahli Programer dan TI)



DAFTAR ISI

Halaman

BAB I KONSEP PENGUKURAN, PENILAIAN DAN EVALUASI.....	63
A. Pengukuran (<i>measurement</i>).....	Error! Bookmark not defined.
B. Penilaian (<i>Assessment</i>).....	Error! Bookmark not defined.
C. Evaluasi (<i>Evaluation</i>).....	Error! Bookmark not defined.
D. Implikasi Pengukuran, Penilaian dan Evaluasi	Error! Bookmark not defined.
BAB II PERAN PENILAIAN DALAM PEMBELAJARAN	Error! Bookmark not defined.
A. Belajar dan Proses Pembelajaran	Error! Bookmark not defined.
B. Faktor Penentu Hasil belajar	Error! Bookmark not defined.
C. Peran Evaluasi dalam Proses Pembelajaran	Error! Bookmark not defined.
BAB III ANALISIS KOMPETENSI DAN TUJUAN PEMBELAJARAN	Error! Bookmark not defined.
A. Analisis Kompetensi	Error! Bookmark not defined.
1. <i>Cognitive Domain</i>	Error! Bookmark not defined.
2. <i>Affective Domain</i>	Error! Bookmark not defined.
3. <i>Psychomotor Domain</i>	Error! Bookmark not defined.
B. Tujuan Pembelajaran	Error! Bookmark not defined.
PENGEMBANGAN TES BAKU	Error! Bookmark not defined.
A. Manfaat Tes Baku	Error! Bookmark not defined.
B. Posedur Pengembangan Tes Baku	Error! Bookmark not defined.
1. <i>Defining the test universe, audience, and purpose</i>	Error! Bookmark not defined.
2. <i>Developing a test plan</i>	Error! Bookmark not defined.
3. <i>Composing the test items</i>	Error! Bookmark not defined.
4. <i>Writing the administration instructions</i>	Error! Bookmark not defined.
5. <i>Conduct Piloting Test</i>	Error! Bookmark not defined.
6. <i>Conduct item analysis</i>	Error! Bookmark not defined.
7. <i>Revising the test</i>	Error! Bookmark not defined.
8. <i>Validation the test</i>	Error! Bookmark not defined.
9. <i>Developing norms</i>	Error! Bookmark not defined.
10. <i>Complete test manual</i>	Error! Bookmark not defined.
BAB V PENGENDALIAN KUALITAS ALAT PENILAIAN	Error! Bookmark not defined.



A.	Reliabilitas	Error! Bookmark not defined.
1.	Stabilitas	Error! Bookmark not defined.
2.	Representatif	Error! Bookmark not defined.
3.	<i>Equivalence</i>	Error! Bookmark not defined.
B.	Validitas	Error! Bookmark not defined.
4.	<i>Face validity</i>	Error! Bookmark not defined.
5.	<i>Content validity</i>	Error! Bookmark not defined.
6.	Validitas kriteria	Error! Bookmark not defined.
7.	Reliabilitas Penilaian Kinerja	Error! Bookmark not defined.
DAFTAR PUSTAKA		Error! Bookmark not defined.

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)¹ 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. Keputusan

¹Griffin, P., & Peter, N. (1991). *Educational assessment and reporting*. Sidney: Harcourt Brace Javanovich Publisher

diambil dengan cara membandingkan hasil pengukurandan asesmendengan standar yang telah ditetapkan (Johnson, 2002: 1)². Untuk lebih memahami makna istilah pengukuran, penilaian dan evaluasi maka berikut ini dibahas lebih mendalam tentang penggunaan masing-masing istilah tersebut di dunia pendidikan.

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 (2012) mendefinisikan *measurement* refers to the process by which the attributes or dimensions of some physical object are determined (pengukuran mengacu pada proses dimana atribut atau dimensi dari beberapa objek fisik ditentukan); (2) Bachman (1990)³ mendefinisikan *measurement* sangat dekat dengan pengkuantifikasian; (3) Okpalla (1999)⁴ 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)⁵ 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 ongoing monitoring and reporting of program accomplishments, particularly progress towards preestablished goals*)

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

²Johnson, D. W., & Johnson, R. T. (2002). *Meaningful assessment, A manageable and cooperative process*. Boston : Allyn and Bacon

³Bachman, L.F. (1990). *Fundamental Considerations in Language Testing*. Oxford: Oxford University Press

⁴Okpalla P. M. et al (1999) *Measurement and Evaluation in Education*. Stching – Horden Publishers (nig.) Ltd. Benin City

⁵Anonim (2005), *Performance Measurement and Evaluation*: United States: General Accounting Office (GAO)

(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

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



Pengukuran ranah pengetahuan

	
<i>Pengukuran tinggi dan berat badan</i>	<i>Pengukuran keterampilan</i>

Gambar 1. Contoh Pengukuran

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 berisitugas-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 psikhis 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.

Penilaian (*Assessment*)

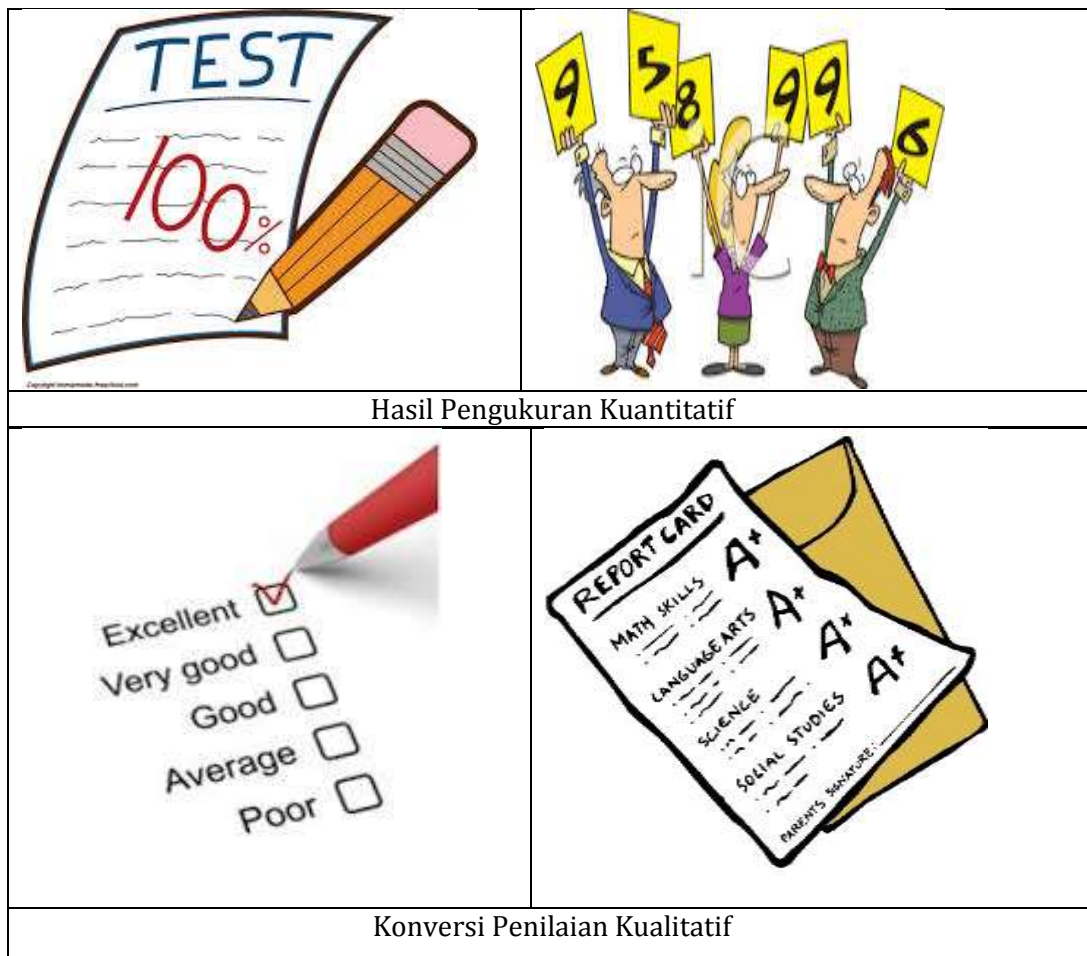
Pembelajaran, belajar, asesmen, dan evaluasi saling berhubungan dan tidak dapat dipisahkan. Guru bertanggung jawab memberi pembelajaran agar siswa dapat belajar. Secara periodik, asesmen dilakukan untuk memberi pertimbangan secara kuantitatif maupun kualitatif tentang hasil belajar atau kompetensi untuk 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 sumberdata 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*. Definisi tersebut mengandung maksud penilaian adalah suatu proses dimana informasi diperoleh untuk mengetahui ketercapaian tujuan. Terry (2011)⁶ 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*. Dalam terjemahan bebas, definisi tersebut 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)⁷ 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.

⁶Terry Overton. (2011). *Assessing Learners with Special Needs: 6th*. Ed diperoleh dari <http://www.enotes.com/ref/q-and-a/distinction-between-assessment-evaluation-201131>

⁷ Johnson. (2002). *Meaningfull assessment, a maanageable and cooperative process*. Boston: Allyn and Bacon

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



Gambar 2. Konversi Penilaian Kuantitatif dan 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

Tabel 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
Sumber: Peraturan Akademik Politeknik Negeri Bandung			0 - 40	E	0.00
Peraturan Akademik UNY					

Setiap lembaga dapat menentukan kriteria sendiri sehingga nilai yang dimunculkan berbeda-beda. Pajang 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 intakepeserta 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 guru memerlukan tingkat pemahaman yang tinggi terhadap kompetensi yang harus dicapai peserta didik, 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. Contoh cara penetapan KKM kompetensi dasar diambil dari materi bimtek Evaluasi: Penetapan KKM pada KTSP oleh Departemen Pendidikan Nasional (Endang Mulyani: 2009)

Tabel 2. Penetapan KKM mata pelajaran

	Indikator	Kriteria nilai		
		Tinggi	Sedang	Rendah
1.	Kompleksitas materi (KM)	1	2	3
2.	Daya dukung sekolah (DD)	3	2	1
3.	Intake kemampuan rata-rata siswa (IS)	3	2	1

Jika indikator memiliki Kriteria: kompleksitas rendah (3), daya dukung tinggi (3) dan intake peserta didik sedang (2) maka nilai KKM-nya adalah

$$\frac{3 + 3 + 2}{9} = 88,9$$

Contoh penerapan penetapan KKM pada beberapa kompetensi dasar (KD) mata pelajaran ilmu gizi dapat disimak pada tabel 2 Kompleksitas materi (KM), daya dukung (DD) dan intake siswa (IS) ditafsirkan sendiri oleh guru karena gurulah yang paling tahu tentang kondisi lingkungan sekolah, kemampuan siswa dan tingkat kesulitan materi.

Tabel 3. Contoh Penetapan KKM pada KD Ilmu Gizi

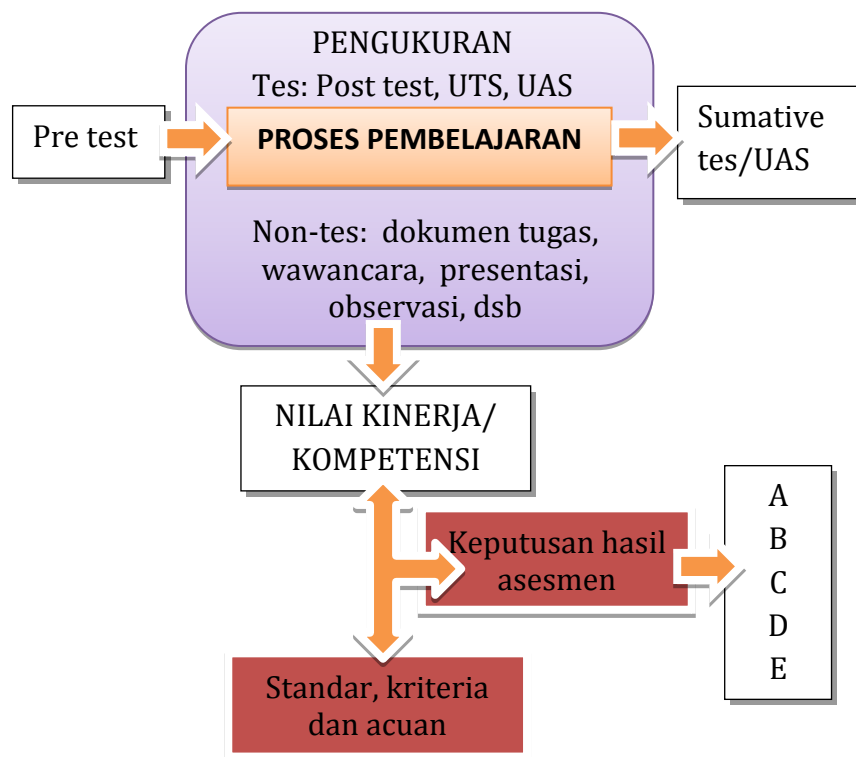
Kompetensi Dasar	Faktor			KKM	
	KM	DD	IS		
4.1. Menganalisis angka kecukupan gizi	2	3	2	7/9	77,8
4.2. Menjelaskan menu seimbang	3	3	2	8/9	88,9
4.3. Menyusun menu untuk bayi dan balita	1	3	2	6/9	66,7

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.

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 gambar3 berikut ini:



Gambar 3: Alur Kegiatan Penilaian Pembelajaran

Penjelasan:

- 1) mengukur hasil belajar dengan menggunakan berbagai macam alat pengukur seperti tes (lisan/tertulis), 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; sangat baik, baik, cukup baik, kurang baik atau menggunakan huruf alphabet (A,B, C, D, E)

Penilaian merupakan tahap yang penting dalam pembelajaran. Suryabrata (1981)⁸ 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.

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.

Belajar pada hakekatnya 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

⁸Sumadi Suryabrata. (1981). *Pengukuran dan penilaian pendidikan*, (Kumpulan makalah penataran bimbingan dan konseling untuk tenaga pengajar di PT se Indonesia, edt). Jakarta: Dirjen Pendidikan Tinggi

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 yang digunakan untuk mengambil keputusan penting (*high stake testing*) seperti penentuan kelulusan, seleksi dan penempatan harus mematuhi prinsip-prinsip penilaian di atas. 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.

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, (1970) 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. (Eds). (1985)⁹ menyatakan evaluasi adalah: *“the process of delineating, obtaining and providing useful information for judging decision alternatives”*. Definisi evaluasi tersebut dapat diartikan sebagai proses yang menggambarkan, memperoleh dan memberikan informasi yang berguna untuk menetapkan beberapa alternatif keputusan; (3) 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 Overton¹⁰ menyatakan *evaluationis 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*. Dalam terjemahan bebas definisi evaluasi tersebut mengandung 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, (1998: 332)¹¹ menambahkan evaluasi diperlukan untuk membantu guru mengetahui tingkat pencapaian tujuan pembelajaran dan membantu guru mengetahui posisi siswa secara individual.

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

⁹Madaus, G. F., & Stufflebeam, D. L. (Eds). (1985). *Systematic Evaluation, evaluation in education and human services*. Massachusetts: Chestnut Hill

¹⁰Terry Overton (2011) *Assessing Learners with Special Needs: 6TH ED* diunduh dari <http://www.enotes.com/ref/q-and-a/distinction-between-assessment-evaluation-2011>

¹¹ Burden, P. R & Byrd, D. M. 1998. *Methods for Effective Teaching*. Needhan Height: Allyn & Bacon

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 evaluasi 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 berkompoten 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 (1984)¹² 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;

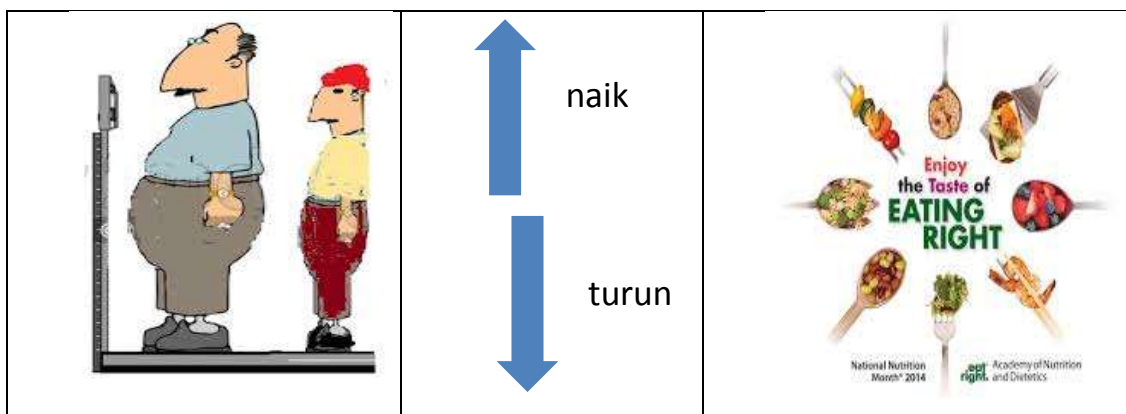
¹²Ogunniyi, M. B. (1984) *Educational Measurement and Evaluation*. Nig Mc. Ibadan: Longman

- 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 obyektif untuk memberi penghargaan, sertifikat, kenaikan kelas
- 14) menjadi dasar untuk membuat ijazah yang berguna untuk melamar pekerjaan

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.

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



Pengukuran	Penilaian	Evaluasi
------------	-----------	----------

Gambar 4. Contoh 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 berikut ini.

	<p>BB sama tapi Lingkar perut beda?</p>	
<p>Pengukuran</p>	<p>Penilaian</p>	<p>Keputusan evaluasi</p>

Gambar 5. Contoh Perbedaan Lingkar Perut Karena Beda Aktivitas

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

Perbedaan antara asesmen dan evaluasi diilustrasikan pada gambar 6. 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 6. Perbedaan Asesmen dan Evaluasi

Implikasi penilaian dan evaluasi dalam kegiatan pembelajaran dirangkum pada tabel berikut:

Tabel 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)

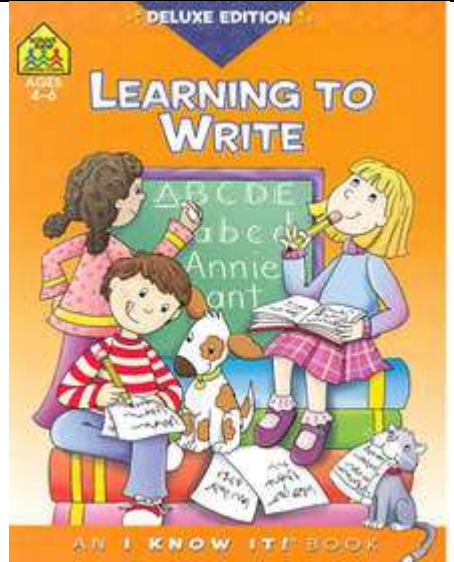
Pengukuran, penilaian, dan evaluasi merupakan serangkaian kegiatan yang dilakukan secara berurutan. Pengukuran dilakukan pertama kali ketika guru ingin mengetahui kemampuan siswa dengan menggunakan alat ukur tes atau non tes. Penilaian dilakukan untuk memberikan informasi tentang kemampuan yang diperoleh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

BAB II PERAN PENILAIAN DALAM PEMBELAJARAN

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 (2013)¹³ 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.

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 metoda dan media agar ilmu pengetahuan, keterampilan, sikap yang ditransferkan lebih mudah diterima peserta didik. Untuk mengetahui hasil belajar dilakukan evaluasi hasil belajar. Berdasarkan ilustrasi tersebut dapat diidentifikasi komponen pembelajaran yaitu terdiri dari: (1) peserta didik dan pendidik; (2) metoda dan media; (3) tujuan pembelajaran, materi yang dipelajari dan evaluasi hasil belajar. Contoh pembelajaran tertera pada gambar 7 Tugas saudara; identifikasilah komponen apa saja yang terdapat pada gambar tersebut.

<p>TUJUAN belajar agar:</p> <ul style="list-style-type: none"> • Tahu • Paham • Bisa • ? 		<p>HASIL belajar sudah:</p> <ul style="list-style-type: none"> • Tahu • Paham • Bisa • Yes, OK
Rencana	Proses	Evaluasi

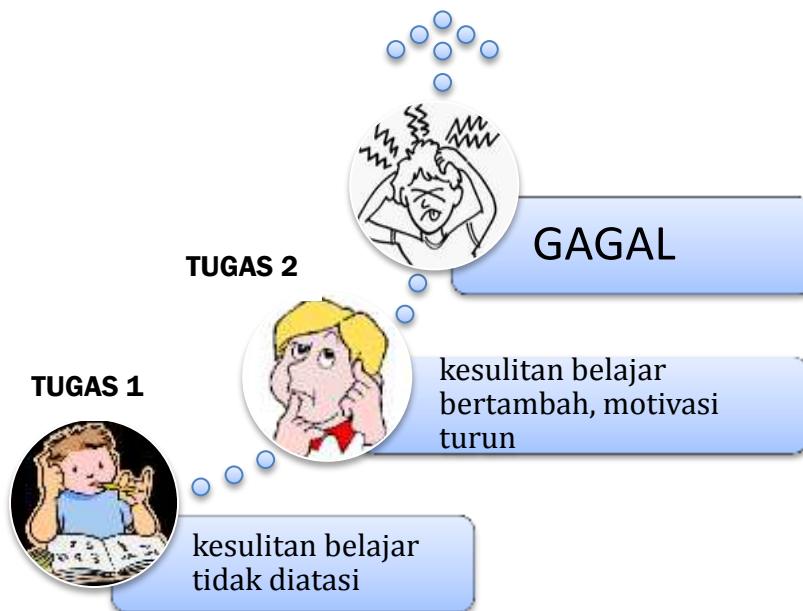
¹³Ady Ferdian Noor, M.Pd, (2013). Modul 1-4: *Pengertian, hakekat, dan teori belajar dan pembelajaran*: Universitas Muhammadiyah Palangkaraya

Gambar 7. Proses Pembelajaran

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

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 pada tahap pertama tidak diatasi maka akan berakibat seperti yang tertera pada ilustrasi di gambar 8.

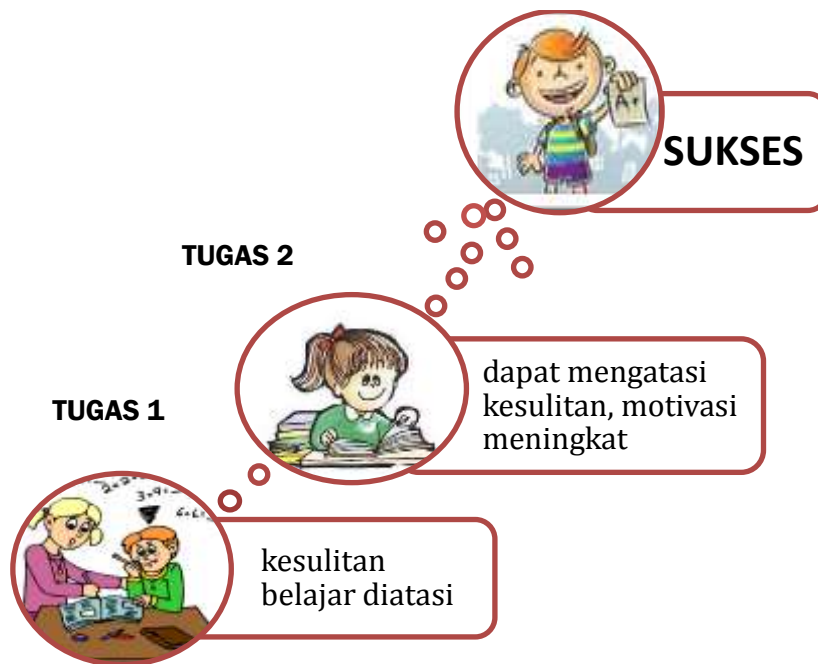


Gambar 8. Pengalaman Gagal

Hasil belajar dapat memberi pengaruh pada motivasi belajar tahap berikutnya. Jika kesulitan belajar tahap pertama tidak diatasi maka belajar tahap kedua akan semakin sulit dan dapat diprediksi hasil belajar akan gagal. 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 9.

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.



Gambar 9. Pengalaman Sukses

Bloom dalam Haladyna (1982: 16)¹⁴menggambarkan model regresi linier untuk menjelaskan tentang pengaruh potensi kognitif dan afektif terhadap prestasi belajar sebagai berikut:

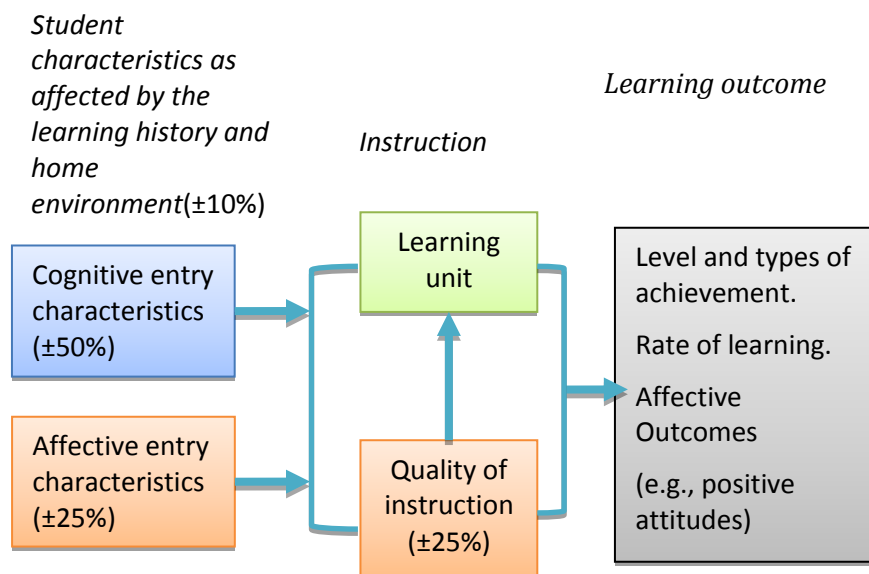
$$\hat{Y} = F(X1, X2, X3)$$

Dimana \hat{Y} adalah prestasi yang diukur dengan test acuan kriteria, $X1$ adalah *cognitive entry characteristics*, $X2$ adalah *affective entry characteristics* seperti sikap dan kepribadian, dan $X3$ 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

¹⁴Roid, G. H., & Haladyna, T. M. (1982). *A technology for test-item writing*. New York: Harcourt Brace Jovanovich

oleh sejarah keluarga, lingkungan rumah, dan latar belakang lain. Bloom (1976: 33)¹⁵ menjelaskan pengaruh kognitif, afektif dan kualitas pembelajaran terhadap prestasi belajar seperti pada gambar 2

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 (1976: 75) 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*).



Gambar 10. Pengaruh Karakteristik Masuk dan Kualitas Pembelajaran pada *Learning Outcome*

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)

¹⁵Bloom, B. S. (1976). *Human characteristic and school learning*. New York: McGraw-Hill book Company

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 bacaan (Bloom's, 1976).

Hamilton (2003: 1-29) melaporkan hasil penelitian ada hubungan antarprestasi siswa dipengaruhi oleh praktek perubahan yang dilakukan guru, latar belakang siswa seperti status sosio ekonomi dan etnik. Willingham, Pollack, Lewis, (2002: 1-3) menemukan enam kelompok variabel yang berpengaruh terhadap prestasi siswa, yaitu: karakteristik kemampuan siswa, inisiatif, kegiatan kompetisi, latar belakang keluarga, sikap dan penilaian guru. 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 subyektif guru. Perbedaan kemampuan antara pria dan wanita dalam matematika diteliti oleh Gierl (2003: 281-306) dan Bielinski (2001: 51-77) 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 dan Phillips (1986: 185-196) 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.

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, (1998: 332) 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 secara tidak langsung 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 berikut:

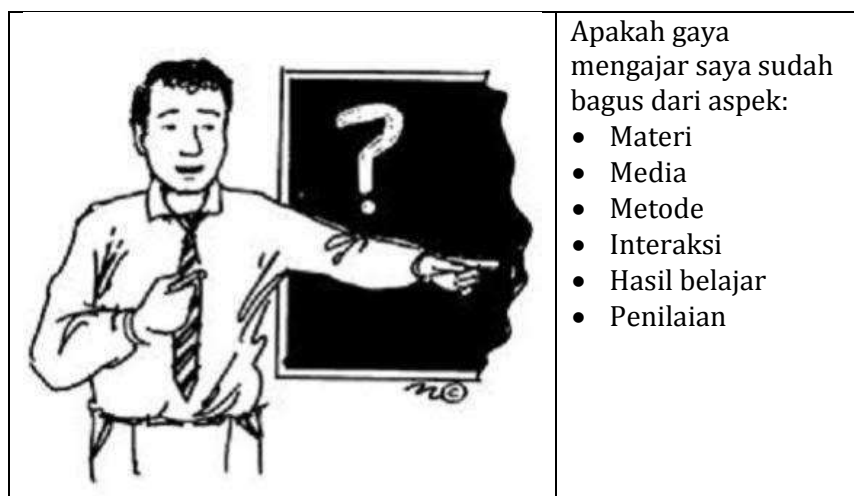
Usaha	Hasil	Dampak
		
<p>Kerja keras, nilai bagus, motivasi meningkat, giat belajar lagi</p>		



Gambar 11. 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?
- (2) Apakah **media/perangkat** pembelajaran yang digunakan oleh pendidik sudah dapat membantu pemahaman peserta didik?
- (3) Apakah **gayamengajar** 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?



Gambar 12. 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 5. Contoh Lembar Observasi Pelaksanaan Pembelajaran

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

NO	INDIKATOR/ASPEK YANG DIAMATI	SKOR
11	Melaksanakan pembelajaran yang memungkinkan timbulnya kebiasaan positif	1 2 3 4 5
12	Melaksanakan pembelajaran yang sesuai dengan alokasi waktu yang direncanakan	1 2 3 4 5
C	Pemanfaatan sumber belajar/media pembelajaran	
13	Menggunakan media secara efektif dan efisien	1 2 3 4 5
14	Menghasilkan pesan yang menarik	1 2 3 4 5
15	Melibatkan siswa dalam pemanfaatan media	1 2 3 4 5
D	Pembelajaran yang memicu dan memelihara keterlibatan siswa	
16	Menumbuhkan partisipasi aktif siswa dalam pembelajaran	1 2 3 4 5
17	Menunjukkan sikap terbuka terhadap respon siswa	1 2 3 4 5
18	Menumbuhkan keceriaan dan antusiasme siswa dalam belajar	1 2 3 4 5
E	Penilaian proses dan hasil belajar	
19	Memantau kemajuan belajar selama proses	1 2 3 4 5
20	Melakukan penilaian akhir sesuai dengan kompetensi (tujuan)	1 2 3 4 5
F	Penggunaan bahasa	
21	Menggunakan bahasa lisan dan tulis secara jelas, baik dan benar	1 2 3 4 5
22	Menyampaikan pesan dengan gaya yang sesuai	1 2 3 4 5
III	PENUTUP	
23	Melakukan refleksi atau membuat rangkuman dengan melibatkan siswa	1 2 3 4 5
24	Melaksanakan tindak lanjut dengan memberi arahan, atau kegiatan atau tugas sebagai bagian remidi/pengayaan	1 2 3 4 5
	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.

BAB III ANALISIS KOMPETENSI DAN TUJUAN PEMBELAJARAN

Analisis Kompetensi

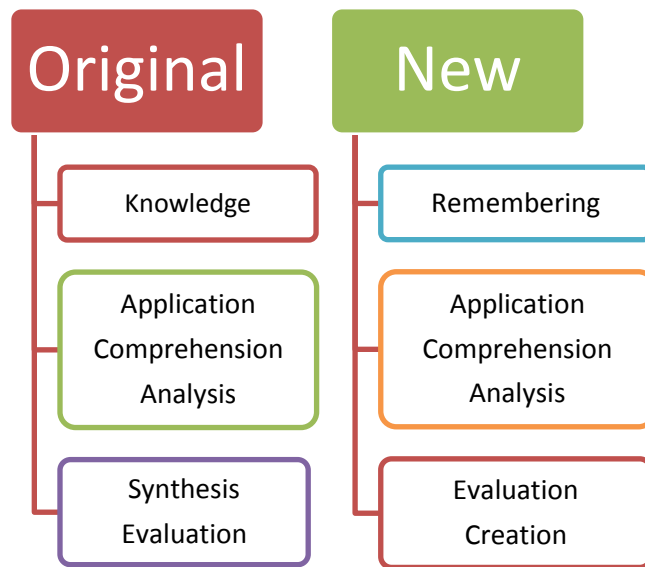
Hasil belajar pendidikan kejuruan dinyatakan dalam rumusan kompetensi. Menurut definisi dari dictionary.com (<http://dictionary.reference.com>) kompetensi adalah kualitas yang menjadikan kompeten, keadeguan, kepemilikan keterampilan, pengetahuan, kualifikasi atau kapasitas yang dituntut (*the quality 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, psicomotor and other*) atau 3H (*Head, Hand, Heart*). Spektrum SMK terdiri dari banyak keahlian, oleh sebab itu selain kompetensi umum yang terdiri dari 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

Cognitive Domain

Bloom (1964) menyusun tingkatan kompetensi *cognitive* menjadi tujuh tingkat dari dasar sampai mahir seperti tertera pada gambar 1. Kompetensi kognitif tingkat paling rendah (mengingat) sedangkan tingkat yang paling tinggi (kreasi). Secara berturut-turut, domain kognitif terdiri dari: pengetahuan/ingatan, pemahaman, aplikasi, analisis, evaluasi dan kreasi. Secara berturut-turut, domain kognitif dari Bloom dinyatakan pada Gambar 13.



Gambar 13. *Cognitive Domain*(Bloom:1994)

Sub-Domain Cognitive

a. Remembering (Ingatan)

“Ingatan” merupakan level domain kognitif yang paling rendah. Evaluasi hasil belajar sebaiknya tidak terlalu banyak mengukur aspek “ingatan” karena dapat merangsang siswa menyontek dan kurang mampu melatih siswa untuk berpikir kreatif memecahkan masalah. Kompetensi kognitif level “ingatan” 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.

Tabel 6. Kata Kerja Operasional

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>)
Menghafal	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 mengubahnya menjadi kata tanya atau perintah pada saat mengevaluasi hasil belajar

Contoh:

Tujuan Pembelajaran

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

Soal evaluasi hasil belajar

- Apa yang dimaksud dengan arti masakan tradisioal?
- 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 mengingat-ingat atau menghafal.

Kompetensi kognitif level “pemahaman” biasanya diukur dari *kemampuanseseorang 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. Tingkat kesulitan kompetensi “ingatan dan pemahaman” masih setara dan bisa digabungkan menjadi satu. 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 sampling yang tepat untuk meneliti studi kasus anak kekurangan energi dan protein
- Siswa dapat menggunakan teori 7P dalam penyusunan strategi pemasaran

Soal evaluasi hasil belajar

- Hitunglah kebutuhan bahan makanan untuk memasak 100 porsi bakmie goreng
- Ubahlah resep kroket kentang menjadi kroket talas
- Pilihlah teknik sampling yang tepat untuk meneliti studi kasus anak kekurangan energi dan protein
- 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
mendiagnosis	Mengilustrasikan
mem-break-down	Mengelompokkan
Mengidentifikasi	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?; Mengapa? ; Bagaimana? Misalnya:

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

f. 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	Menerangkan
Memperbaiki	Membetulkan
Membandingkan	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 design 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.

g. 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 mengkombinasikan 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	Mengkategorikan
Merancang/mendesain	Mengkombinasikan
Membuat (konsep)	Menata
Mengorganisasikan	

Contoh:

Tujuan pembelajaran

- Siswa dapat menyusun menu makan siang keluarga yang memiliki variasi bahan, teknik olah dan rasa
- Siswa dapat menyusun rencana usaha (*bussines plan*) catering
- Siswa dapat merancang strategi pemasaran kue dodol carica

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
- Rancanglah strategi pemasaran dodol carica

Affective Domain

Krathwohl, Bloom, Masia, (1973)¹⁶ 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*),
- ❖ penghargaan (*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 menjadikebiasaan (*habit*) hidup sehari-hari dapat diceritakan sebagai berikut:

- (1) penerimaan 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

¹⁶Krathwohl, D.R., Bloom, B.S., and Masia, B.B. (1964). Taxonomy of educational objectives: Handbook II: Affective domain. New York: David McKay Co.

- melihat orang membuang sampah sembarangan. Siswa masih bersikap pasif yaitu hanya melihat atau mendengar cerita tanpa memberi respon apapun.
- (2) respon 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, dan menjadi sukarelawan, 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 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. 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, p. 28)*¹⁷. Internalisasi mengacu pada proses dimana seseorang bersikap terhadap suatu objekdimana sikap tersebut 'diinternalisasikan' dan secara konsisten memandu atau mengendalikan perilaku seseorang (Seels & Glasgow, 1990, hal. 28).

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 itu pun 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.

Berikut ini diilustrasikan tentang pembentukan sikap/kebiasaan merokok/tidak merokok melalui gambar sebagai berikut

¹⁷Seels and Glasgow (1990). Exercises in instructional design. Columbus OH: Merrill Publishing Company.



Gambar 14. Pembentukan Kebiasaan Merokok

Penilaian domain afektif dalam proses pembelajaran

Domain afektif merupakan domain yang paling sulit diuku 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, kesediaan 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 seksama
- 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 diskusi
- 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 konsistendan 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

Psychomotor Domain

The psychomotor domain (Simpson, 1966)¹⁸ 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.

¹⁸ Simpson, J. S. (1966). The classification of educational objectives, psychomotor domain. Office of Education Project No. 5-85-104. Urbana, IL: University of Illi'nois

Dave's dan Atkinson's (2012)¹⁹ membagi domain psikomotor menjadi lima sub domain yang lebih nyata untuk diukur. Taksonomi sub domain psikomotor dijelaskan pada Tabel 7.

Tabel 7. Tingkat Kompetensi Psikomotorik

Simpson's	Dave's original	Atkinson's adaptation
Perception		
Set/Readiness		
Guided response	Imitation	Imitate
Mechanism	Manipulation	Manipulate
Complex	Precision	Perfect
Adaptation	Articulation	Articulate
Origination	Naturalisation	Embody

Taksonomi kompetensi psikomotor yang dikemukakan oleh tiga ahli pada tabel 7. 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 indramisalnya 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 didalam otak timbul berbagai macam persepsi terhadap objek minuman tersebut; (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 disuruh 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 berikutnyadan kata kerja yang digunakan untuk mengukur ketercapaian kompetensi dijelaskan pada paparan berikut ini.

a. Imitation(peniruan)

¹⁹ Atkinson, S. (2015 Juli 9). Graduate Competencies, Employability and Educational Taxonomies: Critique of Intended Learning Outcomes. *Practice and Evidence of the Scholarship of Teaching and Learning in Higher Education* [Online] 10:2, pp: 154-177 Available: <http://community.dur.ac.uk/pestlhe.learning/index.php/pestlhe/article/view/194/281>

Imitasi atau meniru yaitu menirukan pola perilaku yang telah diamati dari orang lain. Menurut Atkinson's (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 (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 mengoperasikan oven microwave setelah membaca manual
- Siswa dapat memegang pisau dengan benar sesuai dengan petunjuk di gambar
- 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 seragam
- Mahasiswa dapat menata meja makan (table set up) dengan rapi dan tanpa kesalahan

d. Articulation (artikulasi):

Setelah mencapai tingkat presisi, keterampilan (skill) dapat ditingkatkan 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:

- Siswa dapat memegang mixer sambil menuang putih telur ke dalam adonan yang dimixer (ada koordinasi tangan kanan dan kiri dengan gerakan yang berbeda)
- 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 membalik omlet tanpa bantuan alat (keterampilan yang kompleks)
- Siswa dapat bermain juggling (keterampilan yang kompleks)

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, dsb

Contoh tujuan pembelajaran keterampilan level naturalisasi

- Siswa dapat mengoperasikan *food proccesor* dengan lancar

- Siswa dapat mencetak 100 kue semprit dengan bentuk yang seragam dalam waktu 10 menit
- Siswa dapat memotong 10 wortel berbentuk jardiniere sesuai standard dalam waktu 10 menit
- Siswa dapat membungkus 10 lempeng dengan rapi dalam waktu 10 menit
- Siswa dapat memorsi kentang goreng dengan berat yang sama
- Siswa dapat membuat hiasan bentuk kerang pada tart ulang tahun secara simetris dalam waktu 30 menit

Contoh tahap perkembangan belajar keterampilan memfilet ikan diilustrasikan dengan tahap-tahap yang terdapat pada gambar 15.



Gambar 15. Tahap perkembangan belajar keterampilan

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 16.



Gambar 16. Konsistensi Antara Tujuan dan Evaluasi

Untuk mengukur kompetensi sudah/belum tercapai maka ditetapkan indikator pencapaian kompetensi. 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 dari 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* = peringkat/ukuran). Susunan kalimat pada tujuan pembelajaran diawali dengan menyebutkan pelaku (*audience*) yaitu "siswa" kemudian diikuti dengan menyebutkan "kompetensi" atau perilaku akhir yang diharapkan dapat dilakukan siswa dengan menggunakan kata kerja operasional. Contoh SPOK:

- (1) Subjek dengan menyebutkan kata "siswa",
- (2) Predikat menggunakan kata kerja operasional yang dapat diukur dan diamati, dan
- (3) Objek berupa perilaku/kompetensi yang diharapkan akan dicapai.
- (4) Keterangan 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:

- (1) "siswa dapat membuat adonan dasar pizza dengan teknik stright dough; Tujuan pembelajaran dapat dianalisis komponen SPOK yang menyusunnya yaitu: **siswa** (subjek atau audiens) **dapat membuat** (predikat atau behavior) **adonan dasar pizza** (objek) dengan **teknik stright dough** (keterangan)
- (2) siswa dapat membuat 100 gram saus tomat dengan baik selama 20 menit". Siswa (subjek atau audiens) dapat membuat (predikat atau behavior) 100 gram saus tomat (objek) dengan baik selama 20 menit (keterangan dan ukuran)

Prinsip yang digunakan dalam perumusan tujuan adalah **SMART** yaitu: *Spesific, 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. Kriteria tujuan pembelajaran yang telah memenuhi indikator SMART adalah kalimat pada tujuan dapat diubah menjadi pertanyaan, kriteria jawaban benar yang diinginkan dari pertanyaan tersebut jelas, dapat diukur atau diobservasi.



Gambar 17. Prinsip Perumusan Tujuan

Contoh tujuan pembelajaran yang konsisten dengan evaluasi yang memenuhi kriteria spesifik, terukur, rasional, dan dapat tercapai dalam batas waktu yang telah ditetapkan.

Tujuan pembelajaran

'setelah akhir pelajaran, mahasiswa dapat menulis dua butir soal *teknik sampling purposive* dengan bentuk jawaban pilihan ganda (*multiple choise*) yang mengukur ranah cognitive level aplikasi (CA)

Soal evaluasi hasil belajar

Tulislah dua butir soal pilihan ganda untuk materi *teknik sampling* pada level aplikasi (Cognitive, Aplication)

Jawaban soal/tugas:

1. Penelitian tindakan dengan judul "Peningkatan hasil belajar Ilmu Gizi melalui metode *snow ball throwing* pada siswa kelas X jasa Boga SMK N 6 Yogyakarta menggunakan teknik sampling
 - a. Acak sederhana
 - b. Purposive
 - c. Snow ball
 - d. Quota
2. Penelitian dengan judul "Kesulitan belajar yang dialami mahasiswa krisis masa studi" menggunakan teknik sampling

- a. purposive
- b. stratified sampling
- c. aksidental
- d. cluster sampling

Indikator SMART pada tujuan pembelajaran tercermin dari: materi teknik sampling purposivesudah *specific* karena meskipun terdapat banyak teknik sampling tetapi yang ditanyakan hanya teknik sampling purposive. Measurable tercermin dari kemampuan mahasiswa menulis dua butir soal dapat diukur benar atau salah. Achievable, rational dan timely tercermin dari kemampuan menulis 2 butir soal pilihan ganda dapat dicapai dan ditunjukkan setelah akhir 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:

Tabel 8. Contoh KK Operasional dan Kurang Operasional

KK	Kata tanya atau perintah	Keterangan
Menyebutkan	Sebutkan	KKO
Menjelaskan	Jelaskan	KKO
Membuat	Buatlah	KKO
Membedakan	Bedakan	KKO
Membandingkan	Bandingkan	KKO
Memahami		KKKO
Mengetahui		KKKO

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

No	Tujuan Pembelajaran	Domain	Sub domain
1.	Pada akhir pembelajaran, siswa dapat menjelaskan prosedur pelayanan makan ala English	Cog	Compre-hensip
2.	Siswa dapat menganalisis perbedaan sistem pelayanan ala Rusia dan ala Inggris	Cog	Analysis
3.	Siswa dapat menjelaskan tekstur kue bolu panggang yang menggunakan margarin dan tanpa margarin	Cog	Analysis
B.	Siswa bersedia mendiskusikan materi yang diajarkan guru	Aff	Respon-ding
	Siswa dapat mengulangi gerakan membalik telur dadar	Psi	Imitasi
	Siswa dapat menata meja ala rusia	Psik	Manipulasi

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 mencetak tumpeng nasi kuningselama 15 menit.
- Tidak boleh :
- Siswa dapat menjelaskan prosedur pembuatan nasi uduk dan nasi kuning (dua objek yang berbeda)
 - Siswa memiliki kemampuan mengolah nasiberbumbu (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 dapat diketahui, atau diubah menjadi kata tanya pada evaluasi sehingga sebaiknya diganti dengan kata menjelaskan)
 - Siswa dapat membuat garnish dari buah tomat (tidak disebutkan jenis garnish, serta batasan jumlah dan waktunya)

c. Achievable/attainable:

Kompetensi yang tertulis dalam tujuan pembelajaran harus dapat dicapai dan terdapat cara untuk mencapainya. Kompetensi cukup penting untuk dimiliki oleh siswa.

- 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
- 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 tidak perluterlalu banyak sulit dicapai, 10 saja sudah cukupkarena nama minuman panas asli daerah Jawa Tengah tidak sebanyak itu)

d. Realistic

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

- Boleh : • Siswa dapat merangkum materi kuliah bahan pangan dalam bentuk *mind mapping* dalam waktu 5 menit
- Siswa dapat mencetak 100 kue semprit dengan bentuk yang seragam dalam waktu 1 jam.
- Tidak boleh : • Siswa dapat merangkum materi kuliah bahan pangan menjadi 100 halaman (tidak realistik karena jumlah rangkuman masih terlalu banyak)
- Siswa dapat membentuk 100 kue nastar dalam waktu 1 jam (tidak realistik untuk dicapai karena kue dibentuk manual dengan tangan)

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 memadai dan jelas batasannya

- 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)

Tujuan pembelajaran dan indikator pencapaian kompetensi memiliki orientasi yang sama yaitu hasil yang diharapkan dicapai oleh siswa setelah selesai pembelajaran. Oleh sebab itu, penulisan tujuan pembelajaran dan indikator pencapaian kompetensi boleh sama.

D. Kesimpulan

Analisis kompetensi merupakan kegiatan mengenali bagian-bagian kompetensi yang dibutuhkan dalam bekerja dan merumuskan kembali kompetensi tersebut serangkaian kompetensi untuk dapat mencapai kompetensi yang diharapkan. Upaya analisis kompetensi dilakukan secara bertahap mulai dari level penentu kebijakan sampai pada level pelaksana kebijakan. Guru sebagai pelaksana kebijakan memiliki wewenang dan tugas untuk menganalisis kompetensi yang dapat dicapai dalam proses pembelajaran. Kompetensi tersebut dirumuskan dalam tujuan pembelajaran.

E. Soal Latihan

Rumuskan tujuan pembelajaran sesuai dengan contoh dari domain kognitif, afektif dan psikomotor. Ambil kompetensi keahlian Patiseri atau Jasa Boga dari Permendiknas nomor 28 tahun 2009 tentang Standar Kompetensi Lulusan SMK

Lampiran 6. Artikel Jurnal

PENGEMBANGAN ALAT ASESMEN HASIL BELAJAR TATA HIDANG BERBANTUAN KOMPUTER

Prihastuti Ekawatiningsih¹⁾,
Endang Mulyatiningsih²⁾; Sigit Yatmono³⁾

ABSTRAK

Dunia pendidikan menghadapi banyak tantangan akibat perkembangan teknologi informasi berbasis komputer yang sangat pesat. Di masa depan, semakin banyak dibutuhkan perangkat pembelajaran yang dioperasikan dengan komputer. Untuk menjawab tantangan tersebut, pada penelitian ini dikembangkan perangkat pembelajaran yaitu alat asesmen hasil belajar Tata Hidang yang dioperasikan menggunakan komputer. Kegiatan penelitian tahun pertama telah berhasil mengembangkan bank soal ujian Tata Hidang dan merancang program computerized test menggunakan *software wondershare*. Penelitian tahun kedua bertujuan untuk mengimplementasikan penyelenggaraan ujian Tata Hidang berbantuan komputer dan menganalisis validitas *concurrent* hasil ujian menggunakan komputer dengan hasil ujian *pencil and paper test*.

Penelitian ini merupakan penelitian dan pengembangan (R&D) yang menggunakan model ADIE (*analysis, design, implementation, and evaluation*). Pada tahun pertama, penelitian telah sampai tahap *analysis* dan *design*. Penelitian tahun kedua difokuskan pada kegiatan *implementation, and evaluation*. Subjek uji coba hasil rancangan soal ujian Tata Hidang adalah 100 orang mahasiswa yang mengambil mata kuliah Tata Hidang. Metode pengumpulan data menggunakan dokumentasi respon butir soal ujian Tata Hidang. Analisis data respon butir soal ujian menggunakan Analisis Iteman sedangkan validitas *concurrent* menggunakan analisis korelasi *product moment*.

Hasil penelitian menunjukkan bahwa untuk pengembangan alat penilaian masih ada beberapa revisi antara lain: dengan penambahan suara pada kompetensi I untuk pembacaan soalnya. Disamping itu ada beberapa tampilan yang diperbaiki berkaitan dengan lay out program. Program komputer yang digunakan untuk ujian Tata Hidang secara *on line* atau *computerized test* adalah *software Wondershare*. Pada saat merancang program telah disusun langkah-langkah pembuatan untuk programmer dan langkah-langkah pengoperasian pengguna. Telaah kuantitatif menunjukkan bahwa 100 butir soal *multiple choice* sudah diujikan kepada mahasiswa dengan bantuan program komputer. Telaah kualitas butir soal ujian Tata Hidang secara kuantitatif telah membuktikan 73% soal berada pada tingkat kesulitan sedang, 52% butir soal memiliki daya pembeda sangat baik dan 28% soal memiliki daya pembeda yang baik. Nilai reliabilitas soal sebesar 0,945 telah memenuhi kriteria soal yang baik, dengan angka validitas konkuren 0,167.

Kata kunci: *asesmen* hasil belajar, tata hidang, *computerized test*

DEVELOPMENT TOOL OF LEARNING ASSESMENT PROCEDURES FOOD AND BEVERAGES SERVICE COMPUTER AIDED

Prihastuti Ekawatiningsih¹⁾,
Endang Mulyatiningsih²⁾; Sigit Yatmono³⁾

ABSTRACT

This study aims to: (1) analyze the scope of material that would be tested Food and Beverages Service Procedures; (2) examine the quality of the test items were qualitatively and quantitatively; (3) designing a viable software used for computer-aided examination Food and Beverages Service. Research and development Adie model (analysis, design, implementation, and evaluation). Validator draft assesment tool is 3 expert assessment, 2 programmers, and 4-person teams Food and Beverages Servicelecturer. Design validation data collected through the Delphi technique and FGDs (focus group discussion). Subject trial design assesment tool is 100 students who take courses Food and Beverages Service. Analysis of data from trials about the use Iteman analysis. The results of the analysis of the material set 8 basic competence Food and Beverages Service arranged into question multiple choice options. The results of the qualitative study showed about already tested with some improvements feasible alternative answers and statements that are too long. The results of a quantitative study to obtain 77% of items had a moderate level of difficulty and 48% have a distinguishing excellent. Computer aided design of software tests using WondershareQuiz Creator and worth developing.

Keywords: assesment of learning outcomes, food and beverages service , computerized test

PENDAHULUAN

Hasil belajar diukur dan dinilai dengan menggunakan berbagai alat asesmen (penilaian). Hasil belajar dinyatakan bagus jika hasil asesmen menunjukkan nilai yang tinggi. Namun hasil asesmen yang bagus belum tentu menunjukkan kemampuan belajar yang tinggi karena banyak alat asesmen dan asesornya kurang objektif dalam memberikan penilaian. Untuk mengatasi hambatan tersebut, pada saat ini mulai banyak dikembangkan alat-alat asesmen hasil belajar yang dioperasikan dengan komputer (*computerized test*). Dengan menggunakan alat ini, hasil belajar diuji dengan seperangkat tes yang dapat ditampilkan oleh komputer, peserta tes mengerjakan soal tes langsung di komputer dan komputer juga langsung memberikan skor. Hasil tes dapat dilihat oleh peserta tes pada saat itu juga.

Penyelenggaraan tes berbantuan komputer memiliki beberapa keunggulan antara lain: (1) dapat mengurangi biaya cetak soal tes; (2) mengurangi waktu dan tenaga dosen untuk mengoreksi jawaban peserta tes karena komputer langsung dapat memberikan skor tes; (3) penyelenggaraan tes lebih efisien karena semua perangkat tes dapat di *download* dari internet sehingga dapat dikerjakan dari

jarak jauh. Meskipun demikian, ada sedikit kelemahan dari penyelenggaraan tes menggunakan komputer yaitu harus tersedia fasilitas komputer dan jaringan internet jika tes akan diselenggarakan secara *on-line*. Dalam penilaian *on-line*, perlu dirancang sistem yang tidak memberi peluang kepada orang lain yang bukan peserta tes/mahasiswa yang diuji turut membantu pengisian jawaban tes dan tugas melalui jaringan komputer.

Materi kuliah Tata Hidang terdiri dari materi kuliah teori dan praktik. Materi kuliah teori terdiri dari sub materi yang berisi banyak mengenalkan nama alat, bahan, prosedur, dan istilah yang menuntut siswa untuk tahu dan dapat menerapkannya dalam pekerjaan sehari-hari. Materi kuliah Tata hidang berpotensi dikembangkan menjadi bank soal MC (multiple choice) yang disimpan dan dijalankan menggunakan komputer. Untuk menjawab masalah tersebut, penelitian tahun pertama bertujuan:

1. Menganalisis materi Tata Hidang yang layak dikembangkan menjadi kisi-kisi soal tes pilihan ganda
2. Menelaah kualitas butir soal tes secara kualitatif dan kuantitatif
3. Merancang program komputer yang layak digunakan untuk ujisn Tata Hidang berbantuan komputer

TINJAUAN PUSTAKA

Alat asesmen menggunakan bantuan komputer sudah mulai dikembangkan di beberapa negara. Clauser, et.al (2000) pernah mengembangkan *software* evaluasi kinerja di bidang kedokteran. Kompetensi yang diuji adalah keterampilan menangani pasien. Soal tes berupa permasalahan yang dihadapi pasien. Peserta ujian merespon dengan berbagai tindakan penanganan medis. Komputer secara otomatis memberi skor berdasarkan transaksi tindakan yang telah dilakukan peserta ujian. Berdasarkan hasil analisis ternyata skor diperoleh dari penilai ahli (manusia) mempunyai korelasi yang cukup tinggi dengan skor yang dibuat oleh komputer.

Sebuah penelitian yang dilakukan oleh Lunz, & Bergstrom (1994: 251-263) tentang evaluasi penyelenggaraan CAT (*computerized adaptive test*). Hasil penelitian menunjukkan bahwa pengkondisian CAT berpengaruh pada hasil tes. Setelah dicoba dengan empat kondisi 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. Secara berturut-turut hasil tes akan lebih baik jika peserta tes diberi kesempatan

memilih soal yang bisa terlebih dulu (*skip*), kemudian mengoreksi kembali jawabannya (*review*), dan menunda untuk menjawab (*defer*). Hasil yang paling buruk adalah jika peserta tes dibiarkan tanpa kondisi.

Dua hasil penelitian di atas menunjukkan bahwa tes yang menggunakan bantuan komputer lebih baik dan efisien digunakan. Perangkat tes berbantuan komputer lebih banyak pada tes kognitif. Pengukuran hasil belajar kognitif oleh komputer dapat dilakukan jika tes berbentuk pilihan ganda. Pada tes ini harus dipastikan hanya ada satu jawaban yang benar supaya komputer dapat memberi skor yang benar.

Istilah pengukuran (*measurement*), asesmen (*assessment*), dan evaluasi (*evaluation*) menjadi satu rangkaian kegiatan yang tidak dapat dipisahkan. Huitt (2001: 1) menjelaskan bahwa *assessment* mengacu pada pengumpulan data untuk memahami sebuah isu, objek secara lebih baik. *Measurement* adalah proses mengkuantitatifkan data *assessment*, dan *evaluation* menekankan pada perbandingan antara data yang ada dengan standar yang telah ditetapkan. Griffin (1992: 5) memberi contoh penggunaan istilah *assessment*, *measurement* dan *evaluation* sebagai berikut: (1) penetapan angka terhadap objek yang diobservasi atau 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. Evaluasi diperlukan untuk membantu guru mengetahui tingkat pencapaian tujuan pembelajaran dan membantu guru mengetahui posisi siswa secara individual (Burden & Byrd, 1998).

Untuk memperoleh data yang diperlukan dalam evaluasi hasil belajar, diperlukan instrumen atau alat pengukuran. Alat yang digunakan sangat tergantung pada tujuan pengukuran. Alat yang digunakan untuk mengukur aspek kognitif berbeda dengan alat pengukur aspek afektif dan psikomotor. Pendidik dapat menggunakan berbagai alat pengukuran secara komplementer (saling melengkapi) sesuai dengan kompetensi yang dinilai (kognitif, afektif dan psikomotorik).

Alat ukur tes paling sering digunakan untuk mengukur kemampuan kognitif siswa. Dalam bidang evaluasi pembelajaran dikenal tiga macam tes yaitu tes tertulis, tes lisan dan tes kinerja. Tes tertulis terdiri dari beberapa macam bentuk. Menurut karakteristik jawaban, dikenal tes yang jawabannya sudah tersedia dan tes yang jawabannya ditulis sendiri oleh peserta tes. Jenis tes yang

jawabannya sudah tersedia dan peserta tes tinggal memilih jawaban yang benar dikategorikan pada tipe *objective test* karena jawaban benar langsung mendapat skor dan jawaban salah tidak mendapat skor, tanpa ada pertimbangan lain dari penilai. Tes objektif terdiri dari tes pilihan ganda atau MC (*multiple choice*), benar-salah (*true-false items*), menjodohkan atau mencocokkan (*matching exercises*), isian singkat (*short answer, fill-in items*).

Pembuatan perangkat tes baku perlu memperhatikan karakteristik khusus tes seperti konstruk/domain/isi yang akan diukur, bentuk pertanyaan, bentuk penyelenggaraan dan bentuk penyekoran tes. Konstruk tes dikembangkan dari komponen materi atau lebih dikenal dengan kisi-kisi tes. Hasil tes yang dapat diberi skor objektif adalah tes dalam bentuk pilihan ganda/MC (*multiple choice*), menjodohkan atau benar-salah (*true-false*). Butir MC yang terlalu mudah atau terlalu sulit dapat menghasilkan estimasi kemampuan siswa yang kurang tepat. Jika butir soal dengan bentuk jawaban MC terlalu sulit dapat memberi kemungkinan kepada siswa untuk menjawab pertanyaan hanya dengan cara menebak-nebak (*guessing*). Oleh sebab itu, tes dengan bentuk jawaban MC yang baik adalah tes yang memiliki tingkat kesulitan sedang.

Pengembangan tes dilakukan melalui proses yang cukup panjang. Tes yang baik mencakup pengetahuan, keterampilan, kemampuan dan karakteristik yang lain (KSAOs = *knowledge, skill, abilities and other characteristics*) yang dituntut untuk ditunjukkan. McIntire (2000) menetapkan 10 langkah pengembangan tes yang harus dilalui dalam proses pengembangan tes yaitu:

- 11) *Defining the test universe, audience, and purpose*
- 12) *Developing a test plan*
- 13) *Composing the test items*
- 14) *Writing the administration instructions*
- 15) *Conduct piloting test*
- 16) *Conduct item analysis*
- 17) *Revising the test*
- 18) *Validation the test*
- 19) *Developing norms*
- 20) *Complete test manual*

Puslitbang Sisjian (1996: 12) menerbitkan sebuah buku panduan untuk melengkapi kegiatan pengembangan tes dengan petunjuk analisis soal yang dapat dilihat pada Gambar 2. Soal-soal yang akan diujikan harus melewati telaah kualitatif dan kuantitatif untuk membuktikan validitas, reliabilitas dan analisis respon butir. Soal-soal yang sudah mengalami pengujian berkali-kali dan sudah memenuhi kriteria kualitas soal yang baik dapat dibakukan ke dalam bank soal.

Analisis kualitas butir tes secara kuantitatif dilakukan dengan menganalisis respon (jawaban) butir tes dari peserta tes. Substansi yang dianalisis meliputi tingkat kesukaran tes, daya pembeda tes, reliabilitas, dan pengukuran kesalahan baku (*standard error measurement*). Analisis berdasarkan teori respon butir dapat diterapkan hanya pada butir yang mempunyai karakteristik jawaban *local independent* atau *unidimensi* (Hambleton, 1991: 13).

Alat asesmen yang baik harus memenuhi syarat validitas dan reliabilitas. Alat asesmen dikatakan valid jika memberikan hasil ukur yang tepat dan akurat sesuai dengan maksud dikenakannya pengukuran tersebut. Sebuah tes hasil belajar dapat memenuhi persyaratan validitas isi, jika butir-butir soal yang dikembangkan mewakili seluruh materi yang diajarkan. Messick (1989) yang dikutip dari Kane (2006: 19) menganjurkan bahwa pembuktian validitas isi tidak dapat menggunakan data empiris, karena skor tes tidak menunjukkan kesesuaian isi tes dengan materi yang telah diajarkan. Analisis validitas isi banyak dilakukan terhadap perangkat instrumen untuk pengukuran prestasi belajar (*achievement test*). Validitas isi dapat dibuktikan oleh para ahli materi (*judgment experts*).

Validitas konstrak merujuk pada seberapa jauh suatu instrumen mengukur konstrak teori yang hendak diukur. Langkah yang ditempuh untuk mengukur validitas konstrak adalah menghubungkan konstrak (indikator) yang sedang diukur dengan konstrak lainnya menggunakan rumus korelasi. Validitas konstrak dikatakan konvergen apabila indikator ganda saling berasosiasi antara satu dengan yang lain (Neuman, 2003: 168).

Validitas kriteria terdiri dari validitas prediktif dan validitas konkuren (*concurrent*). Validitas prediksi diperoleh jika skor peserta tes saat ini memiliki korelasi yang tinggi terhadap skor tes pada pengukuran kemampuan berikutnya (*longitudinal*). Validitas *concurrent* diperoleh jika skor peserta tes yang diperoleh saat ini konsisten atau berkorelasi tinggi dengan skor tes yang sudah diperoleh sebelumnya. Konsep validitas kriteria diilustrasikan pada gambar 2.

Istilah reliabel dapat diartikan tetap atau konstan. Reliabilitas mengukur kemampuan instrumen untuk menghasilkan data yang mendekati sama bila instrumen tersebut digunakan berulang-ulang pada objek yang sama dan dengan cara yang sama. Pengujian reliabilitas suatu instrumen dapat dikerjakan secara internal dan eksternal. Pengujian reliabilitas secara internal

(*internal consistency*) berkaitan dengan analisis konsistensi butir-butir yang ada dalam instrumen dengan cara membagi satu set butir-butir pertanyaan menjadi dua bagian yang sama, bisa dengan cara membagi butir awal dan akhir atau ganjil dan genap. Tes hasil belajar dinyatakan reliabel apabila koefisien reliabilitasnya (α) > 7 (Nunnally: 1987)

Kompetensi Tata Hidang

Kompetensi merupakan sasaran evaluasi hasil belajar. Kompetensi diartikan sebagai 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). Mahasiswa dinyatakan kompeten dalam pekerjaan tertentu manakala ia memiliki perilaku (pengetahuan, keterampilan dan sikap kerja) minimum yang digunakan untuk menyelesaikan pekerjaan tersebut dalam bentuk unjuk kerja. Dalam pembelajaran, unjuk kerja merupakan penampilan peserta didik dalam mengerjakan sesuatu tugas yang terkait dengan pembelajaran.

Kompetensi yang ingin dicapai setelah mahasiswa belajar mata kuliah Tata Hidang adalah: (1) mampu mengajarkan cara melayani makan para tamu dalam

berbagai acara jamuan makan (formal, informal); (2) menerapkan etika makan jika dia menjadi tamu yang dilayani; (3) mengetahui berbagai macam menu, alat dan perlengkapan Tata Hidang. Kompetensi tersebut dicapai melalui beberapa kompetensi dasar sebagai berikut:

- 1) *Food and Beverage Service*
- 2) Pengetahuan Menu Tata Hidang
- 3) Perlengkapan Tata Hidang
- 4) Persiapan Area Restoran
- 5) Etika dalam Jamuan Makan
- 6) Sistem Pelayanan Makanan
- 7) Transaksi Keuangan

METODE PENELITIAN

H. Jenis Penelitian

Pengembangan alat asesmen Tata Hidang berbantuan komputer dilakukan dengan metode penelitian dan pengembangan (R&D). Tahap pengembangan mengacu pada model ADDIE (*analysis, design, develop, implementation, evaluation*) dari Dick and Carey: 1996). Kegiatan penelitian setiap tahap pengembangan secara rinci diuraikan pada paparan berikut ini:

Analysis

Pada tahap pertama penelitian dengan cara menganalisis materi dan indikator pencapaian kompetensi dasar pengetahuan Tata Hidang. Ruang lingkup materi Tata Hidang terdapat pada silabus dan bahan ajar.

Design

Pada tahap perancangan program disusun

- a. Menulis soal ujian Tata Hidang ranah kognitif dalam bentuk jawaban *multiple choice*. Validasi isi rancangan soal ujian melalui teknik Delphi dan FGD
- b. Menyusun rancangan program dan validasi rancangan program komputer untuk penyelenggaraan ujian on-line mata kuliah Tata-Hidang.

Develop

Menguji coba dan merevisi soal ujian Tata Hidang sampai diperoleh soal ujian yang berkualitas secara teoritis dan empiris sehingga layak masuk didokumentasi ke bank soal.

Implementation

Pada tahap ini, program komputer diaktifkan untuk penyelenggaraan ujian mata kuliah Tata Hidang. Program dinyatakan bagus dan valid jika semua komponen dapat dijalankan sesuai dengan fungsinya.

Evaluation

Evaluasi dilakukan untuk mengetahui kelayakan penyelenggaraan ujian menggunakan komputer. Evaluasi juga dilakukan setelah ujian untuk menguji validitas eksternal menggunakan *concurrent validity*. Bagan alir prosedur pengembangan alat asesmen hasil belajar Tata Hidang terdapat pada gambar 3.

I. Sumberdata Penelitian

Sumberdata penelitian setiap tahap berbeda-beda. Pada tahap analisis, sumberdata yang digunakan adalah dokumen silabus dan bahan ajar. Pada tahap validasi rancangan produk melibatkan berbagai macam tenaga ahli, diantaranya adalah ahli asesmen, ahli materi, ahli bahasa, programer komputer, dan desain grafis. Pengambilan data dilakukan dalam dengan teknik Delphi, dan FGD. Pada develop, sasaran sumberdatanya adalah mahasiswa Pendidikan Teknik Boga sebagai calon pengguna program

J. Metode Pengumpulan Data

Metode pengumpulan data yang digunakan dalam penelitian ini yaitu dokumentasi, teknik Delphi, FGD, tes dan lembar observasi. Metode dokumentasi digunakan pada saat analisis materi dari silabus dan bahan ajar. Teknik Delphi dan FGD (*Focus Group Discussion*) dipakai pada saat menguji kelayakan rancangan program komputer. Metode tes digunakan untuk memperoleh data respon butir peserta tes. Observasi dilakukan untuk mengamati respon peserta tes pada saat mengerjakan ujian dengan komputer.

K. Instrumen Penelitian

Instrumen penelitian terdiri dari alat asesmen (soal ujian) Tata Hidang dan lembar validasi soal ujian. Soal ujian Tata Hidang dikendalikan validitas dan

reliabilitasnya. Proses validasi instrumen dilakukan dengan beberapa cara yaitu validitas isi, butir, konstruk dan kriteria. Anastasi (2003: 86) menjelaskan bahwa pada dasarnya validitas isi adalah pengujian sistematis isi tes untuk menentukan apakah tes tersebut cukup representatif mengukur domain perilaku yang hendak diteliti. Setelah validitas isi terpenuhi, soal tes kemudian diukur berdasarkan validitas butir, konstruk dan kriterianya. Validitas butir dilakukan dengan cara mengkorelasikan skor butir dengan skor total. Validitas kriteria dilakukan untuk membandingkan skor tes yang menggunakan komputer dan skor tes manual menggunakan paper and pencil test.

Reliabilitas tes digunakan untuk mengetahui kestabilan alat ukur/tes ketika tes tersebut digunakan untuk mengukur hasil belajar mahasiswa. Reliabilitas memiliki dua pengertian yaitu konsistensi internal dan konsistensi eksternal. Konsistensi internal berkaitan dengan keandalan butir soal dalam seperangkat tes cukup konsisten untuk mengukur kemampuan siswa. Konsistensi eksternal bermakna pada kemampuan siswa yang cukup konsisten meskipun diukur dengan alat ukur dan waktu pengukuran yang berbeda. Pembuktian reliabilitas tes dalam penelitian ini menggunakan pendekatan

konsistensi internal karena soal hanya diujikan satu kali. Koefisien reliabilitas (Alpha) dihitung dengan rumus Cronbach's Alpha, yaitu:

$$\alpha = \frac{K}{(K-1)} \left(1 - \frac{\sum \sigma_i^2}{\sigma_x^2} \right)$$

Keterangan:

K = jumlah butir dalam tes

σ_i^2 = varian butir soal i

σ_x^2 = varian tes total

Analisis kualitas tes juga dilakukan dengan teori respon butir. Materi yang dianalisis antara lain tingkat kesukaran tes, daya pembeda tes, reliabilitas, dan pengukuran kesalahan baku (*standard error measurement*). Analisis teori respon butir hanya diterapkan pada butir yang mempunyai karakteristik jawaban *local independent* atau *unidimensi* (Hambleton, 1991: 13).

Indeks kesulitan (*difficulty*) masing-masing butir dihitung dengan menemukan persentase peserta tes yang menjawab butir dengan benar. Daya pembeda bekerja untuk dengan cara membandingkan kinerja masing-masing butir yang membuat seseorang memiliki skor sangat tinggi dengan seseorang yang memiliki skor sangat rendah (Puslitbang Sisjian, 1996). Soal yang bagus memiliki daerah

kesulitan pada tingkat rata-rata sekitar 0,5 dan indeks daya pembeda (*discrimination*). Kriteria soal/variabel dinyatakan reliabel apabila mempunyai koefisien Alpha > 0,70 (McIntire, 2000: 122).

L. Metode Analisis Data

Metode analisis data yang digunakan pada tahap analisis menggunakan analisis deskriptif kuantitatif dan kualitatif sedangkan pada akhir pengujian produk menggunakan analisis *Product Moment* untuk mengetahui validitas konkuren antara nilai peserta tes yang berasal dari komputer dengan nilai peserta tes yang diperoleh dari paper and pencil test. Rumus analisis korelasi *Product Moment* yang digunakan adalah sebagai berikut:

$$r_{xy} = \frac{n \sum X_i Y_i - (\sum X_i)(\sum Y_i)}{\sqrt{(n \sum X_i^2 - (\sum X_i)^2)(n \sum Y_i^2 - (\sum Y_i)^2)}}$$

HASIL DAN PEMBAHASAN

Hasil Analisis

Mata kuliah Tata Hidang terdiri dari 1 SKS teori dan 2 SKS praktik. Materi yang dapat disusun menjadi butir soal pilihan ganda adalah materi untuk pelajaran teori yang berisi banyak pengetahuan. Materi memiliki banyak poin-poin penting yang perlu diketahui atau dihafalkan sehingga memberi

kemungkinan untuk disusun pertanyaan-pertanyaan yang memiliki banyak alternatif jawaban pendek. Butir soal *multiple choice* (MC) hanya mengukur aspek pengetahuan (*knowledge*) saja. Soal MC tidak menuntut jawaban yang memberi solusi atas sebuah permasalahan, atau jawaban yang diuraikan. Hasil analisis materi yang dikembangkan menjadi butir soal mata kuliah Tata Hidang dapat disimak pada tabel

Tabel 1: Kisi-kisi Soal Ujian Tata

Hidang

No	Materi Dasar Kompetensi	Jumlah butir
1.	Ruang Lingkup Tata Hidang	10
2.	<i>Personalia Food and Beverage Service</i>	15
3.	Pengetahuan Menu Tata Hidang	15
4.	Perlengkapan Tata Hidang	15
5.	Persiapan Area Restoran	15
6.	Etika dalam Jamuan Makan	10
7.	Sistem Pelayanan Makanan	30
8.	Transaksi Keuangan	15
9.	Campuran dalam 8 Kompetensi Dasar	68
	Jumlah	193

Hasil Telaah Butir

Soal yang akan didokumentasi di dalam bank soal harus telah memiliki kualitas butir yang baik. Untuk menjamin

supaya soal yang masuk di bank soal telah memenuhi kriteria kualitas yang baik, maka soal harus melewati beberapa tahap penelaahan (*review*). Secara berturut-turut dilaporkan hasil telaah teoritik atau kualitatif, telaah empiris atau kuantitatif, validitas dan reliabilitas tes. Telaah butir kualitatif dilakukan dengan teknik Delphi. Tim ahli yang dilibatkan dalam penelaahan butir adalah tiga orang ahli yaitu ahli materi Tata Hidang (A), ahli pengukuran (B) dan ahli bahasa (C).

Telaah butir secara kualitatif dilakukan langsung pada lembar soal ujian. Hasil telaah butir dari ahli materi masih menemukan ada beberapa butir soal yang belum memiliki jawaban yang paling benar. Hal ini bermakna bahwa soal dapat memiliki lebih dari satu jawaban benar atau tidak ada sama sekali jawaban yang cocok untuk masalah yang ditanyakan. Selain itu, pilihan jawaban ada yang belum homogen dan logis ditinjau dari sisi materi. Ahli pengukuran menelaah konstruksi butir soal. Secara umum, butir soal dinyatakan sudah memenuhi kriteria. Hanya ada satu syarat yang belum memenuhi kriteria yaitu masih terdapat banyak rumusan soal dan rumusan jawaban tidak berisi pernyataan yang diperlukan saja. Beberapa butir soal masih menggunakan kalimat yang terlalu

panjang, sehingga perlu dipersingkat supaya lebih efektif

Hasil telaah butir dari ahli bahasa menemukan dua kriteria penyusunan butir soal yang masih perlu diperbaiki yaitu: (1) rumusan soal dan rumusan jawaban tidak hanya merupakan pernyataan yang diperlukan saja; (2) soal belum menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia. Kesalahan bahasa yang terjadi antara lain penulisan ejaan masih banyak yang salah, misalnya penulisan kata depan (di) dan kata asing yang berasal dari bahasa Inggris belum dicetak miring. Hasil telaah butir kualitatif kemudian ditindaklanjuti dengan merevisi butir soal sebelum soal tersebut digunakan untuk ujian.

Setelah melalui penelaahan (validasi) isi butir soal secara kualitatif, soal ujian Tata Hidang kemudian digunakan untuk ujian akhir semester. Soal yang diujicobakan sebanyak 100 butir. Respon jawaban ujian Tata Hidang kemudian dianalisis menggunakan program iteman untuk mengetahui kualitas butir tes dari aspek tingkat kesulitan butir, daya pembeda butir, pengecoh, validitas butir, dan reliabilitasnya. Butir soal yang telah memenuhi kriteria baik akan didokumentasi dalam bank soal. Sementara itu, butir soal yang belum memenuhi

kriteria baik akan ditinjau ulang dan direvisi. Rangkuman hasil dilaporkan pada tabel 2 berikut ini

Tabel 2: Indeks Tingkat Kesulitan Butir (p)

Kategori	Proporsi benar	Jumlah
Mudah	$p > 0,7$	5
Sedang	$0,3 \leq p \leq 0,7$	77
Sulit	$p < 0,3$	17
Salah		1

Kriteria soal yang berkualitas baik adalah jika memiliki indek tingkat kesulitan butir pada kategori sedang ($0,3 \leq p \leq 0,7$). Hasil telaah butir secara kuantitatif menunjukkan 77% butir soal ujian Tata Hidang berada pada tingkat kesulitan sedang. Hasil analisis tersebut menunjukkan kualitas butir sudah memenuhi kriteria soal yang baik. Butir soal yang mudah tetap digunakan untuk memberi kepuasan psikologis kepada mahasiswa yang kurang pandai agar mereka dapat menambah nilai ujian.

Butir soal tes yang baik harus dapat membedakan kemampuan peserta tes diantara peserta tes lainnya. Soal tes yang sulit hanya dapat dijawab benar oleh peserta tes yang pandai. Jika butir soal yang sulit bisa dijawab secara *guessing* oleh siswa kurang pandai maka butir soal tersebut tidak memiliki daya pemda yang

baik. Hasil analisis daya pembeda butir dapat disimak pada tabel ...

Tabel 3. Kriteria Indeks Daya Pembeda Butir (D/r_{pb})

Kategori	Indeks pembeda	Jumlah
Sangat baik	$D \geq 0,4$	48
Baik, tanpa revisi	$0,3 \leq D \leq 0,39$	30
Perbatasan atau perlu revisi	$0,2 \leq D \leq 0,29$	7
Dibuang atau diganti	$D \leq 0,19$	15
Jumlah		100

Hasil analisis menunjukkan terdapat 7% soal perlu direvisi dan 15% soal perlu diganti. Setelah dicermati lebih mendalam, soal yang harus direvisi antara lain karena jawaban ujian memiliki pengecoh dan banyak peserta tes yang terkecoh memilih jawaban tersebut. Hasil analisis ini ditindaklanjuti dengan mereview soal ujian butir demi butir. Soal yang sudah direvisi kemudian didokumentasi ke dalam bank soal dan disimpan ke dalam program komputer.

Pengecekan validitas soal ujian secara empiris dilakukan dengan validitas konstruk yaitu mengkorelasikan kumpulan skor butir pada setiap kompetensi dasar dengan skor total ujian Tata Hidang. Hasil analisis validitas konstruk dapat dilaporkan sebagai berikut:

Tabel 4. Rangkuman validitas Konstruk

No	Materi Dasar Kompetensi	r hit
1.	Ruang Lingkup Tata Hidang	
2.	<i>Personalia Food and Beverage Service</i>	
3.	Pengetahuan Menu Tata Hidang	
4.	Perlengkapan Tata Hidang	
5.	Persiapan Area Restoran	
6.	Etika dalam Jamuan Makan	
7.	Sistem Pelayanan Makanan	
8.	Transaksi Keuangan	

Menurut Hinkle (1979: 85), kriteria yang digunakan untuk menginterpretasikan koefisien korelasi adalah sebagai berikut:

Tabel 5. Interpretasi Hasil Analisis Korelasi Butir Tes

Range $r \pm$	Interpretasi	Jumlah
0,9 – 1	Korelasi sangat tinggi,	
0,7 – 0,9	Korelasi tinggi,	
0,5 – 0,7	Korelasi sedang,	
0,3 – 0,5	Korelasi rendah,	
0,0 – 0,3	Ada korelasi kecil	

Reliabilitas soal ujian dicek menggunakan rumus Alpha Cronbach. Hasil analisis reliabilitas sebesar 0,945. Kriteria soal yang baik adalah memiliki koefisien Alpha $>0,6$ (Nunally, 1978) oleh sebab itu soal telah memiliki kriteria reliabilitas yang ditetapkan.

Rancangan computerized test

Pengembangan alat asesmen pembelajaran Tata Hidang berbantuan

komputer memiliki dua jalur pengembangan yaitu pengembangan butir tes yang akan mengisi program dan pengembangan program komputer yang menjadi media penyelenggaraan tes berbantuan komputer (*computerized test*). Program ujian berbantuan komputer dikembangkan menggunakan *software Wondershare*.

Soal Tata Hidang tipe jawaban pilihan ganda yang telah memenuhi syarat untuk didokumentasikan dalam bank soal kemudian diisikan dan di-up load ke *software Wondershare*. Langkah-langkah yang dilakukan untuk mengisikankan soal ke dalam *software Wondershare* adalah sebagai berikut:

- i) Membuat *User ID field* dan *password field* beserta tombol untuk melanjutkan proses.
- j) Membuat *NIM field*, *First name field* dan *Last Name field* beserta tombol untuk melanjutkan proses.
- k) Memasukkan soal dan jawaban dengan bentuk pilihan ganda.
- l) Merekam suara *dubbing* untuk media pembacaan soal.
- m) Memasukkan gambar/media lain untuk media penjelas soal/jawaban.
- n) Mengidentifikasi dan mengelompokkan soal berdasarkan Kompetensi Dasar dan tingkat kesulitan.
- o) Menerbitkan soal ke dalam bentuk *Flash*.
- p) Instalasi *Flash Player* untuk memutar soal yang telah diterbitkan ke dalam bentuk *Flash*.

Soal ujian yang sudah dimasukkan ke dalam *software Wondershare* telah siap dioperasikan. Ujian Tata Hidang dengan *software Wondershare* diujicobakan pada tahun kedua. Fasilitas soal ujian Tata Hidang berbantuan komputer telah dilengkapi dengan *Audio Dubbing* pembacaan soal untuk soal pada urutan Kompetensi Dasar 1. Soal ujian Tata Hidang hanya bisa diakses oleh mahasiswa yang telah memiliki *User ID* dan *password*. Tampilan soal ujian Tata Hidang dapat disimak pada Gambar berikut ini:



Gambar 1. Tampilan awal untuk untuk masuk sebagai user

Tampilan pertama sebelum peserta tes/ujian memulai mengerjakan soal-soal Tata Hidang, mereka diminta mengisi user dan password terlebih dahulu. Soal ujian hanya bisa diakses oleh mahasiswa yang telah memasukkan identitas user dan

passwordnya. User dan password hanya dapat diperoleh dari administrator program. Langkah berikutnya dapat disimak pada gambar berikut ini



Gambar 2. Tampilan awal untuk untuk masuk sebagai peserta ujian

Mahasiswa yang akan mengikuti ujian Tata Hidang harus mengisi nama dan NIM untuk melanjutkan ke langkah berikutnya. Kompetensi dasar yang mau dikerjakan terlebih dahulu dapat dipilih oleh mahasiswa. Hasil ujian langsung dapat dilihat oleh peserta ujian setelah jawaban dikumpulkan (*Submit*) seperti tertera pada gambar berikut ini.



Gambar 3. Tampilan Hasil Ujian

Passing grade atau yang sering dikenal dengan nama skor potong (*cut score*) merupakan skor pembatas bagi peserta tes dapat dinyatakan lulus/tidak lulus ujian. Skor potong telah ditetapkan sebelumnya. Mahasiswa dapat melihat keputusan hasil ujian dengan membandingkan skor yang diperoleh dari skor potong yang telah ditetapkan. Jika posisi nilai berada di atas skor potong berarti peserta ujian dinyatakan lulus. Jika jawaban mahasiswa banyak yang salah, mahasiswa dapat mengoreksi kembali jawaban yang telah dipilih dengan kunci jawaban seperti terlihat pada gambar berikut ini:



Gambar 4. Tampilan Review Kunci Jawaban

Pada tampilan layar review tampak: (1) tombol *Author info* untuk melihat informasi pembuat soal; (2) tombol *Audio* untuk menghidupkan/mematikan efek suara dan pengaturan volume; dan (3) tombol *Print* untuk mencetak soal.

Ada beberapa macam *software* yang dapat membantu penayangan soal ujian menggunakan komputer. Dalam penelitian ini, program yang digunakan untuk menayangkan soal ujian Tata Hidang dengan komputer adalah *software Wondershare*. Software ini memiliki beberapa kelebihan yaitu: (1) tampilan menarik dan bisa dimodifikasi; (2) pengoperasian sederhana sehingga mudah dilakukan oleh siapa saja; (3) biaya maintenance murah; (4) mudah diakses; (5) kapasitas ruang yang diperlukan untuk penyimpanan cukup kecil yaitu hanya memerlukan 1 KB.

KESIMPULAN DAN SARAN

Kesimpulan

1. Ruang lingkup materi Tata Hidang yang layak dikembangkan menjadi kisi-kisi soal tes pilihan ganda adalah: (1) *food and beverage service*; (2) pengetahuan menu Tata Hidang; (3) perlengkapan Tata Hidang; (4) persiapan area restoran; (5) etika dalam jamuan makan; (6) sistem pelayanan makanan; dan (7) transaksi keuangan;
2. Telaah kualitas butir soal ujian Tata Hidang secara kualitatif telah ditindaklanjuti untuk memperbaiki butir soal yang masih kurang baik yaitu: (1) belum mempunyai satu jawaban yang paling benar; (2)

pilihan jawaban belum homogen dan logis ditinjau dari sisi materi; (3) rumusan soal dan rumusan jawaban tidak hanya merupakan pernyataan yang diperlukan saja; (4) soal menggunakan bahasa yang sesuai dengan kaidah bahasa Indonesia. Telaah kualitas butir soal tes secara kuantitatif telah membuktikan 77% soal berada pada tingkat kesulitan sedang, 48% butir soal memiliki daya pembeda sangat baik dan 30% soal memiliki daya pembeda yang baik. Validitas konstruk soal ujian Tata Hidang dan reliabilitas soal 0,945 telah memenuhi kriteria soal yang baik

3. Program komputer yang digunakan untuk ujian Tata Hidang secara *on line* atau *computerized test* adalah *software Wondershare*. Pada saat merancang program telah disusun langkah-langkah pembuatan untuk programmer dan langkah-langkah pengoperasian untuk pengguna. Program memiliki pengamanan sehingga hanya pengguna yang telah diberi password saja yang dapat mengakses soal. Setelah mengerjakan soal, pengguna dapat langsung melihat skor tes dan kunci jawaban jika diperlukan. Program diterbitkan dalam bentuk flash sehingga mudah diinstall di komputer.

DAFTAR PUSTAKA

- Burden, P. R., & Byrd, D. M. (1998). *Methods for effective teaching*. Boston: Allyn and Bacon.
- Clauser, B. E., Harik, P. & Clyman, S. G., (2000). The generalizability of scores for a performance assessment scored with a computer automated scoring system. *Journal of Educational Measurement*. Fall 2000, Vol. 37, No. 3, pp. 245-261
- Dick, W., & Carey, L. (1996). *The Systematic Design of Instruction* (4th Ed.). New York: Haper Collins College Publishers.
- Griffin, P., & Peter, N. (1991). *Educational assessment and reporting*. Sidney: Harcourt Brace Javanovich Publisher (Hambleton, 1991: 13).
- Huitt, W. (2001). Assessment, measurement and evaluation: Undergraduate version. *Educational Psychology Interactive*. Valdosta, GA: Valdosta State University. Retrieved (17 Nopember 2007), from <http://chiron.valdosta.edu/whuitt/edpsyc/edpmsevl.html>
- Kane, M. T. (2006). Validation. In Brennan, R. B. (Ed). *Educational measurement* (4th ed) pp 17 – 64. American Council on Education Praeger.
- McIntire, S. A., & Miller, L. A. (2000). *Foundation of psychological testing*. Boston: McGraw-Hill.
- Messick, S. (1989). Validity. In Robert L. Linn (Ed). *Educational measurement* (3rded) pp. 13-103. New York: American Council on Education and Macmillan
- Neuman, W. L. (2003). *Social research methods, qualitative and quantitative approaches* (5th). Boston: Pearson Education Inc
- Nunnaly, J. C. (1978). *Psychometric theory*, (2nd ed.). New Delhi: McGraw-Hill Publishing Company Limited
- Yahya Umar. (1996). *Bahan penataran pengujian pendidikan*. Jakarta: Puslitbangsisjian-Depdikbud

