

KUALITAS TES PILIHAN GANDA (*MULTIPLE-CHOICE*) SEBAGAI UPAYA MEMBENTUK PROSES BERFIKIR MAHASISWA

Emy Budiastuti

Jurusan PTBB FT UNY

emy.budiastuti@yahoo.com

Abstrak

Tes merupakan alat untuk mengetahui hasil belajar peserta didik yang utama digunakan. Berdasar hasil survey, di Jurusan PTBB FT UNY, masih jarang bahkan cenderung tidak pernah seorang dosen menganalisis tes yang dikembangkan. Padahal mengembangkan tes merupakan kegiatan yang mutlak dilakukan seorang dosen. Sehingga kualitas tes yang bisa membentuk proses berfikir mahasiswa masih dipertanyakan bahkan sulit diketahui. Tujuan penelitian ini untuk mengetahui kualitas tes ujian bentuk *Multiple-choice* di Jurusan PTBB FT UNY.

Kualitas tes ujian bentuk *Multiple choice* diketahui melalui analisis tingkat kesukaran, daya pembeda, efektifitas distraktor disamping validitas dan reliabilitas. Dengan demikian setiap tes yang dikembangkan dosen perlu dianalisis untuk mengetahui kualitas tes sehingga proses berfikir peserta didik yang sebenarnya dapat diketahui melalui tingkat kesukaran, daya pembeda, efektifitas distraktor, validitas dan reliabilitasnya

Hasil penelitian menunjukkan bahwa tes pilihan ganda yang dikembangkan dosen mempunyai kualitas rendah. Rendahnya kualitas tes diketahui melalui tingkat kesukaran (p) dan daya pembeda rendah ($< 0,25$), distraktor tidak berfungsi dengan baik, dan reliabilitas rendah (< 0.70)

Kata kunci: *kualitas tes, Multiple-choice*

PENDAHULUAN

Dalam pembelajaran, pendidik (dosen) mempunyai kewajiban untuk mengetahui seberapa banyak materi yang telah diajarkan mampu dikuasai peserta didik (mahasiswa). Untuk mengetahui seberapa banyak materi yang bisa dikuasai mahasiswa sebagai hasil pembelajaran diperlukan sarana ataupun cara yang dapat dipertanggung jawabkan secara ilmiah.

Salah satu cara yang digunakan adalah dengan melakukan pengukuran. Pengukuran diperlukan sebagai upaya untuk materi yang dikuasai dan yang belum dikuasai oleh mahasiswa. Alat ukur yang sering digunakan untuk mengetahui hasil belajar mahasiswa adalah instrumen tes.

Pengembangan instrumen tes bentuk tes pilihan ganda, merupakan salah satu perangkat pembelajaran yang perlu dilakukan seorang pengajar (dosen). Dalam mengembangkan tes (tes yang baik dan berkualitas) yang digunakan untuk mengukur hasil belajar diperlukan analisis butir soal. Analisis butir soal pilihan ganda merupakan

langkah penting yang harus dilakukan dosen setelah dosen melakukan evaluasi tes hasil belajar kepada mahasiswa khususnya pada matakuliah teori. Melalui analisis butir soal, akan dapat mengetahui kualitas butir soal yang dibuatnya, yaitu: 1) mengetahui taraf kesukaran soal; 2) menentukan daya pembeda soal (membedakan antara kelompok peserta didik berprestasi tinggi dan berprestasi rendah); 3) mengetahui pola penyebaran jawaban baik sebagai kunci jawaban maupun sebagai pengecoh (efektifitas pengecoh); 4) mengetahui validitas; dan 5) mengetahui reliabilitas tes.

Melalui analisis butir soal, dosen dapat mengetahui gambaran kemampuan peserta didik, sehingga dosen bisa melakukan perbaikan dalam proses pembelajaran selanjutnya. Dengan demikian jika dosen terus melakukan perbaikan khusus dalam pengembangan tes, maka kualitas soal yang dibuatnya semakin lama menjadi semakin sempurna.

Namun kenyataannya khusus di Prodi Pendidikan Teknik Busana, belum

semua butir soal tes terutama pada butir soal ujian dilakukan analisis butir soal. Melalui penelitian ini analisis butir soal ujian perlu dilakukan agar benar-benar diperoleh butir soal yang berkualitas, yang dapat mengukur kemampuan mahasiswa yang sebenarnya, sebagai upaya membentuk proses berfikir mahasiswa

Berdasar latar belakang masalah yang telah diuraikan, maka rumusan masalah dalam penelitian ini adalah

1. Bagaimanakah tingkat kesukaran butir soal ujian matakuliah teori Prodi Pendidikan Teknik Busana?
2. Bagaimanakah daya pembeda soal ujian matakuliah teori Prodi Pendidikan Teknik Busana?
3. Bagaimanakah keefektifan distraktor soal ujian matakuliah teori Prodi PT Busana?
4. Bagaimanakah validitas dan reliabilitas tes pada soal ujian matakuliah teori Prodi PT Busana?

Berdasar rumusan masalah, maka tujuan penelitian ini adalah sebagai berikut

1. Mengetahui tingkat kesukaran butir soal ujian matakuliah teori Prodi PT Busana
2. Mengetahui daya pembeda soal ujian matakuliah teori Prodi PT Busana
3. Mengetahui keefektifan distraktor butir soal ujian matakuliah teori Prodi PT Busana
4. Mengetahui validitas dan reliabilitas tes pada soal ujian matakuliah teori Prodi PT Busana

Manfaat yang diperoleh dalam penelitian ini adalah dosen mengetahui kualitas tes yang dikembangkan sebagai gambaran kemampuan mahasiswa sehingga dapat melakukan langkah-langkah untuk perbaikan pengembangan tes di waktu yang akan datang.

Tes merupakan sejumlah pertanyaan yang mempunyai jawaban benar dan salah. Tes juga diartikan sebagai sejumlah pertanyaan yang membutuhkan jawaban atau tanggapan dengan tujuan mengukur tingkat kemampuan seseorang atau mengungkap aspek tertentu dari orang yang dikenai tes (Djemari Mardapi, 2008:67). Hasil tes bisa digunakan untuk memantau perkembangan mutu pendidikan. Untuk itu maka diperlukan tes yang baik (berkualitas), yang memiliki kesalahan pengukuran yang sekecil mungkin.

Dalam pengukuran terdapat dua kesalahan yaitu kesalahan acak dan kesalahan sistemik. Kesalahan acak lebih mengarah pada pemilihan isi materi, variasi peserta tes, dan variasi pemeriksa lembar jawaban. Sedangkan kesalahan sistemik disebabkan karena butir soal tes terlalu sulit atau terlalu mudah.

Teknik tes merupakan salah satu bentuk penilaian kognitif. Teknik tes ada dua macam, yaitu tes obyektif dan non obyektif. Dikatakan tes obyektif karena siapapun yang melakukan penskoran akan menghasilkan skor yang sama. (Pujiati Suyata, 2011:1). Dipilih tes obyektif apabila pengajar bertujuan mengetahui keseluruhan kompetensi materi yang diajarkan dalam satu semester.

Penguasaan peserta didik dalam aspek pengetahuan tentang suatu pekerjaan merupakan faktor yang penting. Orlich et al. (2007:67-68) menyatakan, "*The cognitive domain encompasses objectives that deal with the recall or recognition of knowledge and the development of intellectual abilities and skills.*" Domain kognitif mencakup sasaran/hasil yang berhubungan dengan daya ingat atau pengenalan pengetahuan dan pengembangan kemampuan intelektual dan keterampilan. Metode yang lazim digunakan dalam mengukur aspek ini adalah tes tertulis dan lisan. Scott (1993:7) menjelaskan berbagai variasi yang mungkin dilakukan dalam mengembangkan tes tertulis, yaitu: *multiple-choice, sentence completion, listing, true-false, matching, essay, dan modified forms.*

Tes obyektif banyak digunakan pada matakuliah teori. Disamping mempunyai kelebihan, tes obyektif juga mempunyai kelemahan. Kelebihan tes obyektif adalah: 1) skor diberikan secara obyektif, 2) dapat mengungkap materi yang banyak, 3) menguji banyak testee, 4) pemeriksaannya mudah dan cepat. Sedangkan kelemahan tes obyektif yaitu cara penyusunannya sulit, karena harus memperhatikan pola penyebaran jawaban baik sebagai kunci jawaban maupun sebagai distraktor, serta mengetahui kondisi tiap butir soal terhadap skor keseluruhan.

Analisis butir soal dengan acuan adalah untuk mengetahui tingkat kesukaran soal, daya pembeda, dan indeks kehandalan

tes dan distribusi respon jika digunakan tes obyektif. Tingkat kesukaran dinyatakan dengan proporsi yang menjawab benar, besarnya mulai 0 sampai 1. Menurut acuan norma, tingkat kesukaran yang baik mulai 0,3 – 0,7. Daya pembeda dihitung dengan korelasi, besarnya mulai – 1,0 sampai 1,0. Indek daya pembeda yang baik besarnya mulai 0,20 atau lebih tinggi. Indek daya beda yang negatif sangat dihindari. Artinya, mereka yang mampu menjawab salah dan yang kurang mampu menjawab benar (Djemari, 1999:2)

Selain butir-butir tes yang baik, tes obyektif dituntut untuk dapat mengetahui kondisi terukur dengan baik. Oleh karena tes obyektif memerlukan persyaratan validitas dan reliabilitas. Kerlinger (1986) menyatakan bahwa validitas suatu tes merupakan kemampuan tes untuk mengukur apa yang seharusnya diukur. Sedangkan menurut Djemari Mardapi (2002), validitas menunjuk pada ketepatan, keberartian, dan kegunaan pemberlakuan yang spesifik dari hasil tes. Dengan demikian syarat suatu tes adalah memenuhi syarat valid, baik validitas isi, validitas konstruk, maupun validitas prediktif.

Validitas isi atau sering disebut validitas logik atau validitas teoritik, berkaitan dengan keseksamaan batasan pada domain yang akan diukur dan merupakan jawaban apakah keseluruhan butir merupakan sampel yang representative dari keseluruhan butir yang mungkin dibuat. Validitas isi dilakukan terhadap aspek materi, konstruksi, dan bahasa (Pujiati Suyata, 2011:4). Menurut Fernandez (1984), validitas konstruk menunjuk pada seberapa jauh suatu instrumen mengukur konstruk teoritik yang

akan diukur. Validitas prediktif merupakan suatu perangkat tes dalam membuat prediksi, dapat meramalkan keberhasilan *testee* pada masa yang akan datang. Berdasar pendapat Allen dan Yen (1979), validitas prediktif melibatkan pengguna skor tes untuk memprediksi perilaku *testee* di masa yang akan datang

Metode Penelitian

Penelitian ini dilaksanakan di jurusan PTBB FT UNY Prodi Pendidikan Teknik Busana. Penelitian ini merupakan penelitian deskriptif, yang bertujuan akan menjelaskan kualitas tes berdasar suatu analisis.

Untuk memperoleh data yang jelas, perlu ditentukan populasi dan sampel penelitiannya. Populasi sebagai sumber informasi dalam penelitian ini yaitu seluruh soal ujian di Prodi PT Busana, sedangkan sampel sebagai sumber informasi adalah soal ujian semester ganjil 2012 yang berbentuk tes Pilihan Ganda (*Multiple-choice Test*), yaitu Matakuliah Komputer. K3, dan Evaluasi pembelajaran. Teknik pengambilan sampel ditentukan secara *Purposive Sampling*.

4. Teknik Analisis Data

Teknik analisis data menggunakan analisis deskriptif berdasar analisis butir soal ujian matakuliah teori dari program ITEMAN.

Hasil Penelitian dan Pembahasan

Soal ujian Mata Kuliah yang telah dianalisis adalah: Komputer Disain, Kesehatan dan Keselamatan Kerja (K3), dan Evaluasi Pembelajaran. Hasil analisis butir soal berdasar Program ITEMAN disajikan dalam tabel berikut.

Tabel 1. Hasil Analisis Soal Ujian

Analisis	Komputer Disain	Kesehatan dan Keselamatan Kerja (K3)	Evaluasi Pembelajaran
N of Items	50	20	40
N of Examinees	33	36	33
Alpha	0,833	0.225	0.638

Mean P	0,742	0.179	0.273
Mean Biserial	0,498	0.422	0.453

Dari hasil ketiga mata kuliah tersebut bisa dinyatakan bahwa tingkat kesukaran soal menunjukkan indek yang masih rendah, artinya dari ketiga soal ujian yang dianalisis, tingkat kesukaran soal ujian < 0.30. Hasil menunjukkan bahwa soal ujian kedua matakuliah termasuk sukar dan perlu direvisi. Sedangkan tingkat kesukaran

untuk mata kuliah Komputer Disain termasuk mudah. Indek daya beda ketiga mata kuliah > 0.40. Indek reliabilitas tes yang memenuhi hanya mata kuliah Komputer disain.

1. Komputer Disain

Hasil analisis butir soal Mata Kuliah Komputer disain disajikan pada tabel 2 berikut

Tabel 2. Kualitas Soal Komputer Disain

No. Butir Soal	Tingkat Kesukaran	Daya Beda	Option (Alternatif Jawaban)	Keterangan Tingkat kesukaran butir
1	0.394	0.373	D belum berfungsi sebagai pengecoh	Baik
2	0.424	0.050	B belum berfungsi sebagai pengecoh	Baik
3	0.909	-0.026	D dan E tidak berfungsi	Sangat mudah
4	0.848	0.085	B dan E tidak berfungsi	Mudah
5	0.939	0.179	C dan D	Sangat mudah
6	0.879	0.063	A,C dan D tidak berfungsi	Mudah
7	0.194	0.181	D belum berfungsi sebagai pengecoh	Sangat sukar
8.	0.000	-9.000	B, D dan E tidak berfungsi	Sangat sukar
9.	0.970	0.084	C, D dan E tidak berfungsi	Sangat Mudah
10.	0.333	-0.151	C tidak berfungsi	Baik
11.	0.056	0.067	D dan E tidak berfungsi	Sangat sukar
12.	0.000	-9.000	D tidak berfungsi	Sangat sukar
13.	0.083	-0.050	Semua berfungsi sebagai pengecoh	Sangat sukar
14.	0.879	0.332	E tidak berfungsi	Mudah
15	0.778	0.134	B dan E tidak berfungsi	Mudah
16	0.061	-0.643	B dan E tidak berfungsi	Sangat sukar
17.	0.000	-9.000	D tidak berfungsi	Sangat sukar
18.	0.000	-9.000	D tidak berfungsi	Sangat sukar
19	0.788	0.145	B dan C tidak berfungsi	Mudah
20	0.636	0.407	B, D dan E tidak berfungsi	Sedang
21	0.576	0.567	E tidak berfungsi	Sedang
22	0.848	0.652	C dan E tidak berfungsi	Mudah
23	0.939	0.643	B, C dan E tidak berfungsi	Sangat mudah
24	0.455	0.354	A dan E tidak berfungsi	Sedang
25	0.758	0.689	B tidak berfungsi	Mudah
26	0.939	0.643	B, C dan E tidak berfungsi	Sangat Mudah
27	0.788	0.507	D dan E tidak berfungsi	Mudah
28	0.788	0.507	A dan B tidak berfungsi	Mudah
29	0.758	0.388	A dan E belum berfungsi	Mudah
30	0.818	0.236	B dan E tidak berfungsi	Mudah
31	0.788	0.631	Semua berfungsi sebagai pengecoh	Mudah
32	0.818	0.595	E tidak berfungsi	Mudah
33	0.667	0.591	B, C dan E tidak berfungsi	Sedang
34	0.576	0.390	D tidak berfungsi	Sedang
35	0.788	0.473	B tidak berfungsi	Mudah
36	0.758	0.560	B, C dan E tidak berfungsi	Mudah
37	0.758	0.657	D dan E tidak berfungsi	Mudah
38	0.879	0.658	A tidak berfungsi	Mudah

39	0.939	0.546	A dan E tidak berfungsi	Sangat mudah
40	0.788	0.676	B dan E tidak berfungsi	Mudah
41	0.939	0.527	B dan E tidak berfungsi	Sangat mudah
42	0.606	0.053	E tidak berfungsi	Sedang
43	0.727	0.208	Semua berfungsi sebagai pengecoh	Mudah
44	0.578	0.323	Semua berfungsi sebagai pengecoh	Sedang
45	0.909	0.279	B dan E tidak berfungsi	Sangat mudah
46	0.909	0.263	B dan E tidak berfungsi	Sangat mudah
47	0.970	0.515	D dan E tidak berfungsi	Sangat mudah
48	0.636	0.177	A dan D tidak berfungsi	Sedang
49	0.909	0.311	B dan D tidak berfungsi	Sangat mudah
50	0.606	0.289	B dan E tidak berfungsi	Sedang

Berdasar analisis data maka mata kuliah Komputer Disain termasuk soal ujian yang mudah. Analisis didasarkan pada butir-butir soal ternyata belum sepenuhnya menerapkan aspek-aspek yang ada dalam ranah kognitif. Butir soal masih mencerminkan tingkat berfikir yang rendah yaitu pengetahuan dan pemahaman, dan masih kurang menerapkan aspek analisis, sintesis dan aplikasi. Hal

demikian akan berpengaruh pada proses berfikir mahasiswa.

2. Kesehatan dan Keselamatan Kerja (K3)

Butir soal Mata Kuliah Kesehatan dan Keselamatan Kerja (K3) terdiri dari 20 butir soal, dengan 5 alternatif jawaban, yang diikuti oleh 36 mahasiswa

Tabel 3. Kualitas Soal Mata Kuliah K3

No. Butir Soal	Tingkat Kesukaran	Daya Beda	Option (Proporsi Jawaban)	Keterangan tingkat kesukaran butir
1.	0.191	0.181	A belum berfungsi	Sukar
2.	0.111	0.505	A dan C tidak berfungsi	Sangat sukar
3.	0.000	-9.000	B,C dan D tidak berfungsi	Sangat sukar
4.	0.028	0.158	A dan C tidak berfungsi	Sangat sukar
5.	0.056	0.067	A dan D tidak berfungsi	Sangat sukar
6.	0.000	-9.000	Semua berfungsi sebagai pengecoh	Sangat sukar
7.	0.000	-9.000	Semua berfungsi sebagai pengecoh	Sangat sukar
8.	0.333	-0.151	A dan E tidak berfungsi	Sedang
9.	0.333	0.583	A dan B tidak berfungsi	Sedang
10.	0.056	0.067	A dan E tidak berfungsi	Sangat sukar
11.	0.556	0.492	A tidak berfungsi	Sedang
12.	0.000	-9.000	C tidak berfungsi	Sangat sukar
13.	0.083	-0.050	B, C dan D tidak berfungsi	Sangat sukar
14.	0.472	0.517	A dan D tidak berfungsi	Sedang
15.	0.556	0.713	C dan D tidak berfungsi	Sedang
16.	0.000	-9.000	A dan E tidak berfungsi	Sangat sukar
17.	0.000	-9.000	A, B dan C tidak berfungsi	Sangat sukar
18.	0.000	-9.000	A, B dan C tidak berfungsi	Sangat sukar
19.	0.806	-9.000	B tidak berfungsi	Mudah
20.	0.000	-9.000	C tidak berfungsi	Sangat sukar

Berdasar hasil analisis, maka tingkat kesukaran butir soal mata kuliah K3 didominasi oleh butir soal yang sukar atau bahkan sangat sukar. Apabila dalam

menyusun soal diawali dengan menyusun kisi-kisi atau tabel spesifikasi sebagai langkah awal, maka akan diperoleh soal yang berkualitas. Pengecoh tidak efektif, artinya

option belum berperan sebagai pengecoh.
Jadi *option* tidak berfungsi secara maksimal.

3. Evaluasi Pembelajaran.

Analisis butir mata kuliah Evaluasi Pembelajaran disajikan pada tabel 4 berikut

Tabel 4. Hasil Analisis Soal Mata Kuliah Evaluasi Pembelajaran

No. Butir Soal	Tingkat Kesukaran	Daya Beda	Option (Proporsi Jawaban)	Keterangan Tingkat kesukaran Soal
1.	0.242	0.181	C tidak berfungsi	Sukar
2.	0.970	0.136	B dan C tidak berfungsi	Sangat mudah
3.	0.545	0,559	D tidak berfungsi	Baik
4.	0.424	0.533	D dan C tidak berfungsi	Baik
5.	0.273	0.270	B dan D tidak berfungsi	sukar
6.	0.091	0,009	Semua berfungsi sebagai pengecoh	Sangat sukar
7.	0,242	0,410	Semua berfungsi sebagai pengecoh	Sukar
8.	0.576	0,261	Semua berfungsi sebagai pengecoh	Baik
9.	0.576	0.201	D tidak berfungsi	Baik
10.	0.121	0.858	D tidak berfungsi	Sukar
11.	0.242	0.484	Semua berfungsi sebagai pengecoh	Sukar
12.	0,545	0,217	Semua berfungsi sebagai pengecoh	Baik
13.	0.121	0,819	Semua berfungsi sebagai pengecoh	Sukar
14.	0.091	1,000	D tidak berfungsi	Sangat sukar
15.	0.242	0.559		Sukar
16.	0.273	0,574	A dan E tidak berfungsi	Sukar
17.	0.152	0.142	A, B dan C belum berfungsi	Sangat sukar
18.	0.242	0,161	A, B dan C belum berfungsi	Sangat sukar
19.	0.061	0.590	B tidak berfungsi	Sangat sukar
20.	0.242	0,261	C tidak berfungsi	Sukar
21.	0.182	0.511	Berfungsi sebagai pengecoh	Sukar
22.	0.303	0,592	Berfungsi sebagai pengecoh	Baik
23.	0.030	1.000	B tidak berfungsi	Sangat sukar
24.	0.091	0.294	Berfungsi sebagai pengecoh	Sangat sukar
25.	0.030	1.000	Berfungsi sebagai pengecoh	Sangat Sukar
26.	0.333	-0.050	Berfungsi sebagai pengecoh	Baik
27.	0.030	1.000	Berfungsi sebagai pengecoh	Sangat sukar
28.	0.061	1.000	B tidak berfungsi	Sangat sukar
29.	0.485	0.058	B tidak berfungsi	Baik
30.	0.242	0.360	Berfungsi sebagai pengecoh	Sukar
31.	0.061	0.526	Berfungsi sebagai pengecoh	Sangat sukar
32.	0.091	0.674	Berfungsi sebagai pengecoh	Sangat sukar
33.	0.061	0.008	Berfungsi sebagai pengecoh	Sangat sukar

34	0.515	0.176	B tidak berfungsi	Baik
35	0.182	0.559	Berfungsi sebagai pengecoh	Sukar
36	0.424	0.354	Berfungsi sebagai pengecoh	Baik
37	0.848	0.553	Berfungsi sebagai pengecoh	Mudah
38	0.697	0.231	Berfungsi sebagai pengecoh	Baik
39	0.000	-9.000	C tidak berfungsi	Sangat sukar
40	0.000	-9000	B tidak berfungsi	Sangat sukar

Hasil analisis soal mata kuliah Evaluasi Pembelajaran menunjukkan bahwa tingkat kesukaran butir soalnya rendah. Tingkat kesukaran butir soal yang rendah menunjukkan bahwa butir soal tersebut sukar. Namun ditinjau dari daya pembeda soal, maka mata kuliah Evaluasi Pembelajaran termasuk baik, artinya butir soal tersebut dapat membedakan antara peserta berprestasi tinggi dengan yang berprestasi tinggi. Pengecoh berfungsi dengan baik, artinya hamper semua option dipilih oleh siswa.

KESIMPULAN

Berdasar hasil penelitian yang telah diuraikan, maka simpulan penelitian ini adalah sebagai berikut

1. Rerata Tingkat kesukaran soal ujian masih rendah yaitu < 0.300
2. Rerata Daya pembeda soal ujian > 0.4
3. Pengecoh belum berfungsi dengan baik
4. Validitas tampak sudah sesuai dengan karakteristik tes pilihan ganda
5. Indek Reliabilitas Soal ujian sebesar < 0.7

REKOMENDASI

1. Sebelum menyusun tes pilihan ganda dianjurkan untuk membuat kisi-kisi atau tabel spesifikasi untuk mempermudah dosen dalam menentukan sebaran butir soal sesuai dengan tingkatan domain kognitif
2. Agar Prodi mempunyai dokumen atau bank soal ujian yang berkualitas, masing-masing dosen diharapkan untuk meningkatkan diri dalam mengembangkan soal pilihan ganda

sesuai dengan kaidah yang dipersyaratkan

Daftar Pustaka

- Allen, M.J & Yen, W.M 1979. *Introduction to measurement theory*. Belmont, CA: Wadsworth, Inc
- Djemari Mardapi. 2008. *Teknik penyusunan instrument Tes dan Non Tes*. Jogjakarta: Mitra Cendekia
- Fernandez, H.J.X. 1984. *Testing and measurement*. Jakarta: National Education Planning, Evaluation and Curriculum Development
- Noordyah. 2012. *Menyusun soal pilihan ganda dan analisis butir*. file:///G:/MENYUSUN%20SOAL%20PILIHAN%20GANDA%20DAN%20ANALI . Diakses 30 Nopember 2012
- Orlich, D. C., et al. 2007. *Teaching strategies: A guide to effective instruction*. New York: Houghton Mifflin Company
- Pujiati Suyata. 2011. Validitas logis dan validitas empiris instrument penilaian dengan teknik tes. *Makalah*.
- Scott, J. L. 1993. *Improving vocational curriculum: Cognitif achievement evaluation*. Illinois: The Goodheart-Willcox Company, Inc
- Sumadi Suryabrata. 1987. *Pengembangan tes hasil belajar*. Jakarta: Rajawali