



## INSTRUMEN PENILAIAN *TWO-TIER TEST* ASPEK PENGETAHUAN UNTUK MENGUKUR KETERAMPILAN PROSES SAINS (KPS) PADA PEMBELAJARAN KIMIA UNTUK SISWA SMA/MA KELAS X

**Renata Kusuma Wardani<sup>1</sup>, Sri Yamtinah<sup>2\*</sup>, Bakti Mulyani<sup>2</sup>**

<sup>1</sup> Mahasiswa Program Studi Pendidikan Kimia, Fakultas Keguruan dan Ilmu Pendidikan,  
Universitas Sebelas Maret, Surakarta

<sup>2</sup> Dosen Program Studi Pendidikan Kimia, Fakultas Keguruan dan Ilmu Pendidikan,  
Universitas Sebelas Maret, Surakarta

\*Keperluan korespondensi, telp : 081227182520, email : jengtina\_sp@yahoo.com

### ABSTRAK

Penelitian ini bertujuan untuk : (1) mengembangkan instrumen penilaian *two-tier test* sebagai penilaian Keterampilan Proses Sains yang berbasis pendekatan saintifik, (2) mengetahui karakteristik instrumen penilaian *two-tier test* aspek pengetahuan yang memiliki syarat validitas, reliabilitas, tingkat kesukaran, daya beda, dan indeks pengecoh sebagai soal yang baik, (3) membuat profil siswa sebagai bentuk laporan yang informatif dan mudah dibuat oleh guru. Penelitian ini merupakan penelitian pengembangan, terdiri dari tahap (1) pemeriksaan pendahuluan, (2) penyesuaian teoritis, (3) uji empiris, (4) proses dan hasil dokumentasi, analisa dan refleksi. Subyek penelitian ini adalah guru kimia dan siswa kelas X. Berdasarkan penelitian yang dilakukan, dapat disimpulkan bahwa : (1) pengembangan *two-tier test* dilakukan berdasarkan model Akker, hingga dihasilkan produk berupa *two-tier test* berjumlah 20 butir soal sebagai instrumen penilaian Keterampilan Proses Sains (KPS), (2) instrumen penilaian *two-tier test* layak dan memenuhi kriteria soal yang baik, diperoleh harga validitas > 0,88; reliabilitas > 0,7; daya beda > 0,20; tingkat kesukaran untuk soal utama mencapai 40% mudah dan 60% sedang, dan tingkat kesukaran untuk soal alasan (*second-tier*) 55% mudah dan 45% sedang, dan indeks pengecoh telah berfungsi dengan baik, (3) profil siswa yang dibuat dapat digunakan sebagai sarana yang informatif untuk mengetahui pemahaman siswa dan keterampilan proses sains.

**Kata Kunci** : *two-tier test*, keterampilan proses sains, stoikiometri, profil siswa

### PENDAHULUAN

Pengetahuan kimia dipelajari pada tiga tingkatan, yaitu makroskopis, sub-mikroskopis, dan simbolis (*representational*). Hubungan antar tingkatan tersebut harus diajarkan secara eksplisit. Interaksi dan perbedaan di antara ketiga tingkatan tersebut merupakan karakteristik penting pada pembelajaran kimia dan hal ini diperlukan untuk memahami konsep-konsep kimia [1].

Kurikulum 2013 yang saat ini digunakan sangat menekankan pendekatan ilmiah pada setiap pembelajaran. Ilmu kimia merupakan ilmu yang didasarkan pada eksperimen,

sehingga keterampilan proses sains sangat diperlukan oleh siswa agar siswa dapat lebih memahami materi kimia secara lebih nyata. Keterampilan Proses Sains (KPS) adalah kemampuan siswa untuk menerapkan metode ilmiah dalam memahami, mengembangkan dan menemukan ilmu pengetahuan [2].

Salah satu hal yang sangat penting dalam pembelajaran adalah penilaian, dimana penilaian berfungsi untuk mengetahui sejauh mana pemahaman yang telah diperoleh siswa. Dalam melakukan penilaian, dibutuhkan alat ukur. Dimana alat ukur yang biasa digunakan guru adalah tes. Pada umumnya, bentuk tes yang selama ini

digunakan oleh guru adalah bentuk tes objektif dan uraian. Tes adalah cara atau prosedur dalam rangka pengukuran dan penilaian di bidang pendidikan, yang berbentuk pemberian tugas atau serangkaian tugas, baik berupa pertanyaan-pertanyaan, atau perintah-perintah oleh testee, sehingga dapat dihasilkan nilai yang melambangkan prestasi testee, dimana nilai tersebut kemudian dibandingkan dengan standar nilai tertentu [3].

Berdasarkan hasil observasi analisis kebutuhan menggunakan angket *preliminary study*, didapatkan hasil bahwa selama ini kebanyakan dari guru-guru, khususnya guru mata pelajaran kimia hanya mengenal instrumen penilaian dalam bentuk pilihan ganda dan uraian, dan beberapa dalam bentuk portofolio dan proyek. Namun terdapat kelemahan pada instrumen penilaian yang selama ini digunakan, misalnya pada instrumen penilaian bentuk pilihan ganda, siswa cenderung menebak jawaban dari *option jawaban* yang diberikan. Selain itu, tes bentuk pilihan ganda dan uraian belum mampu mengukur keterampilan proses sains yang seharusnya dimiliki oleh siswa. Sehingga diperlukannya alternatif instrumen penilaian lainnya yang dapat mengukur keterampilan proses sains. Dalam hal penilaian, kebiasaan mengerjakan tes yang sederhana dapat mempengaruhi tes yang memerlukan pemikiran kompleks jika tes dikaitkan dengan pengalaman belajar [4].

Dalam penelitian ini dikembangkan instrumen penilaian *two-tier test* sebagai alternatif instrumen yang dapat mengukur keterampilan proses sains. Instrumen *two-tier test* ini berupa tes obyektif 2 tingkat, soal utama (*first-tier*) berisi soal pilihan ganda biasa yang mengungkap aspek pengetahuan, sedangkan soal kedua (*second-tier*) berisi seperti layaknya pilihan ganda biasa namun menekankan pada pemikiran tingkat tinggi dan keterampilan dalam memberikan alasan [5].

Penskoran soal *two-tier test* ini dilakukan menggunakan penskoran

*Graded Response Model* (GRM), dimana *Graded Response Model* (GRM) merupakan salah satu model yang dikembangkan untuk menangani skoring pada butir-butir soal politomus [6].

Instrumen penilaian *two-tier test* yang dikembangkan juga dilengkapi dengan software analisis data yang dapat menghasilkan profil individu siswa, sehingga akan menghasilkan laporan yang informatif berkaitan dengan Keterampilan Proses Sains (KPS), letak kesulitan siswa pada materi yang bersangkutan, dan prestasi belajar siswa.

## METODE PENELITIAN

Penelitian dilaksanakan di 3 sekolah, antara lain SMA Negeri 1 Surakarta, SMA Negeri 5 Surakarta, dan SMA Batik 2 Surakarta. Subyek penelitian ini adalah siswa kelas X dan materi yang digunakan adalah Stoikiometri.

Metode penelitian yang digunakan adalah penelitian pengembangan, yang terdiri dari tahapan : pemeriksaan pendahuluan (*preliminary investigation*), penyesuaian teoritis (*theoretical embedding*), uji empiris (*empirical testing*), dan Proses dan hasil dokumentasi, analisa dan refleksi (*documentation, analysis, and reflection on process and outcome*). Pemeriksaan pendahuluan dilakukan dengan melakukan analisis kebutuhan (penyebaran angket) [7]. Pada tahap penyesuaian teoritis dilakukan kajian berkaitan dengan penskoran butir soal, dimana penskoran menggunakan *Graded Response Model* (GRM).

Tahap uji empiris terbagi menjadi 2 tahap, yaitu validasi oleh ahli dan uji coba soal (terbatas, menengah, dan luas). Validasi isi didasarkan pada validitas Aiken, dimana panelis berjumlah 8 orang yaitu 6 guru dan 2 dosen. Pada penelitian ini validitas isi ditentukan dengan menggunakan formula Aiken, yaitu:

$$S = \sum n_i (r - l_0)$$

Dimana:

- V : indeks validitas dari Aiken  
 c : banyaknya kategori/criteria  
 $l_0$  : kategori terendah  
 $n_i$  : banyaknya penilai (*raters*) yang memilih kriteria i  
 r : kriteria ke i  
 n : jumlah seluruh penilai

Nilai V berkisar pada 0 – 1 dan kriteria yang digunakan untuk menyatakan sebuah butir soal dikatakan valid secara isi pada jumlah rater (penilai) sebanyak 8 orang adalah 0,88 [8].

Tahap akhir adalah proses dan hasil dokumentasi, analisa dan refleksi, dimana pada tahap ini dilakukan analisa keefektifan produk yang dikembangkan dengan memperhatikan saran-saran dari guru melalui penyebaran angket kepuasan pengguna.

## HASIL DAN PEMBAHASAN

### Hasil Studi Pendahuluan

Tahap awal dari penelitian ini adalah pemeriksaan pendahuluan atau studi pendahuluan. Studi pendahuluan dilakukan dengan wawancara terhadap guru, menyebarkan angket analisis kebutuhan guru, dan angket analisis kebutuhan siswa.

Berdasarkan hasil wawancara, dapat disimpulkan bahwa sebagian besar guru menggunakan instrumen penilaian bentuk uraian dan pilihan ganda, dimana terdapat beberapa kelemahan pada soal bentuk tersebut, yaitu salah satunya soal bentuk pilihan ganda masih kurang efektif dalam mengukur KPS, sedangkan soal bentuk uraian membutuhkan waktu yang lama dalam pengoreksiannya. Oleh karenanya diperlukan alternatif bentuk soal yang mampu mengukur KPS secara efektif dan efisien.

Sedangkan, berdasarkan hasil angket analisis kebutuhan oleh guru, dapat disimpulkan bahwa kebanyakan guru menggunakan tes dalam evaluasi hasil belajar siswa, dimana bentuk tes yang digunakan adalah uraian dan pilihan ganda. Selain itu, salah satu kemampuan yang sangat perlu untuk dimiliki siswa adalah keterampilan

proses sains, sehingga dibutuhkan instrumen yang mampu mengukur KPS tersebut secara efisien. Salah satu materi yang dianggap cukup sulit adalah materi Stoikiometri, dimana pada materi ini hanya sebagian siswa yang tuntas

Berdasarkan hasil angket analisis kebutuhan oleh siswa dapat disimpulkan bahwa soal yang biasa digunakan oleh guru dalam melakukan penilaian adalah dalam bentuk uraian dan pilihan ganda, dimana soal-soal yang diberikan biasanya memiliki tingkat kesukaran yang sedang. Selain itu, sebagian besar siswa berpendapat bahwa diperlukannya bentuk instrumen yang mampu mengukur keterampilan proses sains secara efisien, sehingga mampu mengetahui sejauh mana pemahaman siswa akan materi yang telah disampaikan oleh guru.

Berdasarkan dari hasil studi pendahuluan, maka peneliti mengusulkan dikembangkannya instrumen alternatif bentuk lain yang mampu mengatasi kelemahan pada bentuk instrumen yang biasa digunakan oleh guru, yaitu instrumen penilaian *two-tier test* (berjumlah 20 butir soal dengan waktu 90 menit) yang mampu mengukur KPS secara efektif dan dilengkapi dengan software analisis data dengan output berupa profil individu siswa.

### Analisis Kualitas Soal

Analisis kualitas soal diperlukan untuk mengetahui kualitas soal yang dikembangkan. Analisis kualitas soal dilakukan berdasarkan pada hasil uji coba soal. Hasil uji coba soal dapat memberikan informasi atau data empirik tentang kebaikan soal yang dikembangkan [9].

Berdasarkan uji validitas yang telah dilakukan, dengan jumlah penilai (*raters*) 8 orang, soal *two-tier test* yang dikembangkan berjumlah 20 butir soal, terdapat 17 butir soal memiliki harga V minimal 0,88 sehingga 17 butir soal tersebut dapat dikatakan valid. Sedangkan 3 butir soal lainnya (butir soal nomor 9, 16, dan 17) memiliki harga V sebesar 0,75 yang menunjukkan 3 butir soal tersebut dikatakan tidak valid, sehingga

dilakukan perbaikan terhadap 3 butir soal tersebut.

### Analisis Hasil Uji Coba Produk

Uji coba produk dilakukan sebanyak 3 tahap, yaitu uji coba skala terbatas, uji coba skala menengah, dan uji coba skala luas.

Uji coba skala terbatas dilakukan untuk mengetahui keterbacaan soal dan kecukupan waktu. Dimana uji coba skala terbatas dilakukan dengan mengujicobakan soal ke lima siswa pada masing-masing sekolah.

Berdasarkan uji coba dapat disimpulkan bahwa hampir seluruh siswa peserta uji coba skala terbatas belum pernah mengerjakan soal model *two-tier test*. Namun dilihat dari soal yang diberikan, sebagian besar siswa menyatakan bahwa soal-soal yang diberikan cukup dapat dipahami, hanya pada beberapa soal terdapat data yang kurang lengkap dan terdapat beberapa soal yang tidak memiliki jawaban pada option yang diberikan, sehingga diperlukan beberapa perbaikan pada soal-soal tersebut. Dilihat dari tingkat kesukaran soal, terdapat beberapa soal yang tergolong sukar sehingga dalam pengerjaannya membutuhkan waktu yang lebih lama dibandingkan dengan soal yang lain. Sedangkan dari segi waktu yang diberikan, yaitu 90 menit dirasa cukup untuk mengerjakan 20 butir soal *two-tier test* yang diberikan.

Adapun dalam pelaksanaannya, penskoran yang digunakan adalah dengan *Graded Response Models* (GRM). Pedoman penskoran tersebut dapat dilihat pada Tabel 1.

Tabel 1. Pedoman Penskoran *Two-Tier Test*

Respon Jawaban Siswa		Skor
<i>First-tier</i>	<i>Second-tier</i>	
Benar	Benar	3
Benar	Salah	2
Salah	Benar	1
Salah/ Tidak	Salah/ Tidak	0
Menjawab	Menjawab	

Tahap selanjutnya adalah uji coba skala menengah. Dimana uji coba skala menengah dilakukan dengan mengujicobakan soal ke satu kelas pada masing-masing sekolah. Tujuan dari dilaksanakannya uji coba ini adalah untuk mengetahui karakteristik soal *two-tier* yang dikembangkan.

Berdasarkan pengujian kualitas soal dilihat dari reliabilitas, tingkat kesukaran, daya beda soal, dan indeks pengecoh, dapat disimpulkan bahwa soal *two-tier test* yang dikembangkan memenuhi persyaratan sebagai soal yang baik. Dimana dari perhitungan, diperoleh reliabilitas sebesar 0,875 untuk soal utama (*first-tier*) dan 0,886 untuk soal alasan (*second-tier*). Sedangkan untuk daya beda soal, butir soal yang dikembangkan memiliki daya beda antara 0,2-0,9, serta untuk indeks pengecoh diperoleh hasil bahwa distraktor soal dapat dikatakan berfungsi dengan baik karena telah dipilih oleh minimal 2,5% peserta. Adapun untuk tingkat kesukaran soal dapat dilihat pada Tabel 2.

Tabel 2. Persentase Tingkat Kesukaran Soal

Tingkat Kesukaran	Persentase Tingkat Kesukaran (%)		
	Mudah	Sedang	Sukar
<i>First-tier</i>	60	35	5
<i>Second-tier</i>	60	20	20

Berdasarkan hasil uji coba skala menengah ini, dapat dikatakan bahwa soal *two-tier test* yang dikembangkan cocok untuk digunakan sebagai instrumen penilaian pada sekolah dengan kategori sedang dan tinggi.

Tahap selanjutnya adalah uji coba skala luas. Dimana uji coba skala luas dilakukan dengan mengujicobakan soal ke dua kelas pada masing-masing sekolah. Tujuan dari dilakukannya uji coba skala luas ini adalah untuk memperoleh tanggapan berkaitan dengan soal *two-tier* yang dikembangkan.

Pada tahap ini diberikan angket tanggapan kepada guru berkaitan dengan soal yang dikembangkan, khususnya soal *two-tier* untuk materi

Stoikiometri. Adapun hasil dari tanggapan guru berkaitan dengan soal adalah bahwa soal *two-tier* merupakan soal yang baik yang dapat digunakan untuk mengukur KPS. Hanya saja penggunaannya yang masih sangat jarang, sehingga diperlukan sosialisasi berkaitan dengan soal dengan model *two-tier*. Selain itu, dilihat dari segi waktu yang diberikan untuk mengerjakan, yaitu 90 menit untuk 20 butir soal dirasa cukup, namun karena siswa atau siswa belum terbiasa dengan soal model *two-tier* ini, maka untuk siswa yang tergolong ke dalam siswa kelompok bawah, membutuhkan waktu yang lebih lama. Sehingga diperlukan pertimbangan kembali berkaitan dengan waktu yang diberikan dan tingkat kesukaran soal.

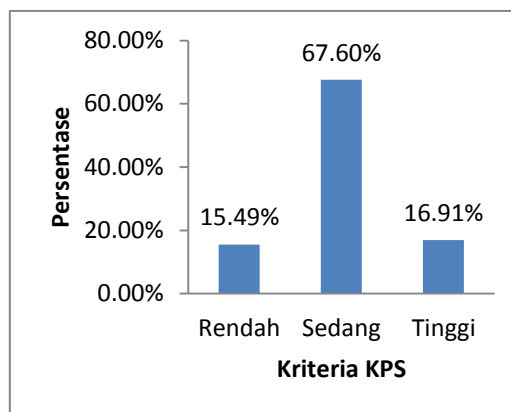
**Analisis Kemampuan Proses Sains**

Soal *two-tier test* yang dikembangkan selain berfungsi sebagai alat evaluasi, juga berfungsi sebagai alat penilaian untuk mengukur keterampilan proses sains siswa. Dimana soal *two-tier test* dikembangkan berdasarkan indikator keterampilan proses sains. Skor siswa yang dihasilkan kemudian dibagi berdasarkan jenjang keterampilan proses sains. Kategori keterampilan proses sains dapat dibagi menjadi 3 jenjang berdasarkan skornya. Pembagian kategori keterampilan proses sains berdasarkan skor dapat dilihat pada Tabel 3.

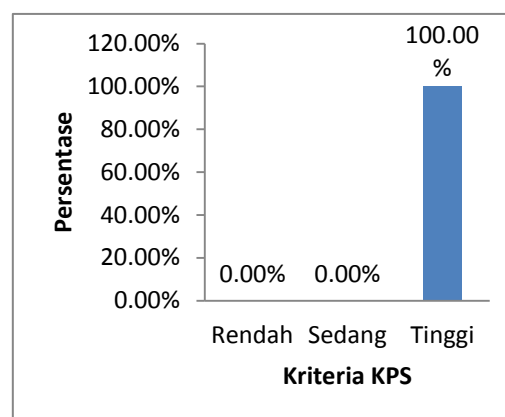
Tabel 3. Skala Kategori Keterampilan Proses Sains (KPS)

Skor (x)	Kategori KPS
$0 < x \leq 20$	Rendah
$21 \leq x \leq 40$	Sedang
$41 \leq x \leq 60$	Tinggi

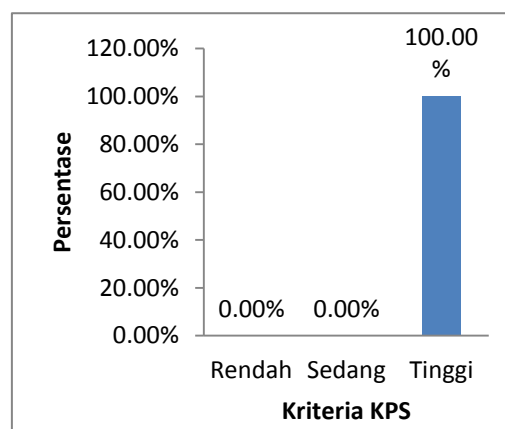
Berdasarkan skala kategori keterampilan proses sains pada Tabel 5, maka hasil pengelompokan siswa berdasarkan kategori keterampilan proses sains dapat dilihat pada Gambar 1, Gambar 2, dan Gambar 3.



Gambar 1. Grafik Keterampilan Proses Sains SMA Batik 2 Surakarta



Gambar 2. Grafik Keterampilan Proses Sains SMA N 5 Surakarta



Gambar 3. Grafik Keterampilan Proses Sains SMA 1 Surakarta

**Profil Siswa**

Profil siswa berfungsi sebagai sarana informatif bagi guru, siswa, dan orang tua untuk mengetahui sejauh mana pemahaman yang diperoleh siswa pada materi yang telah disampaikan.

Selain, memberikan informasi mengenai Keterampilan Proses Sains (KPS) yang dimiliki oleh siswa, profil siswa ini juga dapat digunakan untuk mengetahui prasetasi belajar siswa, serta letak kesulitan siswa pada materi stoikiometri, sehingga guru dapat melakukan usaha lebih lanjut sebagai upaya perbaikan dalam pembelajaran.

Profil siswa yang dikembangkan adalah dalam bentuk program Microsoft Excel. Hasil atau output dari program ini berupa lembar profil siswa, dimana dalam profil individu tersebut telah berisi indikator-indikator mana yang telah dan belum dicapai oleh siswa, prestasi belajar yang diperoleh oleh siswa, beserta kategori Keterampilan Proses Sains (KPS) siswa yang dicapai.

Berdasarkan hasil dari penyebaran angket kepada guru-guru, dapat disimpulkan bahwa profil individu siswa yang dikembangkan cukup mudah dipahami dan digunakan, baik dalam penggunaannya maupun hasil yang diperoleh. Selain itu profil individu siswa tersebut cukup efektif dan efisien dalam memberikan informasi kepada guru, siswa maupun orang tua mengenai sejauh mana pemahaman siswa, hanya saja masih terbatas pada materi tertentu.

## KESIMPULAN

Berdasarkan penelitian yang telah dilakukan, maka dapat disimpulkan sebagai berikut :

1. Pengembangan instrumen penilaian *two-tier test* dilakukan dengan beberapa tahap berdasarkan model Akker, meliputi : uji pendahuluan, penyusunan kisi-kisi dan butir soal, serta uji empiris, hingga dihasilkan produk akhir berupa instrumen *two-tier test* berjumlah 20 butir soal sebagai instrumen penilaian Keterampilan Proses Sains (KPS).
2. Instrumen penilaian *two-tier test* memenuhi kriteria sebagai soal yang baik, dimana harga validitas isi > 0,88; reliabilitas > 0,7; daya beda > 0,20; tingkat kesukaran untuk soal utama (*first-tier*) mencapai 40% mudah dan 60% sedang, dan tingkat kesukaran untuk soal alasan

(*second-tier*) mencapai 55% mudah dan 45% sedang, serta indeks pengecoh telah berfungsi dengan baik.

3. Profil siswa dikembangkan menggunakan program Microsoft Excel dengan beberapa bagian data, sehingga hasilnya layak digunakan sebagai sarana informatif untuk mengetahui keterampilan proses sains siswa.

## UCAPAN TERIMA KASIH

Selama proses penelitian ini, banyak pihak yang turut membantu secara langsung ataupun tidak langsung, memberikan dukungan, saran dan kritik yang membangun. Oleh karena itu, pada kesempatan ini dengan segala ketulusan hati, penulis mengucapkan terimakasih kepada Ibu Dra. Rahayu Sukantari., M.Pd. selaku guru Kimia SMA N 1 Surakarta, Bapak Drs. Ari Harnanto., M.Si dan Bapak Wayudi Padmono., S.Pd selaku guru Kimia SMA N 5 Surakarta, dan Bapak Ispriyanto., S.Pd selaku guru Kimia SMA Batik 2 Surakarta, serta semua pihak yang telah memberikan bantuan dan pengarahan kepada penulis

## DAFTAR RUJUKAN

- [1] Barke, H.D, Al Hazari & Yitbarek S., (2009), *Misconceptions in Chemistry, Addressing Perception in Chemical Education*, German : *Springer-Verlag Berlin Heidelberg*
- [2] Dahar, R.W, (2012), *Teori-Teori Belajar*, Jakarta : Depdikbud Dirjend Lembaga Tenaga Kependidikan
- [3] Sudijono, A., (2008), *Pengantar Evaluasi Pendidikan*. Jakarta: PT Raja Grafindo Persada
- [4] Subali, B., (2011), *Pengukuran Kreativitas Keterampilan Proses Sains dalam Konteks Assesment for Learning*. Jurnal Cakrawala Pendidikan Th XXX No.1
- [5] Adodo, S. O., (2013), *Effects of Two-Tier Multiple Choice Diagnostic Assessment items on Students'*

Learning Outcome in Basic Science Technology. Ondo State: Academic Journal of Interdisciplinary Studies University of Rome, 2 (2), 2281-3993

- [6] De Ayala, R. J. , (1993), An Introduction to Polytomous Item Response Theory Models. Measurement and Evaluation in Counseling and Development, 25, 127-189
- [7] Akker, J, V, (1999), Principles and Method of Development Research. University of Twentee
- [8] Aiken, L. (1985). Three Coefficient for Analyzing The Realibility and Validity of Ratings. Educational and Psychological Measuremen, 45
- [9] Mardapi, D., (2012), Pengukuran Penilaian dan Evaluasi Pendidikan. Yogyakarta : Nuha Litera