



Buku yang dihadapan pembaca berjudul **"Pengembangan Dan Pengujian Instrumen Untuk Penelitian Pendidikan"**. Instrumen merupakan bagian terpenting dalam kegiatan penelitian. Pengalaman selama 13 tahun sebagai dosen di perguruan tinggi, banyak ditemukan mahasiswa yang sulit menyelesaikan penelitiannya. Hal ini disebabkan salah satu faktornya adalah kurangnya pengetahuan dalam merancang instrumen penelitian. Kejadian tersebut terjadi pada mahasiswa yang mengambil S1 bahkan program pascasarjana.

Materi buku ini berusaha merespon kesulitan yang dihadapi oleh mahasiswa dan peneliti pada umumnya. Buku ini disajikan dalam bentuk yang sangat praktis berisi berbagai pijakan teori dan praktek tentang penyusunan instrumen, baik instrumen yang berbentuk angket maupun tes. Selanjutnya disajikan pula aplikasi pengolahan data tentang uji validitas dan reliabilitas instrumen penelitian, baik menggunakan kalkulator, Excel, Rash Model, Anates, dan SPSS.



Media Ilmu Press

Penerbit : Media Ilmu Press
Jl. Conge Ngembalrejo Bae Kudus.
No. Hp 085 641 367 301
<http://mediailmupress.blogspot.com>
Email : intannajma@gmail.com

ISBN: 978-979-16369-8-8



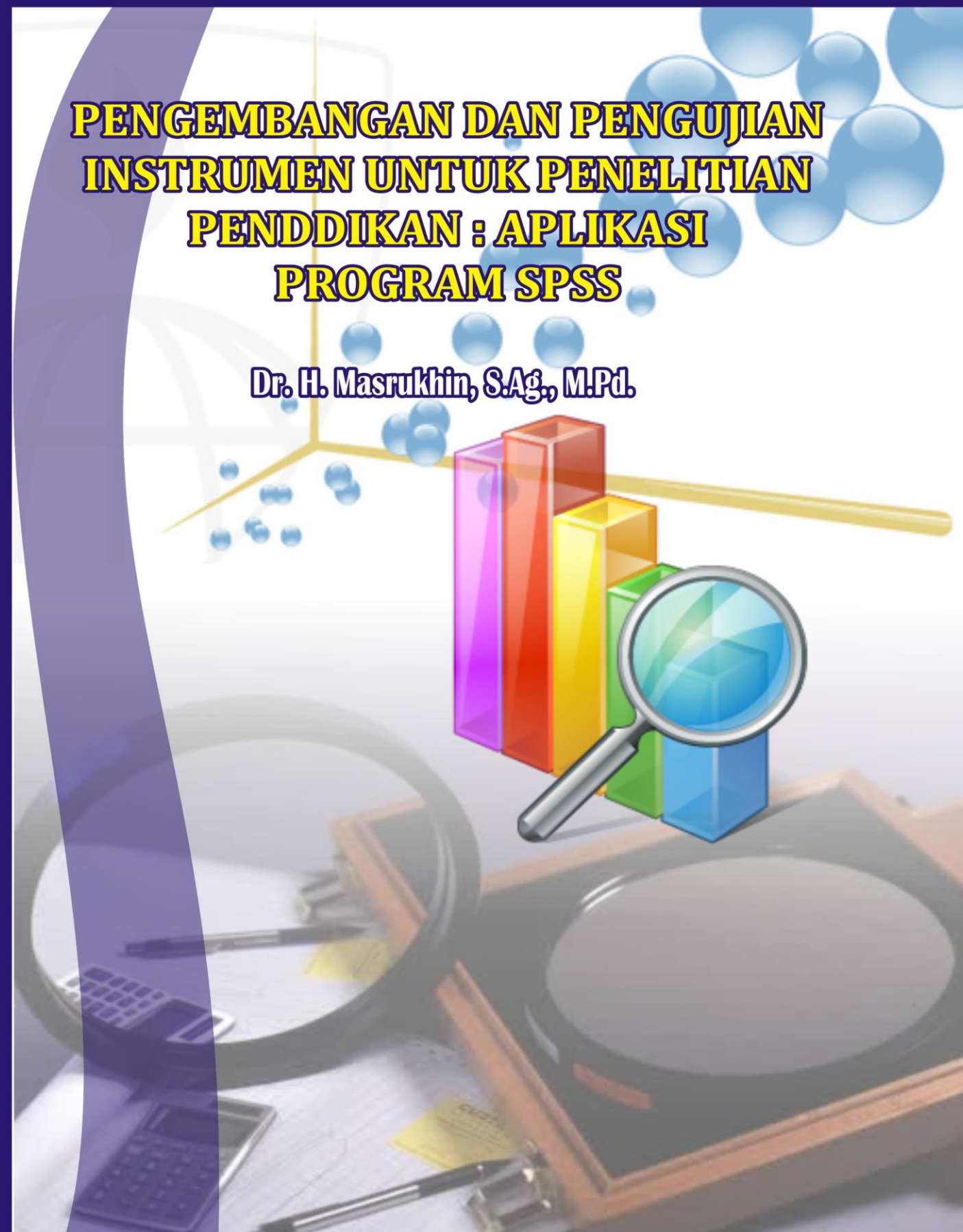
9789791636988-0

PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN PENDIDIKAN : APLIKASI PROGRAM SPSS

Dr. H. Masrukhin, S.Ag., M.Pd.

PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN PENDIDIKAN : APLIKASI PROGRAM SPSS

Dr. H. Masrukhin, S.Ag., M.Pd.



**PENGEMBANGAN
DAN PENGUJIAN INSTRUMEN
UNTUK PENELITIAN PENDIDIKAN :
*Aplikasi Program Excel, Rasch Model, Anates, SPSS,
dan WarpPLS.***

Dr. H. Masrukhin, S.Ag., M.Pd.



2016

MEDIA ILMU PRESS

Jl. Conge Ngembalrejo Depan Kampus
STAIN Kudus Hp. 085 641 367 301

Judul Buku :
**PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK EVALUASI PENELITIAN
PENDIDIKAN**

(Aplikasi Program SPSS, Anates, Rasch Model, dan WarpPLS)

ISBN : 978-979-16369-8-8

Penulis : Dr. H. MASRUKHIN, S.Ag., M.Pd.

Pencetak : UD. MEDIA ILMU PRESS

Cetakan

Hak Cipta @ 2016 pada penulis

Setting : Tim Media Ilmu

Desain Cover : Tim Media Ilmu

Hak Cipta dilindungi Undang-Undang.

Dilarang memperbanyak atau memindahkan sebagian atau seluruh isi buku ini dalam bentuk apapun, baik secara elektronik maupun mekanis, termasuk memfotocopy, merekam atau dengan penyimpangan lainnya, tanpa izin tertulis dari Penulis.

Penerbit: Media Ilmu Press
Jl. Conge Ngembalrejo Bae Kudus.
Hp. 085 641 367 301
<http://mediailmupress.blogspot.com>
Email : intannajma@gmail.com

KATA PENGANTAR

Segala puja dan puji penulis panjatkan ke hadirat Allah SWT, yang telah melimpahkan petunjuk, bimbingan dan kekuatan lahir dan batin kepada diri penulis, sehingga buku ini dapat tersusun dan terbit sebagaimana mestinya. Shalawat dan salam semoga dilimpahkan oleh-Nya kepada junjungan nabi besar Muhammad SAW, para sahabat dan semua pengikutnya yang setia di sepanjang zaman. Amien !

Instrumen merupakan bagian terpenting dalam kegiatan penelitian. Pengalaman selama 13 tahun sebagai dosen di perguruan tinggi, banyak ditemukan mahasiswa yang sulit menyelesaikan penelitiannya. Hal yang disebabkan salah satu faktornya adalah kurangnya pengetahuan dalam merancang instrumen penelitian. Kejadian tersebut terjadi pada mahasiswa yang mengambil S1 bahkan program pascasarjana.

Berdasarkan pengalaman tersebut, maka materi buku ini berusaha merespon kesulitan yang dihadapi oleh mahasiswa, yang diberi Judul “Pengembangan Dan Pengujian Instrumen Untuk Penelitian Pendidikan (Aplikasi Program SPSS)”. Buku ini berisi berbagai pijakan teori dan praktek tentang penyusunan instrumen, baik instrumen yang berbentuk angket maupun tes. Selanjutnya disajikan pula aplikasi pengolahan data tentang uji validitas dan reliabilitas instrumen penelitian, baik menggunakan kalkulator, program komputer excel, Rasch Model, Anates, SPSS, dan WarpPLS.

Dalam kesempatan ini, ucapan terima kasih ingin kami sampaikan kepada Penerbit dan Percetakan MEDIA ILMU PRESS beserta seluruh stafnya yang telah membantu menerbitkan dan menggandakan buku ini. Tak lupa saya sampaikan kepada istri tercinta dan ananda tersayang yang telah memberikan dorongan baik moril maupun materiil saat penulisan buku ini, sehingga selesai.

Disadari bahwa buku ini masih terdapat kekurangan dan kekhilafan. Oleh karena itu saran dan kritik yang bersifat membangun sangat diharapkan demi kesempatan buku ini di masa mendatang. Untuk itu disampaikan ucapan terima kasih yang sebesar-besarnya.

Semoga buku ini benar-benar bermanfaat bagi mahasiswa dimana saja berada. Amin.

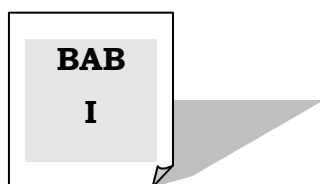
Kudus, Nopember 2016
Penyusun

DAFTAR ISI

	Halaman
Halaman Judul.....	i
Kata Pengantar.....	iii
Daftar Isi.....	iv
 BAB I PENDAHULUAN	 1
A. Rasionalisasi.....	1
B. Ruang Lingkup Materi.....	1
C. Pengertian Validitas dan Reliabilitas dalam Pengukuran.....	2
D. Dasar-Dasar Statistik untuk Penelitian	4
1. Skala Pengukuran	4
2. Tingkat Kepercayaan (<i>Confidence Interval</i>).....	5
3. Signifikansi/Probabilitas (<i>Significance Level</i>).....	5
4. Jumlah Data/Kasus	5
5. Tipe Data.....	5
6. Derajat Kebebasan (<i>Degree of Freedom</i>).....	6
7. Pengertian Hipotesis Serta Uji Hipotesis One Tailed dan Two Tailed.....	7
8. Nilai Kritis (<i>Critical Value</i>).....	12
9. Statistik Parametrik dan Non Parametrik	12
 BAB II TAKSONOMI TUJUAN PEMBELAJARAN.....	 13
A. Deskripsi Taksonomi	13
B. Kawasan Kognitif	14
1. Pengertian Ranah Penilaian Kognitif	14
C. Kawasan Afeksi (Sikap dan Perilaku)	20
D. Kawasan Psikomotor	22
 BAB III PENGEMBANGAN INSTRUMEN PENGUKURAN	 29
A. Pengantar	29
B. Tahap-tahap Pengembangan Alat Ukur	29
1. Perbedaan Pengertian Aspek dan Dimensi dalam Pengembangan Alat Ukur	30
2. Penulisan Butir	31
3. Jumlah Butir dalam Alat Ukur.....	31
4. Uji Coba Alat Ukur.....	32
5. Analisis Butir	33
6. Merakit Alat Ukur.....	34
C. Teknik Pengukuran	34
1. Skala Thurstone.....	35
2. Skala Guttman.....	36
3. Skala Likert	37
4. Tomat Kebahagiaan Thurstone & Likert	37

D. Desain Pembuatan Instrumen	38
1) Kuesioner Juga Punya Tingkat Kesulitan	39
E. Penyusunan Instrumen	39
a) Hal-Hal yang Perlu Dipertimbangkan	40
b) Hal-Hal yang Perlu Diperhatikan	41
F. Penggunaan Opsi Tengah	41
 BAB IV PENGEMBANGAN INSTRUMEN RANAH KOGNITIF	43
A. Pengantar	43
B. Penerapan Pengukuran dalam Domain Kognitif	43
1. Karakteristik Ranah Penilaian Kognitif	43
2. Pengukuran Ranah Penilaian Kognitif	48
3. Instrumen Tes Hasil Belajar (<i>Achievement Test</i>)	50
3.1. Jenis Tes	52
3.1.1. Bentuk tes obyektif	52
a) Tes Benar Salah (<i>True-False</i>)	53
b) Tes Pilihan Ganda (<i>Multiple Choice Test</i>)	53
c) Tes <i>Essay</i> (Tes Subjektif)	55
 BAB V PENGEMBANGAN INSTRUMEN RANAH AFEKTIF	59
A. Pengantar	59
B. Penerapan Pengukuran dalam Domain	59
1. Karakteristik Domain Afektif yang Penting	59
C. Pengembangan Instrumen Ranah Afektif	66
D. Skala Instrumen Penilaian Afektif dan Sistem Pemberian Skorya	67
E. Telaah Instrumen dan Perakitannya	69
F. Uji Coba Instrumen dan Analisis Hasil Uji Coba	70
 BAB VI PENGEMBANGAN INSTRUMEN RANAH PSIKOMOTOR	71
A. Pengertian Ranah Penilaian Psikomotor	71
B. Karakteristik Ranah Penilaian Psikomotor	71
C. Pengukuran Ranah Penilaian Psikomotor	75
D. Penerapan Pengukuran dalam Domain	77
1. Kawasan Psikomotor	77
 BAB VII PENGUJIAN INSTRUMEN PENELITIAN DAN PENILAIAN	79
A. Pengertian Validitas dan Reliabilitas	97
1. Validitas	80
a) Validitas Isi (<i>content validity</i>)	80
b) Validitas Konstruksi (<i>construct validity</i>)	80
c) Validitas Kriteria Terkait (<i>criterion related validity</i>)	81
2. Reliabilitas	81
a) Reliabilitas stabilitas (<i>test-retest method</i>)	82

b) Reliabilitas ekivalensi	83
c) Konsistensi internal	83
1) Metode Belah-Dua (<i>Split-Half Method</i>).....	84
2) Metode Kuder Richadson, Alpha Cronbach's, dan Anova Hoyt.....	84
d) Kesepakatan penilai (<i>Agreement</i>)	85
B. Prosedur Pengujian Instrumen Penelitian	86
a. Analisis Tingkat kesukaran atau disebut dengan istilah (<i>item difficulties</i>)	86
b. Analisis daya pembeda tes (<i>item discriminating power</i>).....	86
c. Analisis efektifitas pengecoh (<i>distractor</i>).....	87
1. Pengujian Instrumen Tes	87
1.1. Pengujian Validitas Instrumen Tes.....	89
1.2. Menentukan Uji reliabilitas instrument Tes	96
1.2.1. Pengujian Reliabilitas Instrumen untuk Item Berbentuk Tes Obyektif	96
a. Formula Spearman Brown.....	96
b. Formula KR-20	96
1.2.2. Pengujian Reliabilitas Instrumen untuk Item Berbentuk Tes Subyektif.....	103
1.3. Uji taraf Kesukaran Item	106
1.4. Analisis Daya Pembeda.....	108
1.5. Analisis Efektifitas Pengecoh (<i>Distractor</i>).....	110
2. Pengujian Instrumen Non Tes.....	111
2.1. Uji Validitas Instrumen Non Tes	111
2.2. Uji Reliabilitas Instrumen Non Tes	115
 BAB VIII APLIKASI PROGRAM DALAM PENGUJIAN INSTRUMEN	
PENILIAN DAN PENELITIAN	121
A. Pengujian Instrumen Penelitian dengan Aplikasi Program SPSS.....	121
1. CONTOH DATA DARI TABULASI Uji Coba Lapangan	121
2. Pengolahan dengan Program SPSS.....	122
2.1. Uji Validitas	124
2.2. Uji Reliabilitas	127
2.3. Interpretasi	131
2.3.1. Uji Reliabilitas	131
2.3.2. Interpretasi	131
B. Pengujian Instrumen Penelitian dengan Aplikasi Program Anates.....	131
C. Pengujian Instrumen Penelitian dengan Aplikasi Program Rasch Model.....	144
D. Langkah-Langkah Menjalankan Program WarpPLS.....	177
 Daftar Pustaka	206
Lampiran	210
1. Table t-test	211
2. Tabel r Product Moment.....	212
3. Tabel Chi Square (χ^2)	213
4. Tabel Rho (Spearman)	214



PENDAHULUAN

A. Rasionalisasi

Instrumen penelitian merupakan faktor yang penting dalam menentukan keberhasilan penelitian, baik penelitian sosial maupun penelitian pendidikan. Instrumen tersebut dapat berupa bentuk tes dan angket. Tes biasanya digunakan untuk mengumpulkan data berupa kemampuan seseorang dalam bidang tertentu, sedangkan angket digunakan untuk mengetahui gambaran data tentang sikap seseorang terhadap sesuatu.

Penyusunan instrumen tersebut tentu membutuhkan kemampuan peneliti atau mahasiswa dalam memahami teori sebagai pijakan dalam mengembangkan dimensi-dimensi, indikator-indikator sampai pada pembuatan soal, baik bentuk tes maupun angket.

Pada umumnya mahasiswa mengalami kesulitan untuk membedakan kedua bentuk instrumen tersebut, diharapkan hadirnya buku ini, akan membantu mahasiswa dalam membuat instrumen tersebut dan menggunkan sesuai dengan obyek yang diteliti.

B. Ruang Lingkup Materi

Ruang lingkup materi dalam pengembangan instrumen penelitian pendidikan, secara umum lingkup bahasan dalam tulisan ini menyangkut masalah penyusunan instrumen penelitian yang dipilah dalam dua bagian yakni : (a) instrumen jenis tes hasil belajar, dan (b) instrumen jenis angket.

Pengembangan Instrumen tes mengacu pada tujuan belajar yaitu taksonomi tujuan pembelajaran, yaitu taksonomi pembelajaran Benyamin S. Bloom untuk domain kognitif, Krathwol untuk afektif, dan psikomotor menggunakan domain Simpson.

Selanjutnya, pada setiap domain dikemukakan dimensinya dimana untuk masing-masing sebagai berikut :

1. domain kognitif memiliki enam dimensi yakni: pengetahuan, pemahaman, penerapan, analisis, sintesis, dan evaluasi.
2. domain afektif memiliki lima dimensi yakni: kemauan menerima, kemampuan menanggapi, keyakinan, mengorganisasi, dan membentuk pola.
3. domain psikomotor terdiri dari enam dimensi yakni: persepsi, kesiapan, gerakan kompleks, dan penyesuaian keaslian.

Untuk mempermudah dalam pengembangan instrumen tes belajar, maka dimensi-dimensi yang ada dikembangkan dalam bentuk indikator-indikator dan kata kerja operasional. Indikator dan kata kerja operasional inilah yang dijadikan acuan untuk menyusun unjuk kerja secara hirarkis dari tingkah laku yang paling rendah sampai pada tingkah laku yang paling tinggi.

Sedangkan pengembangan instrumen jenis angket, diawali dari desain dan definisi operasional variabel yang diteliti. Setiap variabel dikembangkan dalam dimensi-dimensi variabel

dan dalam setiap dimensi dikembangkan dalam indikator-indikator penelitian. Kemudian dikembangkan dalam bentuk soal-soal atau item angket.

Para peneliti untuk mempermudah dalam kegiatan pengembangan instrumen, disusun dalam bentuk kisi-kisi instrumen penelitian atau *layout instrument*.

C. Pengertian Validitas dan Reliabilitas dalam Pengukuran

Setiap penelitian yang menggunakan pendekatan kuantitatif dalam berbagai macam bidang studi pada umumnya variabel-variabel penelitiannya dirumuskan sebagai variabel laten (atau sering juga disebut faktor atau konstruk) yaitu variabel yang tidak diukur secara langsung, tetapi dibentuk melalui dimensi-dimensi yang diamati atau indikator-indikator yang diamati, dengan menggunakan kuesioner/angket.

Skala yang sering dipakai dalam penyusunan kuesioner adalah skala ordinal atau skala LIKERT, yaitu skala yang berisi lima tingkat jawaban yang merupakan skala jenis ordinal dengan pilihan sebagai berikut : (1) sangat setuju, (2) setuju (3) ragu-ragu, (4) tidak setuju, (5) sangat tidak setuju.

Pembuatan alat atau instrumen sebagai pengumpul data memiliki peran yang sangat besar bagi kesuksesan dalam memperoleh data-data penelitian yang shahih dan objektif pada kegiatan penelitian. Oleh karena itu instrumen tersebut perlu dilakukan pengujian validatas dan reliabilitasnya.

Instrumen yang valid berarti alat ukur yang digunakan untuk mendapatkan data (mengukur) itu valid. Valid berarti instrumen tersebut dapat digunakan untuk mengukur apa yang hendak diukur. Meteran yang valid dapat digunakan untuk mengukur panjang dengan teliti, karena meteran memang alat untuk mengukur panjang Meteran tersebut menjadi tidak valid jika digunakan untul mengukur berat. Instrumen yang reliabel berarti instrumen yang bila digunakan beberapa kali untuk mengukur obyek yang sama, akan menghasilkan data yang sama. Alat ukur panjang dari karet adalah contoh instrumen yang tidak reliabel.

Dengan menggunakan instrumen yang valid dan reliabel dalam pengumpulan data, maka diharapkan hasil penelitian akan menjadi valid dan reliabel. Jadi instrumen yang valid dan reliabel merupakan syarat untuk mendapatkan hasil penelitian yang valid dan reliabel. Hal ini tidak berarti bahwa dengan menggunakan instrumen yang telah teruji validitas dan reliabilitasnya, otomatis hasil (data) penelitian menjadi valid dan reliabel.

Hal ini masih akan dipengaruhi oleh kondisi obyek yang diteliti. Peneliti harus mampu mengendalikan obyek yang diteliti dan meningkatkan kemampuan dan menggunakan instrumen untuk mengukur variabel yang diteliti.

Instrumen yang reliabel belum tentu valid. Meteran yang putus di bagian ujungnya, bila digunakan berkali-kali akan menghasilkan data yang sama (reliabel) tetapi selalu tidak valid. Hal ini disebabkan karena instrumen (meteran) tersebut rusak. Reliabilitas instrumen merupakan syarat untuk pengujiannya validitas instrumen. Oleh karena itu walaupun instrumen yang valid umumnya pasti reliabel, tetapi pengujian reliabilitas instrumen perlu dilakukan.

Pada dasarnya terdapat dua macam instrumen, yaitu instrumen yang berbentuk tes untuk mengukur prestasi belajar dan instrumen yang non tes untuk mengukur sikap. Instrumen yang berupa tes jawabannya adalah "salah atau benar", sedangkan instrumen sikap jawabanya; tidak ada yang "salah atau benar" tetapi bersifat "positif dan negatif.

Pada gambar tersebut ditunjukkan bahwa instrumen yang baik (yang berupa tes maupun non tes), harus valid dan reliabel. Instrumen yang valid harus mempunyai validitas internal dan eksternal. Instrumen yang mempunyai validitas internal dan eksternal. Instrumen yang mempunyai validitas internal atau rasional, bila kriteria yang ada dalam instrumen secara rasional (teoritis) telah mencerminkan apa yang diukur. Jadi kriterianya ada di dalam instrumen itu.

Sedangkan instrumen yang mempunyai validitas eksternal bila kriteria di dalam instrumen disusun berdasarkan luar atau fakta-fakta empiris yang telah ada. Kalau validitas internal instrumen dikembangkan menurut teori yang relevan, maka validitas eksternal instrumen dikembangkan dari fakta empiris. Misalnya akan mengukur kinerja (*performance*) sekelompok pegawai, maka tolak ukur (kriteria) yang digunakan berdasarkan pada tolok ukur yang telah ditetapkan di kepegawaian itu, sedangkan validitas internal dikembangkan dari teori-teori.

Validitas internal instrumen yang berupa tes harus memenuhi *construct validity* (validitas konstruk) dan *content validity* (validitas isi). Sedangkan untuk instrumen yang non tes yang digunakan untuk mengukur sikap cukup memenuhi validitas konstruksi. Sutrisno Hadi (1986) menyamakan *construct validity* dengan *logical validity* dan *validity by definition*. Instrumen yang mempunyai validitas konstruksi, jika instrumen tersebut dapat digunakan untuk mengukur gejala sesuai dengan yang didefinisikannya. Misalnya akan mengukur efektivitas kerja, maka perlu didefinisikan terlebih dahulu apa itu efektivitas kerja. Setelah itu disiapkan instrumen yang digunakan untuk mengukur efektivitas kerja sesuai dengan definisi. Untuk melahirkan definisi, maka diperlukan teori-teori. Dalam hal ini Sutrisno Hadi menyatakan bahwa "bila bangunan teorinya sudah benar maka hasil pengukuran dengan alat ukur (instrumen) yang berbasis pada teori itu sudah dipandang sebagai hasil yang valid".

Instrumen yang harus mempunyai validitas ini adalah instrumen yang digunakan untuk mengukur prestasi belajar (*achievement*) dan mengukur efektivitas pelaksanaan program: dan tujuan. Untuk menyusun instrumen prestasi belajar yang mempunyai validitas isi (*content validity*), maka instrumen harus disusun berdasarkan materi pelajaran yang telah diajarkan. Sedangkan instrumen yang digunakan untuk mengetahui pelaksanaan program, maka instrumen disusun berdasarkan program yang telah direncanakan. Selanjutnya instrumen yang digunakan untuk mengukur tingkat tercapainya tujuan (efektivitas) maka instrumen harus disusun berdasarkan tujuan yang telah dirumuskan.

Dalam uji reliabilitas sebenarnya adalah alat untuk mengukur suatu kuesioner yang merupakan indikator dari variabel atau konstruk. Suatu kuesioner dikatakan reliabel atau handal, jika jawaban seseorang terhadap kenyataan konsisten atau stabil dari waktu-kewaktu.

Pengukuran reliabilitas dapat dilakukan dengan dua cara yaitu :

1. *Repeated Measure* atau pengukuran ulang. Disini seseorang akan diberikan pertanyaan yang sama pada waktu yang berbeda, dan dilihat apakah ia tetap konsisten dengan jawabannya.
2. *One Shot* atau pengukuran sekali saja. Pengukuran dilakukan sekali saja dan kemudian hasilnya dibandingkan dengan pertanyaan lain atau mengukur korelasi antar jawaban pertanyaan.

Untuk melakukan uji reliabilitas dapat digunakan program SPSS dengan menggunakan uji statistik Cronbach Alpha. Adapun kriteria bahwa instrumen itu dikatakan reliabel, apabila nilai yang didapat dalam proses pengujian dengan uji statistik Cronbach Alpha $> 0,60$ (Nunnally, 1969). Dan sebaliknya jika Cronbach Alpha ditemukan angka koefisien lebih kecil ($< 0,60$), maka dikatakan tidak reliabel.

D. Dasar-Dasar Statistik untuk Penelitian

1. Skala Pengukuran

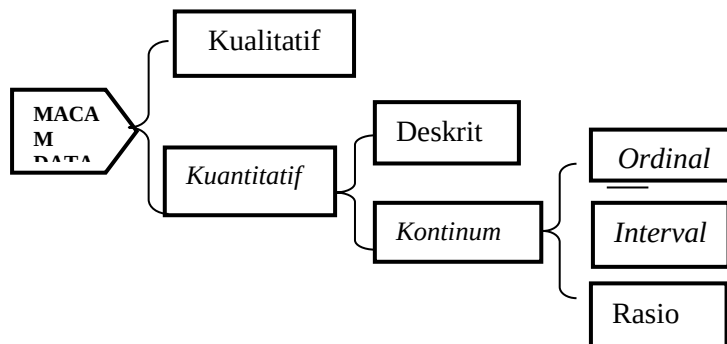
Ada empat tipe skala pengukuran dalam penelitian, yaitu nominal, ordinal, interval dan rasio.

Dalam kehidupan sehari-hari kita akan menjumpai berbagai macam data atau fakta pada lingkungan sekitar kita. Data atau fakta dalam ilmu-ilmu sosial dapat ditangkap dalam wujud gejala sosial yang ada. Gejala tersebut dianggap sebagai gejala normal, yang dapat digolong-golongkan secara terpisah (dikhotomik), diskrit, kategorik. Kita hanya bisa menghitung banyaknya subjek. Problema pengukuran tidak ada di dalam gejala ini. Mengukur berarti mengidentifikasi besar kecilnya suatu gejala, tidak menghitung berapa jumlah atau berapa banyaknya, dan tidak berlaku persekataan atau penjajagan pada data ini, tidak dikenal skala nominal, sebab skala selalu menunjuk kepada jenjang, proporsi atau ratio. Skala nominal hanya untuk mengidentifikasi kelompok. (Hadi, 1977 : 106). Skala pengukuran nominal digunakan untuk mengklasifikasi objek, individual atau kelompok.

Gejala ditangkap sebagai gejala kontinum, yaitu gejala yang bervariasi menurut tingkatan, ia memiliki kontinuitas dengan ciri-ciri untuk menggolong-golongkan subjek pendukung gejala ini menjadi beberapa tingkatan, derajat atau jenjang, seperti kepandaian : pandai sekali, pandai, cukup pandai, kurang pandai, dan tidak pandai.

Dalam gejala kontinum dikenal persekataan sebagai berikut :

1. Skala ordinal (berjenjang), menunjukkan posisi penjenjangan sebagai urutan-urutan dari derajat tertinggi ke derajat terendah. Skala pengukuran ordinal memberikan informasi tentang jumlah relatif karakteristik berbeda yang dimiliki oleh objek atau individu tertentu. Pada skala ini proses penghitungan dengan menggunakan rumus-rumus statistik nonparametrik untuk mencari signifikansi, yaitu median, persentil, Spearman r, Kendal T dan Kendal W.



Gambar 1. Macam Data

2. Skala interval mempunyai karakteristik seperti yang dimiliki oleh skala nominal dan ordinal dengan ditambah karakteristik lain, yaitu berupa adanya interval yang tetap. Skala interval adalah jarak pengukuran antara dua angka, dengan jarak yang sama antara interval satu dengan lainnya. Contoh : Umur : 0-4; 5-9, 10-14, dan seterusnya. Rumus statistik non parametrik dapat digunakan yaitu mean, deviasi standar, Korelasi Pearson Phi, Fisher , dan dapat juga menggunakan statistik parameterik.

3. Skala ratio mempunyai karakteristik seperti yang dimiliki oleh skala nominal, ordinal, dan interval dengan ditambah karakteristik lain, yaitu memiliki nilai 0 (nol) empiris absolut. Skala ratio, berarti pengukuran berdasarkan perbandingan berawal dari angka atau titik nol (0) yang absolut. Alat pengukur misalnya meteran, timbangan berat. Untuk ilmu social kebanyakan adalah alat ukur yang tidak langsung. Umpamanya tinggi badan 175 cm lebih tinggi daripada 150 cm. Rumus yang dapat digunakan adalah statistik non parametrik atau parametrik, yaitu geometric mean dan coefficient of variation (Siegel, 1985 : 37).

2. Tingkat Kepercayaan (*Confidence Interval*)

Tingkat kepercayaan atau *confidence interval* atau *risk level* didasarkan pada gagasan yang berasal dari Teorema Batas Sentral (*Central Limit Theorem*) yaitu apabila suatu populasi secara berulang-ulang ditarik sampel, maka nilai rata-rata atribut yang diperoleh dari sampel-sampel tersebut sejajar dengan nilai populasi yang sebenarnya. Selanjutnya, nilai-nilai yang diperoleh tersebut yang berasal dari sampel-sampel yang sudah ditarik didistribusikan secara normal dalam bentuk nilai benar/nyata. Bentuk nilai-nilai yang diperoleh tersebut akan menjadi nilai-nilai sampel yang lebih tinggi atau lebih rendah jika dibandingkan dengan nilai populasinya.

Dalam suatu distribusi normal, sekitar 95% nilai-nilai sampel berada dalam dua simpangan baku (*standard deviation*) dari nilai populasi sebenarnya. Dengan kata lain jika tingkat kepercayaannya sebesar 95% dipilih, maka 95 dari 100 sampel akan mempunyai nilai populasi yang sebenarnya dalam jangkauan ketepatan sebagaimana sudah dispesifikasi sebelumnya. Adakalanya bahwa sampel yang kita peroleh tidak mewakili nilai populasi yang sebenarnya. Tingkat kepercayaan berkisar antara 99% yang tertinggi dan 90% yang terendah. Dalam SPSS tingkat kepercayaan secara default diisi 95%.

3. Signifikansi/Probabilitas (*Significance Level*)

Signifikansi atau disebut juga probabilitas merupakan tingkat ketepatan (*presisi*) dalam kaitannya dengan kesalahan pengambilan sampel (*sampling error*), merupakan jangkauan dimana nilai populasi yang tepat diperkirakan. Jangkauan ini sering diekspresikan dengan menggunakan poin-poin persentase, misalnya 1% atau 5%. Oleh karena itu, jika seorang peneliti menemukan bahwa 75% guru tertentu yang digunakan sebagai sampel sudah mengadopsi suatu metode pembelajaran berbasis informasi teknologi (IT) yang direkomendasikan dengan tingkat ketepatan sebesar $\pm 1\%$, maka peneliti tersebut dapat menyimpulkan bahwa antara 74% dan 76% dari guru tersebut menjadi populasi sudah mengadopsi metode pembelajaran berbasis IT tersebut. Dalam program SPSS signifikansi ditulis secara default sebagai 0,05 (5%).

4. Jumlah Data/Kasus

Dalam SPSS jumlah data disebut sebagai kasus. Cara membacanya ialah dengan melihat baris. Jadi jumlah baris nama dengan jumlah kasus/data. Di SPSS jumlah data ini diberi simbol N. SPSS tidak membedakan antara N (populasi) dan n (sampel).

5. Tipe Data

Data atau fakta yang ada dalam fenomena dapat dikelompokkan ke dalam bentuk *kategori* dan berbentuk *bilangan*.

Data yang berbentuk bilangan disebut *data kuantitatif* yang mempunyai variasi nilai, yang dikenal menjadi dua golongan yaitu : (1) Data diskrit (hasil menghitung); (2) Data kontinu (hasil pengukuran).

Contoh : *Data* yang bukan kuantitatif disebut *data kualitatif*, data yang bersifat ketagorik sebagai lukisan kualitas objek yang dipelajari. Golongan ini dieknl pula dengan nama *atribut*. Misalnya : sembuh, rusak, gagal, berhasil, dll.

Menurut sumbernya dapat di pilah menjadi dua bagian yaitu :

- 1) *data intern* yaitu data yang diperoleh dari hasil pencatatan dalam kegiatan suatu instansi.
- 2) *data ekstern* yaitu data yang diperoleh dari luar suatu instansi, yang biasanya dibutuhkan untuk perbandingan. Pada data ini pilah menjadi dua bagian yaitu :
 - a. *data ekstern primer (data primer)*, yaitu data yang diperoleh dari instansi yang sama;
 - b. *data ekstern sekunder (data sekunder)*, yaitu data yang diperoleh dari instansi yang berbeda (luar instansi).

Tipe data yang ada pada program SPSS 10.0 For Windows adalah :

- 1) *Numeric* merupakan tipe angka dengan tanda plus dan tnda minus di depan angka serta indikator desmal. Lebar maksimal 40 karakter.
- 2) *Comma* merupakan tipe yang termasuk angka, tanda plus dan tanda minus didepan angka, indikator desimal serta pemisah ribuan.
- 3) *Dot*, tipe ini sama dengan tipe comma yang membedakan hanyalah pemisah ribua, yang digunakan adalah titik.
- 4) *Secientific notation*, merupakan tipe data yang menggunakan lambang atau notasi ilmiah seperti log, alfa dll.
- 5) *Date*, tipe ini menampilkan data dalam format tanggal atau waktu.
- 6) *Dollar*, tipe ini adalah tanda \$, sebuah titik sebagai indikator desimal dan beberapa tanda koma pemisah ribuan.
- 7) *Custom Crency*, tipe ini digunakan untuk menampilkan format mata uang seperti Rp. 5000.
- 8) *String* digunakan untuk huruf dan karakter lainnya.

6. Derajat Kebebasan (*Degree of Freedom*)

Derajat kebebasan mempunyai dua makna yang berbeda dalam kaitanya dengan distribusi statistik untuk memberikan nama dari salah satu parameternya. Dalam kecocokan model, derajat kebebasan menunjuk pada jumlah informasi yang lain. Umumnya kita memulai jumlah derajat kebebasan (dk) dengan data. Semakin suatu prosedur atau model cocok, maka jumlah derajat kebebasan semakin kecil. Penghitungan dk dilakukan melalui ukuran sampel. Derajat kebebasan merupakan pengukuran jumlah informasi dari data sampel yang telah digunakan. Setiap penghitungan statistik dilakukan dari suatu sampel tertentu, maka satu derajat dk digunakan. Setiap rumus dalam SPSS cara menghitung dk (*DF/Degree of Freedom*) berbeda, misalnya dalam Chi Square untuk menghitung DF digunakan rumus $(C-1) \times (R-1)$; sedang untuk uji t sampel bebas untuk $df = n-2$; untuk uji t sampel berpasangan untuk menghitung DF digunakan rumus $n-1$ dan seterusnya.

7. Pengertian Hipotesis Serta Uji Hipotesis *One Tailed* dan *Two Tailed*

Pada bagian ini akan memberikan pengertian dasar mengenai hipotesis, cara membuat hipotesis, taraf kesalahan dalam merumuskan hipotesis dan pengujian hipotesis.

Penelaahan yang mendalam terhadap berbagai macam sumber berkaitan dengan fokus penelitian, diharapkan dapat menentukan anggapan dasar, dan kemudian dilanjutkan dengan merumuskan hipotesis penelitian.

Anggapan dasar tersebut dapat diupayakan dengan cara :

1. Membaca buku, surat kabar, atau terbitan lain yang berkaitan dengan focus penelitian, baik sumber secara umum atau khusus.
2. Dengan banyak membaca berita, ceramah, diskusi dll.
3. Banyak berkunjung ke tempat yang akan diteliti.
4. mengadakan pendugaan, mengabstraksi berdasarkan perbendaharaan pengetahuannya.

Dari arti katanya hipotesis memang berasal dari 2 penggalan kata yaitu “hypo” yang artinya “dibawah” dan “thesa” yang artinya “kebenaran”. Dengan demikian **hipotesis** dapat diartikan : *sebagai suatu jawaban yang bersifat sementara terhadap permasalahan penelitian, sampai terbukti melalui data yang terkumpul.*

Lebih lanjut dijelaskan oleh Nasution (2000) mengatakan bahwa definisi hipotesis ialah pernyataan tentatif yang merupakan dugaan mengenai apa saja yang sedang kita amati dalam usaha memahaminya. Sedangkan menurut Zikmund (1997: 112) mendefinisikan hipotesis sebagai: *“Unproven proposition or supposition that tentatively explains certain facts or phenomena; a probable answer to a research question”*. Hipotesis merupakan proposisi atau dugaan yang belum terbukti yang secara tentatif menerangkan fakta-fakta atau fenomena atau fenomena tertentu dan juga merupakan jawaban yang memungkinkan terhadap sesuatu pertanyaan riset.

Sikap peneliti terhadap hipotesis yang dibuat :

1. menerima keputusan seperti apa adanya seandainya hipotesisnya tidak terbukti pada akhir penelitian.
2. mengganti hipotesis seandainya melihat tanda-tanda bahwa data yang terkumpul tidak mendukung terbtinya hipotesis (pada saat penelitian berlangsung) serta dalam laporan penelitian harus dituliskan proses penggantian ini, dengan bertindak jujur dan tegas.

1) Merumuskan Hipotesis

Pada umumnya hipotesis dirumuskan untuk menggambarkan hubungan dua variable akibat. Namun demikian ada hipotesis yang menggambarkan perbandingan satu variable dari dua sample.

Hipotesis merupakan suatu pernyataan yang penting kedudukannya dalam penelitian. Oleh karena itu peneliti hendaknya merumuskan hipotesisnya dengan jelas. Sebagaimana disarankan oleh Borg dan Gall (1979 :61) mengajukan adanya persyaratan untuk hipotesis yaitu :

- a) Hipotesis harus dirumuskan dengan singkat tetapi jelas;
- b) Hipotesis harus dengan nyata menunjukkan adanya antara dua atau lebih variable;
- c) hipotesis harus didukung oleh teori-teori yang dikemukakan oleh para ahli atau hasil peneliti yang relevan.

2) Jenis-Jenis Hipotesis

Hipotesa, yang isi dan rumusannya bermacam-macam dapat dibedakan menjadi tiga bagian yaitu :

a) Hipotesa tentang perbedaan versus (vs) hubungan

Pada hipotesa ini, peneliti akan melihat apakah pernyataan sementara yang diberikan ada hubungan atau perbedaan. Hipotesa hubungan berkaitan dengan pernyataan rekaan yang menyatakan hubungan dua variable atau lebih. Hipotesis ini sebagai konsep dasar dari teknik korelasi dan regresi. Sedangkan perbedaan menyatakan adanya ketidaksamaan antar variable tertentu disebabkan oleh adanya pengaruh variable yang berbeda-beda. Hipotesis ini sebagai dasar konsep teknik penelitian komparatif.

b) Hipotesa kerja (H_a) versus (vs) hipotesa nully (H_n)/ H_o .

Hipotesa dilihat dari cara menyusun pernyataan hipotesis dalam penelitian dapat dibedakan menjadi hipotesa kerja dan hipotesis nully.

Hipotesa kerja dirumuskan oleh peneliti untuk menetapkan pernyataan dalam pengujian untuk *diterima* dalam penelitian. Hipotesa nully biasanya digunakan dalam penelitian eksperimental, akan tetapi sekarang ini sudah banyak yang digunakan dalam penelitian sosiologi ataupun pendidikan. Hipotesa null pertama kali diperkenalkan oleh **Fisher**, diformulasikan untuk *ditolak* dalam penelitian. Hipotesa kerja dan hipotesa nully dapat dirumuskan sebagai berikut :

H_a : “Ada.....dengan.....”

H_a : “Andaikata....., maka.....”

H_o : “Tidak ada beda antara dengan.....”

H_o : “.....tidak mem.....”

c) Hipotesa *common sense* dan ideal

Hipotesa *common sense* (akal sehat) yang berkaitan dengan terkaan tentang dalil pemikiran bersahaja dan *common sense* (akal sehat). Hipotesa ini berkaitan dengan pernyataan *hubungan keseragaman kegiatan terapan*. Contoh : hipotesa tentang produksi dan status pemilikan tanah. Input siswa dalam pendidikan dan kualitas output pendidikan.

Hipotesa jenis ideal yaitu hipotesa yang menyatakan hubungan yang kompleks. Hipotesa ini bertujuan untuk menguji adanya hubungan yang logis antara keseragaman-keseragaman pengalaman empiris. Peningkatan dari hipotesa ini adalah *hipotesa analitis*.

Contoh : Tentang hubungan letak kios dengan peningkatan omset dalam perdagangan. Hubungan antara potensi siswa dengan keberhasilan siswa dalam pendidikan. jika kita rinci, maka hubungan diatas akan menjadi hipotesa analitis yaitu mencari hubungan kompleksitas dari unsure variable dalam penelitian tersebut.

3) Tiga Bentuk Rumusan Hipotesa

Menurut tingkat eksplanasi hipotesis yang akan diuji, maka rumusan hipotesis dapat dikelompokkan menjadi 3 macam yaitu :

a) *Hipotesis Deskriptif*

Adalah dugaan tentang nilai suatu variable mandiri, tidak membuat perbandingan atau hubungan.

Contoh : Rumusan masalah dalam penelitian :

Seberapa tinggi daya tahan lampu merk X ?

Seberapa tinggi produktivitas padi di Kabupaten X ?

Berapa lama daya tahan lampu merk A dan B ?

Seberapa baik gaya kepemimpinan di lembaga X ?

Maka dapat dirumuskan hipotesis sebagai berikut :

- Daya tahan lamp merk X = 800 jam
- Produktivitas padi di Kab. X 8 ton/ha.
- Daya tahan lampu merk A=450 jam dan merk B=600 jam
- Gaya kepemimpinan di lembaga X telah mencapai 70% yang diharapkan.

Rumusan hipotesis statistiknya adalah :

$H_o : \mu \leq 800 \text{ jam}$

$H_a : \mu \geq 800 \text{ jam dst.}$

b) Hipotesis Korelasi

Adalah suatu pernyataan yang menunjukkan dugaan tentang hubungan antara dua variable atau lebih.

Contoh :

Rumusan masalah penelitian : “Adakah hubungan antara Gaya Kepemimpinan dengan Efektifitas Kerja ?”

H_o : “Tidak ada hubungan antar gaya kepemimpinan dengan efektifitas kerja”.

H_a : “Ada hubungan antar gaya kepemimpinan dengan efektivitas kerja”.

Rumusan hipotesis statistiknya adalah :

$H_o : \rho = 0$

$H_a : \rho \neq 0$ (ρ = symbol yang menunjukkan kuatnya hubungan)

c) Hipotesis Komparatif

Adalah pernyataan yang menunjukkan dugaan nilai dalam satu variable atau lebih pada sample yang berbeda.

Contoh :

Rumusan masalah penelitin :

Adakah perbedaan produktifitas kejra antara pegawai gholongan I, II, dan III ?

Rumusan hipotesis adalah :

Tidak terdapat perbedaan (persamaan) produktivitas kerja antara pegawai golongn I, II, III.

Rumusan hipotesis statistiknya adalah :

$H_o = \mu_1 = \mu_2 = \mu_3$

$H_a = \mu_1 \neq \mu_2 \neq \mu_3$ (salah satu berbeda sudah merupakan H_a).

4) Pengujian Hipotesis

Untuk mempermudah pemahaman teknik statistik yang digunakan dalam menguji hipotesis dapat dilihat pada tabel di bawah ini.

Tabel 1. Statistik Yang Digunakan Untuk Menguji Hipotesis

Jenis/Tingkatan Data	Teknik Statistik Yang Digunakan Untuk Pengujian
Nominal	Test Binomial Chi Kuadrat (1 sampel)
Ordinal	Run test
Interval/Rasio	t-test (1 sampel)

Tabel 2
Pedoman Untuk Memilih Teknik Korelasi Dalam Pengujian Hipotesis

Macam/Tingkatan Data	Teknik Korelasi Yang Digunakan
Nominal	1. Koefisien Kontingency
Ordinal	Spearman Rank Kendal Tau
Interval dan rasio	1. Pearson Product Moment 2. Korelasi ganda 3. Korelasi Parsial

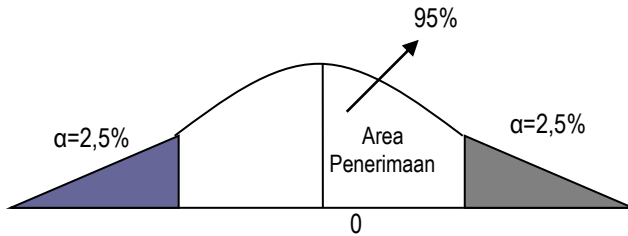
Tabel 3
Berbagai Teknik Statistik Untuk Menguji Hipotesis Komparatif

MACAM DATA	BENTUK KOMPARASI			
	Dua Sampel		k Sampel	
	Korelasi	Independen	Korelasi	Independen
Interval/ Rasio	t-test dua sampel	t-test dua sample	One Way Anova Two Way Anova	One Way Anova Two Way Anova
Nominal	Mc Nemar	Fisher Exact Chi Kuadrat Two sample	Chi Kuadrat for k sample Cochran Q	Chi Kuadrat for k sample
Ordinal	Sign test Wicoxon Matched Pairs	Median Test Mann-Whitney U test Kolmogorov Smirnov Wald-Wolfowitz	Friedman Two Way Anova	Median Extension Kruskal-Walls One Way Anova

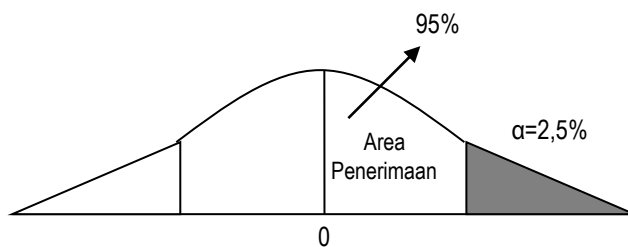
Menggunakan kurva untuk menguji hipotesis dapat digambarkan sebagai berikut:

Gambar 2. Uji Dua Sisi

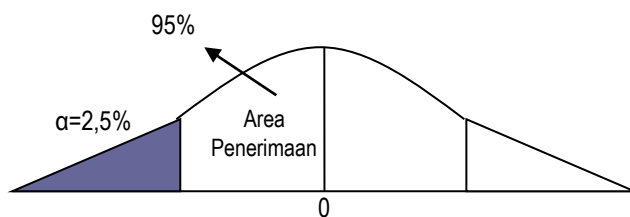
a) Untuk uji dua sisi



b) Untuk uji satu sisi sebelah kanan



c) Untuk uji satu sisi sebelah kiri



5) Taraf Kesalahan dalam Merumuskan Hipotesis

Proses penelitian dilakukan dengan langkah-langkah pengumpulan data, pengolahan, uji hipotesis, analisis dan penafsiran, dan pengambilan kesimpulan dan pemecahan masalah, yang bersifat terus-menerus (kontinu) sepanjang hidup manusia, untuk dapat mencari dan menemukan kebenaran, atau teori dan ilmu. Tingkat Kebenaran tersebut dalam teori statistika tidak dapat mencapai 100 %. Derajat kebenaran tersebut hasil penelitian selalu dibawah 100 %, mis. 99 %, 95 % dan lain-lain. Untuk memperjelas tingkat kekeliruan dalam membuat kesimpulan dapat dilihat dalam tabel 4.

Tabel. 1. Tingkat Kekeliruan dalam Hipotesis Penelitian

Kesimpulann	Keadaan Sebenarnya	
	Hipotesis Benar	Hipotesis Salah
Terima hipotesis	Benar	Keliru (Kekeliruan tipe II)
Tolak hipotesis	Keliru (Kekeliruan Tipe I)	Benar

Hipotesis dapat dijadikan sebagai arah dalam menetapkan variabel, mengumpulkan data, mengolah data dan mengambil kesimpulan sebagai bentuk terakhir temuan. Pada dasarnya, pekerjaan meneliti adalah usaha untuk membuktikan hipotesis.

Ada dua macam hipotesis kerja atau hipotesis alternatif (H_a), dan H_0 (hipotesis nihil) atau hipotesis statistis.

Ada dua kekeliruan yang kita buat dalam membuat hipotesis yaitu :

1. Menolak hipotesis yang seharusnya diterima, disebut kekeliruan alpha (α).
2. Menerima hipotesis yang seharusnya di tolak, disebut kekeliruan beta (β)

Tingkat kesalahan ini dinamakan *level of signifikan* atau tingkat signifikansi. Dalam prakteknya tingkat signifikansi telah ditetapkan oleh peneliti terlebih dahulu

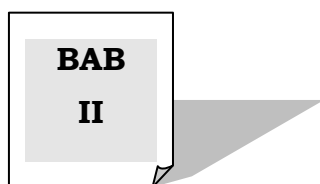
Pada umumnya cara menguji hipotesis yaitu Menggunakan daerah kurva normal. Apabila harga Z- Score terletak di daerah penerimaan H_0 , maka H_a yang dirumuskan tidak diterima.

8. Nilai Kritis (*Critical Value*)

Nilai kritis digunakan untuk pengujian signifikansi. Nilai di mana pengujian statistik harus melampaui nilai tertentu agar hipotesis H_0 ditolak. Misalnya nilai kritis t dengan derajat kebebasan (dk) = 12 dan pada tingkat signifikansi sebesar 0,05 adalah 1,98. Nilai absolut t atau t hitung harus lebih besar dari 1,98 agar H_0 ditolak. Nilai kritis diambil dari tabel nilai kritis t, sedang nilai absolut berasal dari data.

9. Statistik Parametrik dan Non Parametrik

Statistik parametrik merupakan statistik di mana populasi diasumsikan cocok dengan setiap distribusi yang diukur, umumnya distribusi normal. Metode statistik inferensi parametrik merupakan prosedur matematik untuk pengujian hipotesis statistik yang mengasumsikan bahwa distribusi variabel-variabel tersebut sedang dinilai sesuai dengan kelompok parameter yang berdistribusi normal. Contoh dalam SPSS: ANOVA dan Korelasi Pearson. Sebaliknya, statistik nonparametrik berkaitan dengan model-model non parametrik dan inferensi non parametrik. Pengertian ini juga mengacu pada statistik yang interpretasinya tidak tergantung pada populasi yang dicocokkan dengan setiap distribusi normal. Contoh dalam SPSS : Korelasi Phi dan Spearman.



TAKSONOMI TUJUAN PEMBELAJARAN

A. Deskripsi Taksonomi

Dalam perkembangannya banyak teori tentang pembelajaran, tetapi hanya ada tiga teori pembelajaran yang menurut Shrock dan Coscarelli (1989) penting untuk menjadi landasan evaluasi atau penelitian yang bertujuan untuk mengukur keberhasilan dalam pelaksanaan pendidikan. Adapun ketiga teori tersebut adalah taksonomi pembelajaran dari Bloom, teori kemampuan belajar (*learning capabilities*) dari Gagne, dan teori pajangan komponen (*component display theory* atau *component design theory*) dari Merrill. Ketiga teori tersebut memiliki kesamaan karakter, yaitu adanya hierarki dari sejumlah konsep.

Penilaian adalah upaya atau tindakan untuk mengetahui sejauh mana tujuan yang telah ditetapkan itu tercapai atau tidak. Dengan kata lain, penilaian berfungsi sebagai alat untuk mengetahui keberhasilan proses dan hasil belajar siswa. Dalam sistem pendidikan nasional rumusan tujuan pendidikan, baik tujuan kurikuler maupun tujuan instruksional, menggunakan klasifikasi hasil belajar dari Benyamin Bloom yang secara garis besar membaginya menjadi tiga ranah, yakni ranah kognitif, ranah afektif, dan ranah psikomotorik.

Salah satu prinsip dasar yang harus senantiasa diperhatikan dan dipegangi dalam rangka evaluasi hasil belajar adalah prinsip kebulatan, dengan prinsip evaluator dalam melaksanakan evaluasi hasil belajar dituntut untuk mengevaluasi secara menyeluruh terhadap peserta didik, baik dari segi pemahamannya terhadap materi atau bahan pelajaran yang telah diberikan (aspek kognitif), maupun dari segi penghayatan (aspek afektif), dan pengamalannya (aspek psikomotor).

Ketiga aspek atau ranah kejiwaan itu erat sekali dan bahkan tidak mungkin dapat dilepaskan dari kegiatan atau proses evaluasi hasil belajar. Benjamin S. Bloom dan kawan-kawannya itu berpendapat bahwa pengelompokan tujuan pendidikan itu harus senantiasa mengacu kepada tiga jenis *domain* (daerah binaan atau ranah) yang melekat pada diri peserta didik, yaitu:

- a) Ranah proses berfikir (*cognitive domain*)
- b) Ranah nilai atau sikap (*affective domain*)
- c) Ranah keterampilan (*psychomotor domain*)

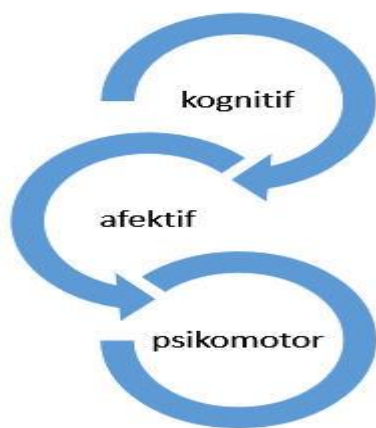
Dalam konteks evaluasi hasil belajar, maka ketiga domain atau ranah itulah yang harus dijadikan sasaran dalam setiap kegiatan evaluasi hasil belajar. Sasaran kegiatan evaluasi hasil belajar adalah:

1. Apakah peserta didik sudah dapat memahami semua bahan atau materi pelajaran yang telah diberikan pada mereka?
2. Apakah peserta didik sudah dapat menghayatinya?
3. Apakah materi pelajaran yang telah diberikan itu sudah dapat diamalkan secara kongkret dalam praktek atau dalam kehidupannya sehari-hari?

Ketiga ranah tersebut menjadi obyek penilaian hasil belajar. Diantara ketiga ranah itu, ranah kognitiflah yang paling banyak dinilai oleh para guru disekolah karena berkaitan dengan kemampuan para siswa dalam menguasai isi bahan pengajaran.

Tujuan pembelajaran secara umum biasanya diarahkan pada kawasan dari taksonomi tujuan pembelajaran Benyamin S Bloom yang meliputi kawasan (1) kognitif, (2) afektif, (3) psikomotor. Meskipun demikian Benyamin S. Bloom hanya mengkonsentrasikan pada domain kognitif, sementara domain afektif dikembangkan oleh Krathwohl, dan domain psikomotor dikembangkan oleh Simpson.

Hubungan ketiga taksonomi tersebut bersifat sirkuler diantara ranah kognitif, afektif dan psikomotor. Hal ini memiliki makna bahwa dalam memberikan aspek kemampuan dan ketrampilan pada siswa memerlukan proses yang terus-terus (*continuity*) dan dilakukan tahap demi tahap (*step by step*). Sebagaimana dapat dilihat pada gambar 3 dibawah ini.



Gambar 3. Hubungan Sirkulir antara Tiga Taksonomi

B. Kawasan Kognitif

1. Pengertian Ranah Penilaian Kognitif

Ranah kognitif adalah ranah yang mencakup kegiatan mental (otak). Menurut Bloom, segala upaya yang menyangkut aktivitas otak adalah termasuk dalam ranah kognitif. Ranah kognitif berhubungan dengan kemampuan berfikir, termasuk didalamnya kemampuan menghafal, memahami, mengaplikasi, menganalisis, mensintesis, dan kemampuan mengevaluasi. Dalam ranah kognitif itu terdapat enam aspek atau jenjang proses berfikir, mulai dari jenjang terendah sampai dengan jenjang yang paling tinggi. Keenam jenjang atau aspek yang dimaksud adalah:

- 1) Pengetahuan/hafalan/ingatan (*knowledge*) adalah kemampuan seseorang untuk mengingat-ingat kembali (*recall*) atau mengenali kembali tentang nama, istilah, ide, rumus-rumus, dan sebagainya, tanpa mengharapkan kemampuan untuk menggunkannya. Pengetahuan atau ingatan adalah merupakan proses berfikir yang paling rendah. Salah satu contoh hasil belajar kognitif pada jenjang pengetahuan adalah dapat menghafal surat al-‘Ashar, menerjemahkan dan menuliskannya secara baik dan benar, sebagai salah satu materi pelajaran kedisiplinan yang diberikan oleh guru Pendidikan Agama Islam di sekolah.
- 2) Pemahaman (*comprehension*) adalah kemampuan seseorang untuk mengerti atau memahami sesuatu setelah sesuatu itu diketahui dan diingat. Dengan kata lain, memahami adalah mengetahui tentang sesuatu dan dapat melihatnya dari berbagai segi. Seseorang peserta didik dikatakan memahami sesuatu apabila ia dapat memberikan penjelasan atau memberi uraian

yang lebih rinci tentang hal itu dengan menggunakan kata-katanya sendiri. Pemahaman merupakan jenjang kemampuan berfikir yang setingkat lebih tinggi dari ingatan atau hafalan. Salah satu contoh hasil belajar ranah kognitif pada jenjang pemahaman ini misalnya: Peserta didik atas pertanyaan Guru Pendidikan Agama Islam dapat menguraikan tentang makna kedisiplinan yang terkandung dalam surat al-‘Ashar secara lancar dan jelas.

- 3) Penerapan (*application*) adalah kesanggupan seseorang untuk menerapkan atau menggunakan ide-ide umum, tata cara ataupun metode-metode, prinsip-prinsip, rumus-rumus, teori-teori dan sebagainya, dalam situasi yang baru dan kongkret. Penerapan ini adalah merupakan proses berfikir setingkat lebih tinggi ketimbang pemahaman. Salah satu contoh hasil belajar kognitif jenjang penerapan misalnya: Peserta didik mampu memikirkan tentang penerapan konsep kedisiplinan yang diajarkan Islam dalam kehidupan sehari-hari baik dilingkungan keluarga, sekolah, maupun masyarakat.
- 4) Analisis (*analysis*) adalah kemampuan seseorang untuk merinci atau menguraikan suatu bahan atau keadaan menurut bagian-bagian yang lebih kecil dan mampu memahami hubungan di antara bagian-bagian atau faktor-faktor yang satu dengan faktor-faktor lainnya. Jenjang analisis adalah setingkat lebih tinggi ketimbang jenjang aplikasi. Contoh: Peserta didik dapat merenung dan memikirkan dengan baik tentang wujud nyata dari kedisiplinan seorang siswa di rumah, di sekolah, dan dalam kehidupan sehari-hari di tengah-tengah masyarakat, sebagai bagian dari ajaran Islam.
- 5) Sintesis (*syntesis*) adalah kemampuan berfikir yang merupakan kebalikan dari proses berfikir analisis. Sintesis merupakan suatu proses yang memadukan bagian-bagian atau unsur-unsur secara logis, sehingga menjelma menjadi suatu pola yang berstruktur atau berbentuk pola baru. Jenjang sintesis kedudukannya setingkat lebih tinggi daripada jenjang analisis. Salah satu hasil belajar kognitif dari jenjang sintesis ini adalah: peserta didik dapat menulis karangan tentang pentingnya kedisiplinan sebagaimana telah diajarkan oleh Islam.
- 6) Penilaian/penghargaan/evaluasi (*evaluation*) adalah merupakan jenjang berpikir paling tinggi dalam ranah kognitif dalam taksonomi Bloom. Penilaian/evaluasi disini merupakan kemampuan seseorang untuk membuat pertimbangan terhadap suatu kondisi, nilai atau ide, misalkan jika seseorang dihadapkan pada beberapa pilihan maka ia akan mampu memilih satu pilihan yang terbaik sesuai dengan patokan-patokan atau kriteria yang ada. Salah satu contoh hasil belajar kognitif jenjang evaluasi adalah: peserta didik mampu menimbang-nimbang tentang manfaat yang dapat dipetik oleh seseorang yang berlaku disiplin dan dapat menunjukkan mudharat atau akibat-akibat negatif yang akan menimpa seseorang yang bersifat malas atau tidak disiplin, sehingga pada akhirnya sampai pada kesimpulan penilaian, bahwa kedisiplinan merupakan perintah Allah SWT yang wajib dilaksanakan dalam sehari-hari.

Keenam jenjang berpikir ranah kognitif bersifat kontinum dan overlap (tumpang tindih), dimana ranah yang lebih tinggi meliputi semua ranah yang ada dibawahnya. *Pengetahuan* (1) adalah merupakan jenjang berpikir paling dasar. *Pemahaman* (2) mencakup pengetahuan (1). *Aplikasi atau penerapan* (3) mencakup pemahaman (2) dan pengetahuan (1). *Sintesis* (5) meliputi juga analisis (4), aplikasi (3), pemahaman (2) dan pengetahuan (1). *Evaluasi* (6) meliputi juga sintesis (5), analisis (4), aplikasi (3), pemahaman (2) dan pengetahuan (1).

Tujuan aspek kognitif berorientasi pada kemampuan berfikir yang mencakup kemampuan intelektual yang lebih sederhana, yaitu mengingat, sampai pada kemampuan memecahkan masalah

yang menuntut siswa untuk menghubungkan dan menggabungkan beberapa ide, gagasan, metode atau prosedur yang dipelajari untuk memecahkan masalah tersebut. Dengan demikian aspek kognitif adalah subtaksonomi yang mengungkapkan tentang kegiatan mental yang sering berawal dari tingkat pengetahuan sampai ke tingkat yang paling tinggi yaitu evaluasi.

Teori taksonomi pembelajaran yang disampaikan oleh Benjamin S. Bloom (1956) berjudul *Taxonomy of Educational Objectives, Handbook 1: Cognitive Domain*. Beberapa terminologi yang menggambarkan kawasan kognitif adalah sebagai berikut : (1) mendefinisikan istilah teknis dengan memberikan atribut, sifat, atau relasi, (2) kemampuan untuk membedakan referensi untuk kata-kata dan membangun batasan agar istilah biologis memiliki arti (3) keakraban dengan sejumlah besar kata-kata dalam rentangan maknanya, (4) pengetahuan tentang perbendaharaan kata tentang seni yang bias dibaca dan dikonversikan dengan cermat, (5) mengakui pengertian perbendaharaan kata dalam pemikiran kuantitatif, (6) pengetahuan tentang istilah-istilah akuntansi yang penting, (7) penguasaan tentang istilah-istilah untuk bekerja dalam bidang ilmu pengetahuan, (8) memahami pengertian terminologi berkaitan dengan bangun-bangun geometrik.

Kawasan kognitif adalah kawasan yang membahas tentang tujuan pembelajaran yang berkenaan dengan proses mental yang berawal dari tingkat pengetahuan sampai ke tingkat yang lebih tinggi yakni evaluasi. Kawasan kognitif ini terdiri dari 6 tingkatan yang secara hirarkis berurut dari yang paling rendah (pengetahuan) sampai ke yang paling tinggi (evaluasi). Tahapan ketrampilan kognitif sebagaimana Hamzah B. Uno, Herminanto Soyan, dan I Made Candiasa (2001:6-7) menjelaskan yaitu ada enam tingkatan atau tahap bersifat hierarki yaitu :

1. **Pengetahuan (*Knowledge*)**
2. **Pemahaman (*Comprehension*)**
3. **Aplikasi (*Application*)**
4. **Analisis (*Analysis*)**
5. **Sintesis (*Synthesis*)**
6. **Evaluasi (*Evaluation*)**

1. *Tingkat Pengetahuan (knowledge)*

Pengetahuan di sini diartikan sebagai kemampuan seseorang dalam menghafal atau mengingat kembali atau mengulang kembali pengetahuan yang pernah diterimanya.

2. *Tingkat Pemahaman (Comprehension)*

Pemahaman disini diartikan sebagai kemampuan seseorang dalam mengartikan, menafsirkan, menerjemahkan atau menyatakan sesuatu dengan caranya sendiri tentang pengetahuan yang pernah diterimanya.

3. *Tingkat Penerapan (Application)*

Penerapan di sini diartikan sebagai kemampuan seseorang dalam menggunakan pengetahuan dalam memecahkan berbagai masalah yang timbul dalam kehidupan sehari-hari.

4. *Tingkat Analisis (Anaysis)*

Penerapan di sini diartikan sebagai kemampuan seseorang dalam menggunakan pengetahuan dalam memecahkan berbagai masalah yang timbul dalam kehidupan sehari-hari.

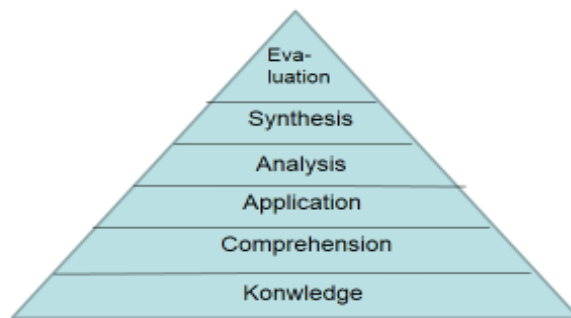
5. *Tingkat Sintesis (Synthesis)*

Sintesis di sini diartikan sebagai kemampuan seseorang dalam mengaitkan dan menyatukan berbagai elemen dan unsur pengetahuan yang ada sehingga terbentuk pola baru yang lebih menyeluruh.

6. *Tingkat Evaluasi (Evaluation)*

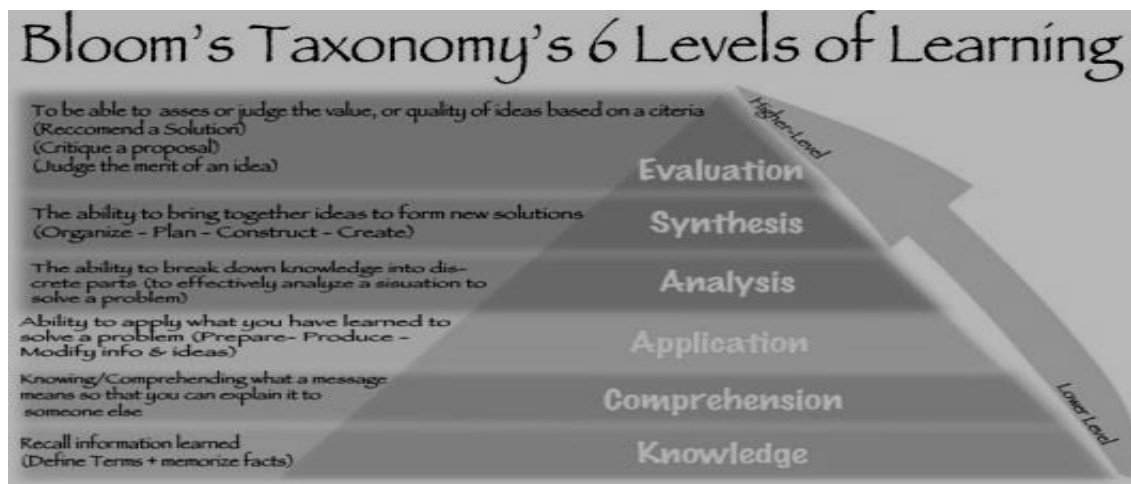
Evaluasi di sini diartikan sebagai kemampuan seseorang dalam membuat perkiraan atau keputusan yang tepat berdasarkan kriteria atau pengetahuan yang dimilikinya.

Untuk memperjelas gambaran yang hierarki pada taksonomi domain kognitif, dapat dilihat pada gambar sebagai berikut :

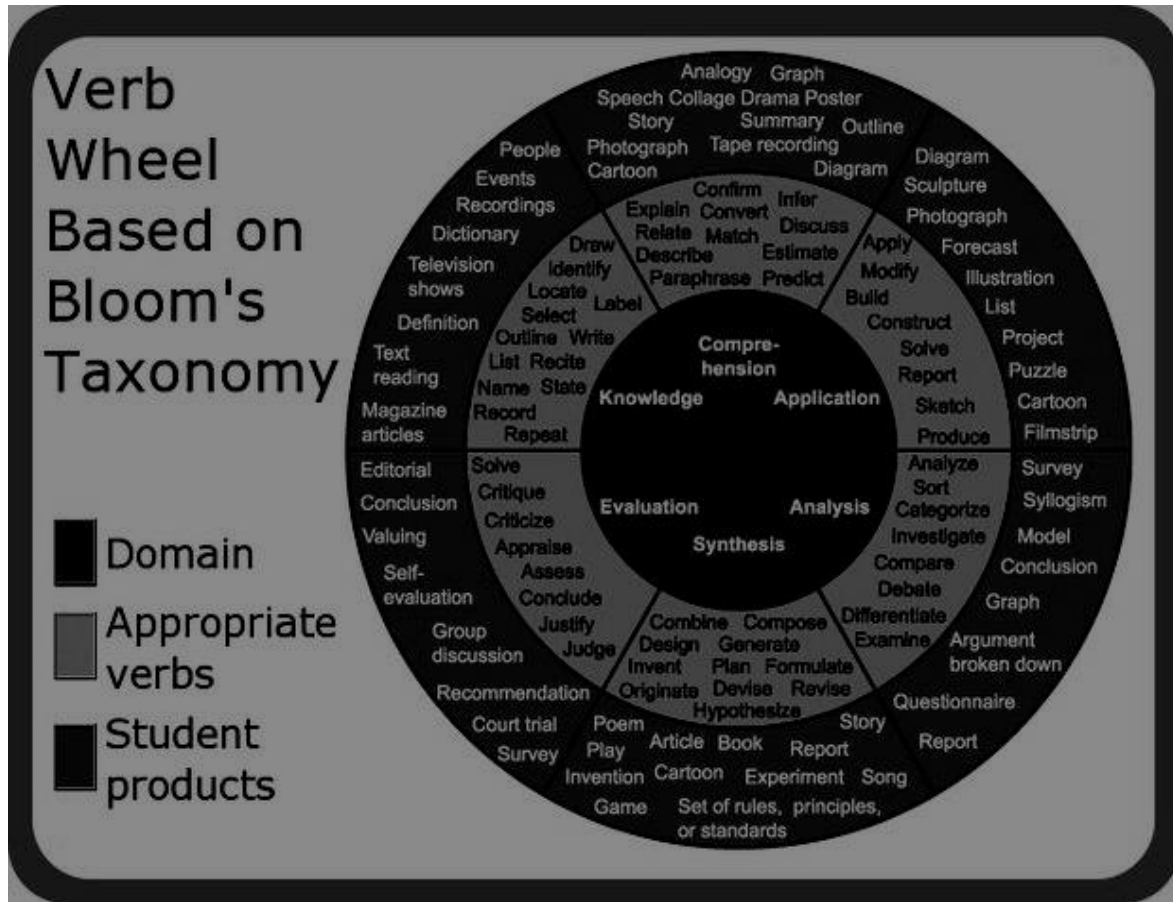


Gambar 4. Hierarki Taksonomi Bloom

Keenam jenjang berpikir yang terdapat pada ranah kognitif menurut Taksonomi Bloom itu, jika diurutkan secara hirarki piramidal adalah sebagai tertulis pada gambar 4.

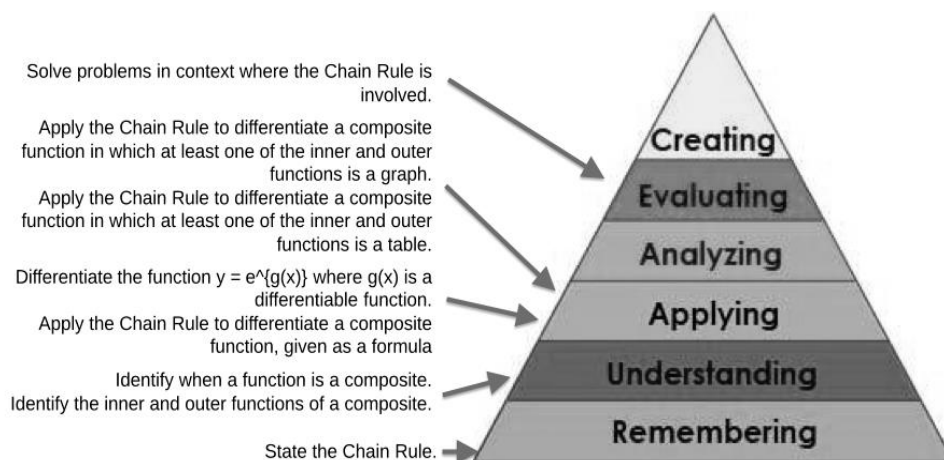


Gambar 5. Enam Level Pembelajaran dalam Taksonomi Bloom

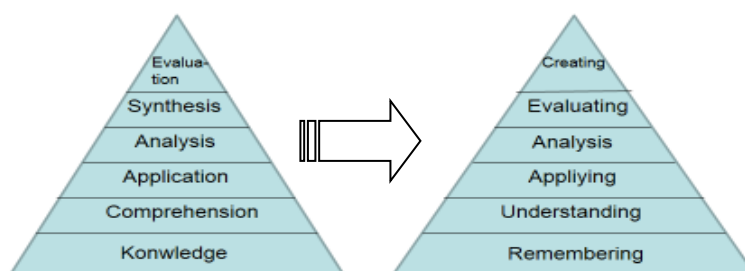


Gambar 6. Kata Operasional Berdasarkan Taksonomi Bloom

Dalam perkembangannya taksonomi Bloom tersebut sudah direvisi oleh Krathwohl (2001). Namun perlu disampaikan bahwa banyak pihak, banyak ahli pendidikan yang belum menerima konsep revisi dari Krathwohl. Sehingga ada dua publikasi dalam menggunakan taksonomi Bloom dan Revisi dari Krathwohl. Sebagaimana pada gambar dibawah ini.

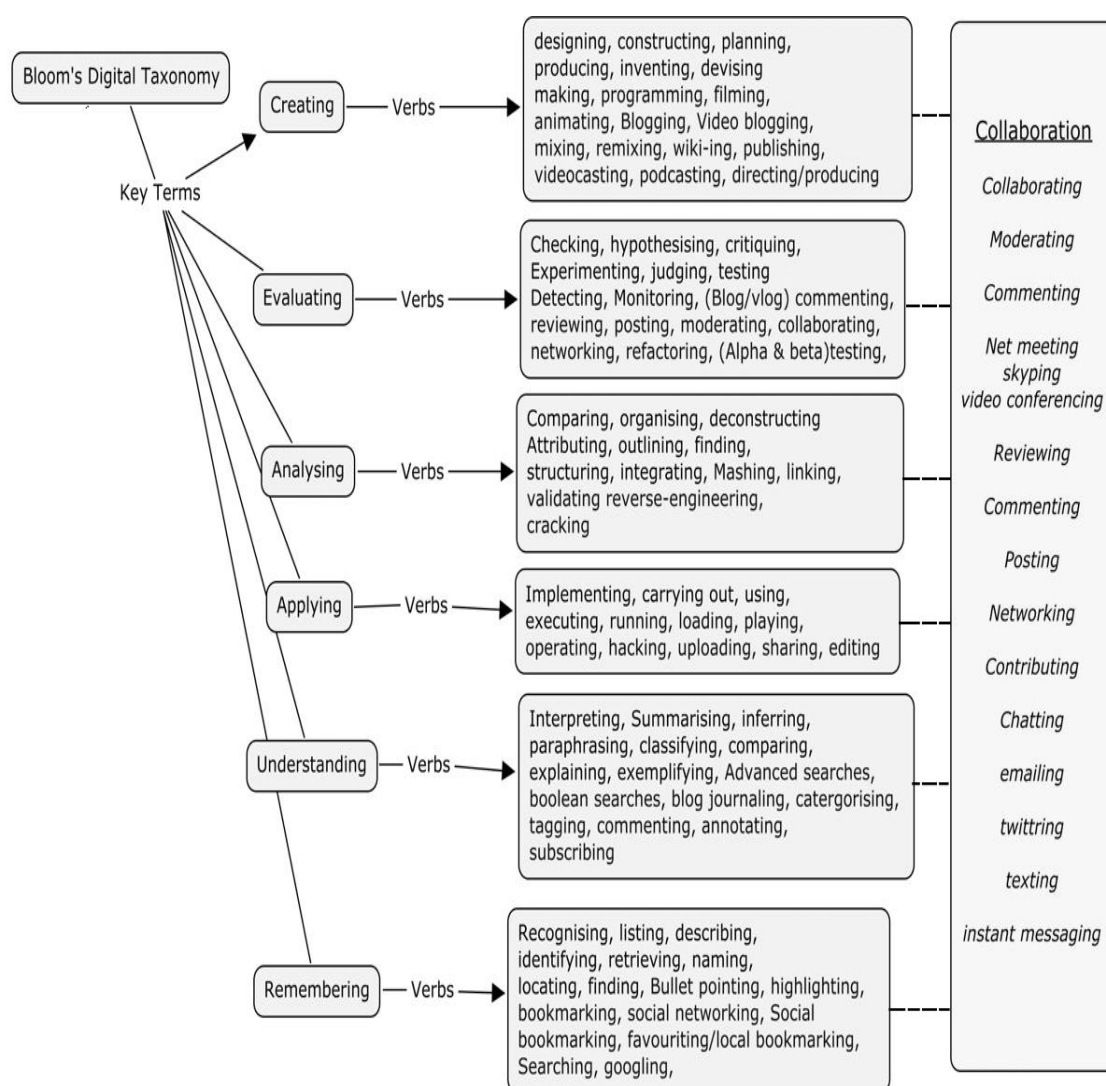


Gambar 7 Hierarki dalam Taksonomi Bloom Versi Revisi



Gambar 8. Perbedaan Taksonomi Bloom Versi Lama dan Versi Revisi

Sedangkan untuk taksonomi Bloom versi revisi dalam domain, kata kerja operasional dan produk-produk siswa, dapat dilihat pada gambar 9 dibawah ini.



Gambar 9. Domain, Kata Operasional yang cocok, dan produk-produk siswa

Contoh Penulisan Tujuan Dengan Kata-Kata Operasional

Kawasan Kognitif

Pengetahuan (C1)

- 1). siswa dapat menyebutkan kembali bangun-bangun geometri yang ber-dimensi tiga
- 2). Siswa dapat menggambarkan satu buah segi tiga sembarang

Pemahaman (C2)

- 1). Siswa dapat menjelaskan dengan kata-katanya sendiri tentang perbedaan bangun-bangun geometri yang berdimensi dua dan berdimensi tiga
- 2). Siswa dapat menerjemahkan arti kode-kode (berita morse) yang dikirim oleh kapal laut yang akan berlabuh

Tingkat Penerapan (C3)

- 1). Siswa dapat menentukan salah satu sudut dari tiga jika diketahui sudut-sudut lainnya
- 2). Siswa dapat menghitung panjang sisi miring dari suatu segi tiga siku-siku jika diketahui sisi lainnya

Tingkat Analisis (C4)

- 1). Siswa dapat mengolah data mentah melalui statistika, sehingga dapat diperoleh harga-harga range, interval kelas, panjang kelas, rata-rata dan standar deviasinya
- 2). Siswa dapat menganalisis sejauh mana dalam dan luasnya pembahasan diskusi yang mereka laksanakan

Tingkat Sintesis (C5)

- 1). Siswa dapat menyusun rencana belajar masing-masing dengan kebijakan yang berlaku di sekolah.
- 2). Siswa dapat mengemukakan formula baru dalam menyelesaikan suatu masalah.

Tingkat Evaluasi (C6)

- 1). Siswa dapat menilai unsur kepadatan isi, cakupan materi, kualitas analisis dan gaya bahasa yang dipakai oleh seprang penulis makalah tertentu
- 2). Siswa dapat menilai kualitas kemampuan pemikiran temannya berdasarkan kemampuan dirinya.

C. Kawasan Afeksi (Sikap dan Perilaku)

- a) Bloom menyatakan bahwa domain afektif, sama halnya dengan domain kognitif, tersusun dalam urutan hirarkis sedemikian sehingga masing-masing kategori perilaku akan diasumsikan merupakan hasil dari kategori perilaku di bawahnya. Akan tetapi tidak tampak bahwa domain afektif didasari oleh prinsip dari sederhana ke kompleks atau prinsip dari konkrit ke abstrak, seperti pada domain kognitif. Beberapa susunan tambahan lain ditemukan di sini.

- b) Analisis tujuan afektif dilakukan untuk menentukan¹ karakteristiknya yang unik. Harapannya adalah untuk menemukan faktor-faktor yang diperlukan agar domain afektif menjadi kontinum. Melalui pengkombinasian dengan struktur domain kognitif, diharapkan struktur domain afektif mulai dengan perilaku sederhana, konkrit, dan tidak terlalu rumit dengan beberapa karakteristik yang belum teridentifikasi. Tingkat perilaku ini akan menjadi bangunan dasar dari perilaku yang kompleks, abstrak, dan rumit dengan jauh lebih banyak karakter yang tidak teridentifikasi. Masalahnya sekarang adalah mendefinisikan perilaku yang belum teridentifikasi dan menemukan prinsip yang menyatakan bahwa perilaku tersebut kontinum.
- c) Kawasan afektif adalah satu domain yang berkaitan dengan sikap, nilai-nilai interest, apresiasi (penghargaan) dan penyesuaian perasaan sosial. Tingkatan afeksi ini ada lima, dari yang paling sederhana ke yang kompleks adalah sebagai berikut :

1) *Kemauan menerima*

Kemauan menerima merupakan keinginan untuk memperhatikan suatu gejala atau rancangan tertentu, seperti keinginan membaca buku, mendengar musik atau bergaul dengan orang yang mempunyai ras berbeda.

2) *Kemauan Menanggapi*

Kemauan menanggapi merupakan kegiatan yang menunjuk pada partisipasi aktif dalam kegiatan tertentu, seperti menyelesaikan tugas terstruktur, menaati peraturan, mengikuti diskusi kelas, menyelesaikan tugas di laboratorium atau menolong orang lain.

3) *Berkeyakinan*

Berkeyakinan dimaksud adalah suatu sikap yang berkenaan dengan kemauan menerima sistem nilai tertentu pada diri individu. Seperti menunjukkan kepercayaan terhadap sesuatu, apresiasi (penghargaan) terhadap sesuatu, sikap ilmiah atau kesungguhan (komitmen) untuk melakukan suatu kehidupan social.

4) *Mengorganisasi*

Pengorganisasian berkenaan dengan penerimaan terhadap berbagai sistem nilai yang berbeda-beda berdasarkan pada suatu sistem nilai yang lebih tinggi. Seperti menyadari pentingnya keselarasan antara hak dan tanggung jawab, bertanggung jawab terhadap hal yang telah dilakukan, memahami dan menerima kelebihan dan kekurangan diri sendiri, atau menyadari peranan perencanaan dalam memecahkan suatu permasalahan.

5) *Tingkat Karakteristik/Pembentukan Pola*

Ini adalah tingkatan afeksi yang tertinggi. Pada taraf ini individu yang sudah memiliki sistem nilai selalu menyelaraskan perilakunya sesuai dengan sistem nilai yang dipegangnya. Seperti bersikap objektif terhadap segala hal.



Gambar 10. Taksonomi Domain Afektif bersifat Hierarki

Kawasan Afektif

Penerimaan (A1)

- 1). Siswa dapat menyatakan setuju terhadap pendapat temannya
- 2). Siswa dapat mengikuti ajakan temannya untuk belajar bersama

Tanggapan (A2)

- 1). Siswa bersedia menyelesaikan tugas terstruktur yang diberikan gurunya
- 2). Siswa sanggup menolong kesulitan yang dialami temannya

Berkeyakinan (A3)

- 1). Siswa memutuskan untuk mengikuti studi tour yang diadakan oleh sekolah
- 2). Siswa mengambil prakarsa untuk membersihkan kelas yang kotor

Pengorganisasian (A4)

- 1). Siswa bersedia melengkapi kekurangan pada tugas yang diberikan guru
- 2). Siswa sanggup menyesuaikan cara belajarnya dengan peraturan yang ada di sekolah

Tingkat karakteristik / pembentuk pola (A5)

- 1). Siswa bersikap sopan santun dalam pergaulan dengan gurunya
- 2). Siswa mempersoalkan nilai yang diberikan oleh gurunya

D. Kawasan Psikomotor

Kawasan Psikomotor

Simson (1966) menyebutkan bahwa domain psikomotor meliputi enam domain mulai dari tingkat yang paling rendah, yaitu persepsi sampai pada tingkat ketrampilan tertinggi, yaitu penyesuaian dan keaslian, meskipun demikian Simson masih mempertanyakan satu tingkat

terakhir yaitu penyesuaian dan keaslian. Oleh karena itu Simson belum memasukkan secara sistematis dalam klasifikasinya. Secara lengkap domain psikomotor adalah :

1. Persepsi

Persepsi berkenaan dengan penggunaan indra dalam melakukan kegiatan. Seperti mengenal kerusakan mesin dari suaranya yang sumbang, atau menghubungkan suara music dengan tari tertentu.

Dimensi dari persepsi adalah :

- Sensori *stimulasi*, adalah sensori yang berkaitan dengan sebuah stimulasi yang berkaitan dengan organ tubuh, yaitu :
 - a) Auditori
 - b) Visual
 - c) Taktile (“ancang-ancang” untuk bertindak)
 - d) Taste (Rasa)
 - e) Smell (Bau)
 - f) Kinestetik
- *Seleksi isyarat*: Menetapkan terhadap isyarat mana orang harus merespon untuk melakukan tugas tertentu dari suatu kinerja. Pemilihan isyarat meliputi identifikasi isyarat dan mengasosiasikannya dengan tugas yang akan dilakukannya. Selain itu pemilihan isyarat juga mencakup pengelompokan isyarat-isyarat dalam bentuk pengalaman dan pengetahuan masa lalu. Isyarat yang relevan dengan situasi dipilih sebagai panduan untuk melakukan gerakan, sedangkan isyarat yang tidak relevan diabaikan atau dihilangkan.
- *Translasi*: Berhubungan dengan persepsi terhadap aksi dalam membentuk gerakan. Ini merupakan proses mental dalam menentukan arti dari isyarat yang diterima dengan aksi. Translasi meliputi translasi simbolik, yaitu memiliki image atau menjadi teringat terhadap sesuatu, memiliki ide, sebagai hasil dari isyarat yang diterima. Translasi juga mencakup insight yang amat esensial dalam pemecahan masalah dengan mencari faktor esensial yang sangat berhubungan dengan penyelesaian. Translasi sensori merupakan satu aspek dari level ini. Translasi sensori meliputi umpan balik, yaitu pengetahuan tentang efek dari suatu proses. Translasi merupakan bagian kontinyu dari gerakan yang sedang dilakukan.

2. Kesiapan

Kesiapan di sini adalah perilaku persiapan atau kesiapan untuk kegiatan atau pengalaman tertentu. Termasuk di dalamnya *mental set* (kesiapan mental), *Physical set* (kesiapan fisik) atau *emotional set* (kesiapan emosi perasaan) untuk melakukan suatu tindakan.

3. Gerakan terbimbing

Gerakan terbimbing gerakan yang berada pada tingkat mengikuti suatu model dan ia lakukan dengan cara meniru model tersebut dengan cara sampai dapat menguasai benar gerakan itu.

4. Gerakan Terbiasa

Gerakan terbiasa adalah berkenaan dengan penampilan respon yang sudah dipelajari dan sudah menjadi kebiasaan, sehingga gerakan yang ditampilkan menunjukkan suatu kemahiran. Seperti menulis halus, menari atau mengatur/menata laboratorium.

5. Gerakan yang komplek

Gerakan yang komplek adalah suatu gerakan yang berada pada tingkat yang terampil yang tinggi. Ia dapat menampilkan suatu tindakan motorik menurut pola tertentu dengan tingkat kecermatan dan atau keluwesan serta efisiensi yang tinggi.

6. Penyesuaian dan Keaslian

Pada tingkat ini individu sudah berada pada tingkat yang terampil sehingga ia sudah dapat menyesuaikan tindakannya untuk situasi-situasi yang menurut persyaratan tertentu. Individu sudah dapat mengembangkan tindakan/ketrampilan baru untuk memecahkan masalah-masalah tertentu.

Kawasan Psikomotor

Persepsi (P1)

- 1). Siswa dapat menunjukkan obeng pipih dengan tepat setelah melihat demonstrasi guru praktek
- 2). Siswa dapat memilih baju yang pantas untuk dirinya

Kesiapan (P2)

- 1). Siswa dapat menunjukkan cara menggunakan keyboard komputer dengan tepat
- 2). Siswa mampu menanggapi pesan kesalahan di layar komputer

Respon terbimbing (P3)

- 1). Siswa dapat melepas disket dari diskdrive dengan tepat setelah melihat modul
- 2). Siswa dapat memasang kembali disket kembali pada diskdrive setelah mendapat petunjuk instruktur

Gerakan Mekanisme (P4)

- 1). Siswa dapat membongkar karburator sepeda motor dengan tepat
- 2). Siswa dapat memasang kembali karburator sepeda motor pada dudukannya dengan tepat.

Respon yang kompleks (P5)

- 1). Siswa dengan lancar dapat mendemonstrasikan cara mengatur format keterampilan di layar komputer
- 2). Siswa dapat melakukan gerakan dasar tari dengan luwes.

Penyesuaian dan keaslian (P6)

- 1). Siswa dapat membuat variasi tampilan storyboard pada komputer.
- 2). Siswa dapat mengatur kembali posisi buku di perpustakaan berdasarkan kartu katalog.

Taksonomi Merrill

Di atas telah diuraikan bahwa domain taksonomi Bloom yang terdiri atas enam tingkatan, untuk mengukur kemampuan siswa. Taksonomi lain yang tidak kalah pentingnya digunakan untuk mengukur kemampuan seseorang dalam belajar adalah taksonomi Merrill.

Robert Gagne (1965) mempostulatkan bahwa tipe berbeda dari domain hasil belajar menghendaki kondisi belajar yang berbeda. Component Display Theory (CDT) dilibatkan dalam usaha untuk mengklarifikasikan teori Gagne. CDT dibangun diatas asumsi yang sama dengan yang dilakukan Gagne, yaitu terdapat kategori outcome yang berbeda, dan masing-masing kategori itu menghendaki prosedur yang berbeda untuk mencapai prestasi dan prosedur yang berbeda untuk mencapai prestasi dan prosedur yang berbeda untuk mencapai prestasi dan prosedur yang berbeda untuk mempromosikan kemampuan yang direpresentasikan oleh kategori. Akan tetapi system klasifikasi satu dimensi yang aslinya diusulkan oleh Gagne (1965) tampaknya sangat terbatas, oleh karena itu system dua dimensi yang aslinya diusulkan oleh Merrill (1973), dimana kinerja (dimensi dari Gagne) sebagai satu dimensi dan isi (content) sebagai dimensi yang lain. Taksonomi tersebut oleh Merrill dinamakan Component Display Theory (CDT). CDT menyatakan bahwa hasil belajar dapat direpresentasikan dengan tujuan belajar atau dengan item tes. Diasumsikan bahwa pada setiap kategori, yaitu kinerja dan isi terdapat kombinasi dari bentuk Presentasi primer dan sekunder, yang akan menghasilkan pemahaman yang lebih efektif dan efisien .

Kategori Hasil Belajar

Merrill seperti sudah disebutkan , melalui CDT memilah taksonomi hasil belajar dalam dua dimensi , yaitu : (1) kinerja, yang terdiri dari mengingat , menggunakan , dan menemukan.; dan (2) isi pelajaran , yang terdiri dari fakta ,konsep, prosedur , dan prinsip. CDT hanya cocok untuk hasil belajar kognitif, dan tidak mencakup tujuan afektif atau psikomotor.

Kategori Isi

Fakta:

Fakta adalah asosiasi satu-satu antara objek, peristiwa, atau simbol yang ada , atau mungkin ada, di dalam lingkungan riil, atau imajinasi .Misalnya, itik (Fakta riil), Bilangan Satu(Fakta Imajinasi). Beberapa contoh item tes yang menyangkut fakta adalah sebagai berikut.

- Simbol tahanan adalah ...
- Sebutkan bagian-bagian utama dari mata
- Siapakah presiden Indonesia sekarang?
- Berapakah nilai , bilang adalah konstanta gravitasi

Konsep :

Konsep adalah sekelompok objek, peristiwa, atau simbol yang memiliki karakteristik yang sama yang dapat diidentifikasi dengan nama yang sama. Misalnya, Itik, Merpati, Ayam (konsep Burung). Beberapa contoh item tes yang menyangkut konsep adalah sebagai berikut.

- Karakteristik apa yang membedakan lukisan impresionis dan lukisan renaissance ?
- Mana dari foto-foto yang menunjukkan awan kumulus ?
- Apakah analisis data yang di berikan berikut memerlukan analisis statistik satu ekor atau dua ekor ?

Prinsip:

Prinsip adalah hubungan sebab akibat dari konsep-konsep . Misalnya, prinsip penawaran dan prinsip permintaan dalam ekonomi, atau prinsip penggunaan media. Beberapa contoh item tes yang menyangkut fakta adalah sebagai berikut.

- Pada sirkuit berikut jelaskan apa yang akan terjadi pada intensitas cahaya, dengan menjelaskan aliran electron dalam sirkuit
- Pada kasus berikut, jelaskan perilaku anti social dari subyek menurut teori *reinforcement*
- Jelaskan secara singkat hukum pertama dari gerakan...

Prosedur:

Prosedur adalah urutan langkah untuk mencapai satu tujuan , memecahkan masalah tertentu, atau membuat sesuatu. Misalnya, prosedur pengolahan data, prosedur penelitian: Contoh item tes yang menyangkut fakta adalah sebagai berikut.

- Langkah-langkah apakah yang diperlukan untuk menemukan arus dalam sirkuit DC
- Gunakan kalkulator tangan untuk menemukan mean dan standar deviasi dari sejumlah bilangan berikut:
5 6 7 8 7 7 6 7 6 6 8 9 6 7
5 4 5 5 6 7 6 6 6 7 6 7 8 5
- Selesaikan persamaan linier berikut :
 $X^2 + 2x + 2 = 0$

Kategori Kinerja

Mengingat:

Mengingat, merupakan unjuk kerja yang menuntut siswa melakukan penelusuran struktur ingatan agar dapat mengungkapkan kembali konstruk-konstruk yang telah di simpan di dalamnya.

Contoh item tes:

Simbol dari tahanan adalah

Menggunakan:

Menggunakan merupakan unjuk kerja yang menuntut siswa menerapkan suatu abstraksi pada kasus-kasus khusus.

Contoh item tes:

Apa yang akan terjadi apabila dalam sirkuit di bawah ini tahanan yang terpasang di buat hubungan pendek? (gambar sirkuit disajikan)

Menemukan:

Menemukan, merupakan unjuk kerja yang menuntut siswa untuk dapat menemukan atau mengembangkan suatu abstraksi yang baru.

Contoh item tes:

Ciptakan satu sirkuit sederhana yang secara perlahan-lahan memperlambat motor arus searah sampai berhenti

Contoh Item Tes Gabungan

Pada masing-masing dimensi telah diberikan contoh item tes. *Component Display Theory (CDT)* juga memungkinkan tujuan atau item tes ke dalam lebih dari satu sel matrik kinerja dan isi . Berikut ini adalah contoh item tes yang menggambarkan klasifikasi dua dimensi.

Mengingat-fakta

- Dalam peta topografi, apakah symbol dari masjid?
- Apa peran nilai dalam kehidupan?
Fakta tidak memiliki representasi abstrak atau umum, sehingga tidak ada klasifikasi menggunakan fakta atau menemukan fakta.

Mengingat- Konsep

- Apakah karakteristik dari mamalia
- Definisikan apa itu penguatan positif

Menggunakan – Konsep

- Apakah binatang yang tergambar dalam foto ini merupakan contoh dari mamalia
- Baca cerita pendek berikut. Identifikasikan paragraf yang terbaik dalam menggambarkan klimaks dari cerita.

Menemukan – Konsep

- Urutkan batu-batu yang di sediakan menjadi beberapa tumpukan yang berbeda . Tunjukan karakteristik , dengan cara apa teman sekelas anda dapat mengurutkan batu menjadi urutan yang sama.
- Gambarkan cara untuk mengelompokkan siswa di dalam kelas yang menggambarkan rentangan kemampuan dan jenis kelamin pada masing-masing kelompok.

Mengingat- Prosedur

- Bagaimana langkah untuk menyeimbangkan buku tas?
- Jelaskan langkah-langkah untuk membuat cetakan berwarna pada printer berwarna

Menggunakan – Prosedur

- Demonstrasikan cara membersihkan clarinet
- Buat lidah pencabangan pada pohon buah

Menemukan- Prosedur

- Buat program computer untuk mengindeks dan memanggil resep makanan
- Ciptakan teknik untuk memberikan tugas secara acak kepada siswa

Mengingat- Prinsip

- Jelaskan masing-masing tiga teknik proyeksi untuk membuat peta permukaan bumi
- Apa yang terjadi saat air menguap ? Jelaskan dari lingkup gerakan molekul dan pemanasan

Menggunakan- Prinsip

- Berikut ini adalah gambar dari dua perahu.satu mengapung sangat tinggi sedangkan yang satu lagi mengapung sangat rendah. Jelaskan minimum tiga alasan yang bias menjelaskan perbedaan itu.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

- Baca studi kasus berikut tentang system ekologi. Pada system ini jumlah hewan pengerat bertambah. Prediksikan beberapa hipotesis berdasarkan pengetahuan anda tentang siklus kehidupan dan interdependensi dari spesies-spesies dalam system ini.

Menemukan- Prinsip

- Lakukan satu eksperimen untuk menguji pengaruh asap rokok terhadap pertumbuhan tanaman . Laporkan penemuan anda.
- Lakukan demonstrasi yang memungkinkan anda untuk mengilustrasikan bagaimana air didapatkan dari sumur.

**BAB
III**

PENGEMBANGAN INSTRUMEN PENGUKURAN

A. Pengantar

Proses pengukuran dalam riset kuantitatif adalah hal yang sangat esensial. Dari sanalah angka didapatkan untuk diolah dan dianalisis sehingga mendapatkan jawaban untuk pertanyaan risetnya. Tanpa pengukuran maka tidak akan ada data didapatkan dan riset tidak akan berlanjut (kecuali mengubah pendekatan menjadi riset kuantitatif). Proses pengukuran tidak lain adalah proses memasukan angka pada konsep sesuai dengan indikator yang telah dibuat ketika kontinum sudah ditetapkan. Biasanya. Kita mengatakan bias “mengukur” sesuatu yang tidak terlihat berdasarkan kesimpulan dari apa yang bisa diamati dengan mata secara langsung.

Dalam ilmu pengetahuan sosial seperti psikologi dan pendidikan, objek yang diukur biasanya objek tidak kasatmata., atau sesuatu yang tersembunyi (laten). Contohnya, dalam bidang pendidikan, atribut seperti kecerdasan, tidaklah dapat diukur dengan hanya mengandalkan pengamatan tampilan luar. Dengan kata lain, mengukur kecerdasan tidak bias dilakukan dengan mengidentifikasi melalui satu indikator saja (penampilan), tetapi dari berbagai indikator. Untuk mengukur kecerdasan diperlukan sebuah alat ukur yang bisa mengungkap berbagai indikator kecerdasan serta menempatkannya dalam rentang yang disepakati.

Dalam pemodelan Rasch, ide dasar dari proses pengukuran ini disebut dengan ciri-sifat laten (latent traits) yang tidak lain merupakan konsep utama dari teori respons butir (IRT). Penjelasannya yakni walaupun ciri-sifat itu tidak bisa diamati secara empirik, manifestasi ciri-sifat tersebut dan interaksinya dengan lingkungan akan memunculkan suatu indikator empirik yang bisa diukur atau diobservasi. Dengan demikian, bisa digunakan untuk menyimpulkan keberadaan atau tidakberadaan ciri-sifat tadi, atau menempatkannya dalam satu kontinum. Dalam kerangka inilah instrument pengukuran dibuat menggunakan berbagai butir (item atau aitem) sebagai dasar untuk membuat simpulan tentang ciri-sifat laten tersebut.

Dalam konsep pengembangan instrument pengukuran psikologi, konstruk (atribut) yang terkait dengan manusia memiliki banyak karakteristik, salah satunya adalah konstruk yang bersifat laten. Artinya, atribut tersebut bersifat hipotetik dan tidak dapat diamati secara langsung. Anda tidak bisa mengetahui agresivitas secara langsung. Contohnya, Si A agresif, tetapi si B tidak. Anda baru mengetahui apakah si A atau tidak melalui seperangkat indikator perilaku yang merupakan manifestasi dari agresivitas. Contohnya, menendang, memukul, menghujat, dan melukai.

B. Tahap-tahap Pengembangan Alat Ukur

Untuk memudahkan penjelasan ini, Anda dapat mengambil contoh tentang konsep harga diri (*self esteem*). Walaupun ciri-ciri akan konsep harga diri ini mungkin bisa berdasarkan pengamatan, hasilnya bersifat subyektif dan kualitatif. Dalam konteks penelitian kuantitatif, hal ini haruslah terukur dan ditentukan dalam kontinum yang disepakati.

Konseptualisasi Konstruk Ukur. Hal pertama dalam proses pengukuran untuk konsep harga diri ini adalah Anda harus mendefinisikannya dan menjabarkannya ke dalam bentuk indikator yang operasional. Contohnya, Anda menggunakan pengertian harga diri berdasarkan teori dari

Rosenberg. Harga diri didefinisikan sebagai komponen evaluatif dari konsep diri (*self concept*) Harga diri merupakan suatu penilaian umum seseorang terhadap dirinya, sejauh mana kompetensi yang dimiliki serta seberapa jauh penghargaan terhadap keberadaan dirinya. Seseorang yang memiliki harga diri yang tinggi menganggap dirinya sangat berharga, namun bukanlah sebagai pribadi yang sombong, congkak, merasa hebat, atau menghina orang lain. Sebaliknya, seseorang dengan harga diri yang rendah adalah seseorang yang kurang menghargai diri sendiri dan kemampuannya—menganggap dirinya tidak berharga dan pribadi yang banyak memiliki kekurangan.

Beberapa poin yang perlu diperhatikan pada tahap ini adalah sebagai berikut :

- Selain dapat didasarkan pada teori tunggal, alat ukur dapat juga dapat dikembangkan berdasarkan gabungan beberapa teori.
- Untuk memudahkan penulisan butir, banyak peneliti yang memecah konsep teoretis menjadi beberapa indikator berperilaku yang bisa dinamakan dengan aspek, komponen, atau domain ukur. Contohnya, domain-domain komunikasi efektif adalah (a) kejelasan ide yang disampaikan, (b) kelancaran dalam menyapaikan ide, dan (c) ketepatan ide penyampaian dengan kondisi audiens.

1. Perbedaan Pengertian Aspek dan Dimensi dalam Pengembangan Alat Ukur

Dalam proses penyusunan alat ukur, sebelum menulis aitem, peneliti biasanya mengeksplorasi teori mengenai konstruk yang hendak diukur (pengertian dan aspek-aspek-nya). Penentuan butir berdasarkan teori adalah karakteristik dari pendekatan penyusunan secara deduktif (*top down*), penyusunan alat ukur disetir oleh teori. Ada juga penyusunan skala yang disetir oleh penemuan-penemuan hasil eksplorasi di lapangan. Namanya pendekatan induktif (*bottom up*).

1.1. Aspek

Aspek adalah bagian konstruk ukur yang lebih operasional sebelum dijabarkan lagi menjadi indikator-indikator perilaku. Kebanyakan peneliti menamakan penjabaran ini sebagai aspek. Contohnya, Betancourt dan Lopez (1993) mengembangkan pengukuran tentang akulturasi mengawali dengan menjabarkan teori menjadi *behavioral aspect of acculturation*. Aspeknya adalah *food preference, friendship patterns, and language usage*. Aspek-aspek inilah yang kemudian dijabarkan menjadi butir-butir di dalam skala. Jika peneliti tidak mendapatkan informasi mengenai bobot aspek-aspek tersebut., diharapkan jumlah tiap butir dalam tiap aspek adalah sama. Beberapa ahli lainnya menyatakan bahwa penyusunan skala dapat langsung dari definisi teoritis yang didapatkan dari kajian literatur secara komprehensif. Dalam pendekatan ini, penjabaran teori menjadi aspek-aspek tidak diperlukan.

1.2. Dimensi

Dimensi menjelaskan komponen-komponen ukur unik yang tidak memiliki korelasi antara satu dan lainnya. Untuk mengetahui apakah sebuah komponen tersebut merupakan dimensi, dapat dilihat melalui dua cara : dari teori dan dari analisis faktor. Analisis faktor akan menghasilkan faktor-faktor (kadang dinamakan juga dengan dimensi). Alat ukur yang memiliki satu dimensi dinamakan dengan unidimensi, sedangkan yang memiliki banyak dimensi dinamakan dengan multidimensi. Banyak alat ukur yang bersifat multidimensi. Contohnya, pengukuran kecemasan (*multidimensional scale of anxiety*) dari Bystritsky et. Al., (1990). Skala tersebut memiliki dimensi lebih dari satu sehingga dinamakan dengan skala multidimensi. Pengukuran kepribadian yang menggunakan teori kepribadian lima faktor, misalnya NEO-FFI juga merupakan pengukuran

multidimensi. Pada banyak alat ukur, dimensi-dimensi di dalamnya kemudian dijadikan menjadi satu sub-skala atau sub-tes.

Kesimpulan, gunakan saja pengertian bahwa aspek adalah domain-domain ukur teoritis yang belum diuji apakah memiliki idependensi ataukah tidak masing-masing. Jika tiap-tiap domain teoretis tersebut terbukti memiliki idependensi yang dibuktikan dengan analisis faktor seperti yang dilakukan oleh Bystritsky et. Al., (1990).ia dinamakan dimensi atau faktor.

2. Penulisan Butir

Tahap berikutnya setelah mendefinisikan konsep harga diri, Anda menyeleksi indikator apa yang bisa memberikan representasi empiris terhadap domain-domain ukur yang telah ditetapkan sebelumnya. Menyusun indikator harus mempertimbangkan teori yang diacu. Butir-butir ditulis berdasarkan indikator keprilakuan yang telah didapatkan dari teori. Contohnya, untuk salah satu indikator orang yang memiliki harga diri tinggi adalah dia menilai dirinya memiliki kelebihan yang unik. Dari indikator ini, kita menuliskan butir “Saya memiliki kelebihan yang saya banggakan”. Beberapa butir dapat ditulis dalam bentuk pertanyaan positif (*favorable*) tentang diri sendiri, seperti “Secara keseluruhan, saya puas dengan diri saya sendiri“, sedangkan indikator lain berisi pertanyaan yang bersifat negative (*unfavorable*), “Saya merasa tidak mempunyai sesuatu yang bisa dibanggakan”. Penskoran butir yang berarah negatif ini adalah berkebalikan dari butir yang berarah positif. Pada butir berarah negatif respons persetujuan (contohnya “setuju”) mendapatkan skor yang lebih rendah dibandingkan respons ketidaksetujuannya (contohnya “tidak Setuju”).

3. Jumlah Butir dalam Alat Ukur

Tidak ada aturan baku mengenai beberapa butir yang dilibatkan dalam skala. Walaupun demikian, ada beberapa hal yang dapat dipakai sebagai pertimbangan.

- a) **Tujuan pengukuran dan jenis keputusan yang akan dibuat.** Instrumen yang dipakai untuk menyeleksi memiliki butir lebih banyak dibandingkan untuk survei. Instrumen yang dipakai untuk membuat klasifikasi pasien lebih mendetail memerlukan lebih banyak butir dibandingkan dengan yang dipakai untuk mendeskripsikan saja. Semakin berisiko keputusan yang diambil berdasarkan pengukuran, semakin banyak butir yang perlu dilibatkan. Tes yang dipakai untuk menyeleksi (high stake testing) memerlukan butir yang lebih banyak.
- b) **Keluasan Domain ukur dan Dimensionalitas Pengukuran.** Domain yang luas (*wide domain*) dan pengukuran dengan dimensi majemuk (*multiple dimension*) memerlukan jumlah butir yang dibandingkan dengan domainnya sempit berdimensi tunggal.
- c) **Karakteristik Responden.** Kelelahan Responden dan kemauan untuk bekerja sama juga perlu dipertimbangkan. Terkait dengan kebosanan dan kelelahan responden, Swanson dan Holton (2009) menyarankan agar kita menggunakan pengukuran pendek, tetapi efektif sehingga kebosanan dan kelelahan tersebut dapat diatasi.

Empat hingga enam butir cukup untuk mengukur satu domain ukur. Biasanya, satu alat ukur memiliki 3-5 domain sehingga rata-rata jumlah butir yang dibutuhkan adalah sekitar 15 hingga 20 butir. Untuk mengantisipasi butir yang gugur maka jumlah butir yang diuji cobakan adalah 30 hingga 40 butir. Menurut Swanson dan Holton (2009), Jumlah ini dapat dipakai untuk mengukur konstruk apa saja yang spesifik. Neill (2011) menulis bahwa dalam konteks analisis factor terhadap pengukuran yang memiliki berbagai domain, setidaknya satu domain minimal memuat 3 butir. Namun, secara umum biasanya jumlah butir dalam satu factor adalah 4 hingga 10

butir. Dalam skala Big Five Inventory (BFI) misalnya, jumlah butir dalam satu domain adalah antara 8 hingga 10 butir.

Beberapa poin yang perlu diperhatikan pada tahap penulisan butir sebagai berikut.

- Jumlah butir yang dibuat oleh peneliti (item pool) harus dua-tiga kali lebih banyak daripada jumlah target butir. Tujuannya, ketika banyak butir yang tidak lolos dalam seleksi butir, adalah 10 maka peneliti dapat menyiapkan 40 butir untuk diujicobakan.
- Ketika peneliti akan menggunakan analisis butir dengan teori kasik, beberapa peneliti menyarankan untuk melibatkan butir yang memiliki kemiripan (DeVellis, 2011). Contohnya, butir “Saya memiliki wajah yang tampan” dan butir “ Saya memiliki wajah yang rupawan“. Tujuannya adalah untuk mencapai konsisten internal. Ketika analisis yang dilakukan berbasis analisis faktor (CFA) atau IRT/Rasch maka kemiripan tersebut tidak seberapa dipentingkan. Bahkan, Pemodelan Rasch menghendaki butir-butir yang memiliki intensitas atau tingkat kesulitan yang berbeda-beda agar menjangkau berbagi tingkat atribut yang diukur (“stepping stones” of item difficulty).

4. Uji Coba Alat Ukur

Tahap selanjutnya adalah mendapatkan informasi secara empiris dengan memberikannya kepada partisipan agar memberikan jawaban berdasarkan deretan pertanyaan indikator tadi. Peneliti bisa merancang bahwa pernyataan negatif dan positif untuk konsep ini sama banyak dan meminta partisipan untuk menilai kesesuaiannya. Namun, sebelum butir tersebut diujicobakan, ada dua tahap yang perlu dilalui. **Pertama**, mengundang pakar atau ahli di bidang atribut yang kita ukur untuk mengevaluasi butir pernyataan yang kita tulis. Contohnya, pengukuran tentang kepuasan pelanggan dapat mengundang beberapa ahli pemasaran untuk menilai kesesuaian pernyataan kita dengan teori dan target populasi yang diukur. Prosedur ini merupakan bagian dari upaya mendapatkan validitas isi. **Kedua**, melakukan uji coba kecil-kecilan (mini pilot test) dengan mengundang beberapa subyek (5-10 orang) untuk menilai seberapa mereka memahami butir pernyataan yang tertulis. Kadang apa yang kita maknai pada suatu butir dimaknai beberapa oleh responden kita. Pada tahap ini, tidak ada proses penskoran karena tujuannya adalah mendapatkan masukan tentang pernyataan di dalam butir. Beberapa hal yang perlu diperhatikan pada tahap ini sebagai berikut.

- Subyek uji coba adalah sampel dari populasi ukur. Jika Anda menembangkan alat ukur untuk remaja., subjek uji coba adalah remaja. Subjek uji coba tidak harus benar-benareksak mewakili target yang akan diteliti. Contohnya, Anda mengembangkan alat ukur untu dipakai pada penelitian di perusahaan A. Asalkan butir-butir di dalam alat ukur tidak spesifik mengukur kondisi di perusahaan A maka peneliti dapat mengujicobakan di tempat lain yang memiliki karakteristik sama dengan karyawan di perusahaan A.
- Variasi subjek uji coba perlu dipertimbangkan. Contohnya, kita mengembangkan pengukuran religiositas remaja, kita libatkan remaja dari berbagi kalangan yang secara teoretis memiliki religiositas heterogen. Uji coba dapat dilakukan pada satu kalangan saja, tetapi secara teoretis memiliki religiositas yang heterogen. Uji coba pada sampel yang bervariasi akan menghasilkan skor yang bervariasi pula. Skor yang bervariasi akan memperkaya informasi mengenai butir yang diuji.

5. Analisis Butir

Analisis butir dilakukan terhadap data hasil uji coba. Properti psikometris setiap butir dikaji sesuai dengan pendekatan analisis yang dipakai. Hasil dari uji coba dapat memberikan masukan yang berharga untuk merevisi butir yang diujicobakan karena berbagi pernyataan yang dibuat sudah diuji secara empiris. Bagaimana membuatnya menghasilkan angka sebagaimana tuntutan riset kuantitatif? Untuk itu, langkah berikutnya melakukan penskoran.

Teknik pengukuran yang paling terkenal adalah Thurstone, Guttman, dan Likert (akan dijelaskan kemudian). Tiap-tiap model mempunyai cara tersendiri untuk mengubah jawaban responden yang nantinya akan dikonversi menjadi angka yang dinamakan dengan skor. **Skor** adalah harga yang kita berikan terhadap respons atau jawaban yang diberikan oleh responden. Ada bagaimana macam cara penskoran. Jawaban benar dalam soal matematika dapat kita hargai angka 1 atau 2, sedangkan jawaban salah satu dapat kita hargai dengan angka 0 atau -1.

Dalam pengukuran sikap atau kepribadian, yang paling populer adalah pemeringkatan Likert; dalam hal ini respons terhadap pernyataan diberikan pilihan jawaban, seperti sangat setuju (STS), tidak setuju (TS), netral (N), setuju (S), dan sangat setuju (SS). Kemudian, jawaban nanti dikonversi menjadi angka yang sesuai dengan jenis pernyataan yang diberikan.

Setelah proses penskoran, langkah selanjutnya adalah melakukan analisis butir. Pada pengukuran atribut maksimal (*maximal performance*) (contoh : tes prestasi kemampuan matematika), kita mengestimasi tingkat kesulitan, daya diskriminasi, dan efektivitas distraktor. Tingkat kesulitan menunjukkan seberapa sulit butir tersebut dihadapan subjek, daya diskriminasi menunjukkan seberapa jauh butir mampu membedakan subjek yang memiliki kemampuan tinggi dan rendah, sedangkan efektivitas distraktor menunjukkan seberapa banyak subjek memilih opsi distraktor. Pada tes pilihan ganda, opsi distraktor adalah opsi yang disediakan selain opsi yang merupakan kunci jawaban. Opsi distraktor yang tidak banyak dipilih oleh subjek direkomendasikan untuk diganti.

Pada pengukuran performansi tipikal (*typical Performance*), properti psikometris butir yang dikaji adalah daya diskriminasi butir. Tingkat kesulitan dan efektivitas distraktor tidak dibatalkan. Properti psikometris yang dijelaskan sebelumnya adalah properti psikometris berbasis teori klasik. Analisis dengan menggunakan pendekatan teori modern (seperti IRT.Rasch) juga mengkaji properti yang sama dengan teori klasik, namun ada beberapa kriteria tambahan yang perlu dipenuhi. Beberapa hal yang perlu dipertimbangkan pada tahap ini antara lain sebagai berikut.

- Tidak semua subjek memberikan respons yang valid karena faktor kelelahan, mengantuk, kurang motivasi, atau mengidentifikasi subjek tersebut dan mengeluarkannya dari analisis jika memberikan dampak yang substansial terhadap properti psikometris butir (Harnisch & Linn, 1981). Beberapa koefisien, misalnya indeks ketepatan personal (*person fit*) dapat mengidentifikasi subyek yang berperilaku menyimpang (*aberrant*). Indeks ketepatan personal (*person fit*) akan dibahas pada bagian lain di buku ini.
- Dimensionalitas pengukuran. Analisis butir hendaknya dilakukan pada data yang telah memenuhi asumsi unidimensionalitas (Netemeyer, Bearden & Sharma, 2003). Analisis faktor dapat menunjukkan apakah data pengukuran yang kita analisis bersifat unidimensi ataukah multidimensi. Jika data kita bersifat multidimensi, analisis butir perlu dilakukan pada tiap dimensi. Asumsi ini juga harus dipenuhi jika menggunakan analisis berbasis teori modern (IRT/Rasch). Pada pendekatan ini selain unidimensionalitas pengukuran, asumsi independensi

local juga harus dipenuhi. Beberapa dengan teknik yang lain, pemodelan Rasch tidak membutuhkan asumsi unidimensional yang sangat ketat.

6. Merakit Alat Ukur

Proses analisis butir adalah proses yang dilakukan secara berputar dan berulang (iterative) hingga mendapatkan alat ukur yang valid dan reliable. Ketika butir-butir yang tidak memenuhi kriteria telah digugurkan, property psikometris alat ukur secara keseluruhan (reliabilitas) akan mengalami peningkatan. Namun, validitas juga perlu diperhatikan, karena terkadang realibilitas yang tinggi tidak secara otomatis menghasilkan validitas yang tinggi pula. Bisa jadi 40 butir yang diuji, Anda dapatkan 20 butir yang keseluruhannya menghasilkan realibilitas yang tinggi dengan validitas rendah. Butir-butir yang tersisa semuanya tidak mewakili domain ukur secara lengkap sehingga dapat dikatakan validitas isinya tidak terpenuhi. Oleh karena itu, proses analisis butir dilakukan lagi hingga semua kriteria yang ditetapkan telah terpenuhi. Proses yang dilakukan pada tahap ini adalah proses perakitan alat ukur.

Pada analisis butir yang menggunakan pendekatan IRT/Rasch, proses perakitan dilakukan dengan menggunakan kriteria berupa indeks fungsi informasi (*information function*) yang nilainya setara dengan realibilitas. Pada tahap ini, kita memilih butir-butir yang telah memiliki property yang baik (*fit* dengan model) untuk dipilih lagi dengan mempertimbangkan nilai fungsi informasi.

Oleh karena itu, proses pengukuran ciri-sifat laten ini paling tidak ada tiga hal yang harus terpenuhi. *Pertama*, adalah definisi yang dirujuk akan konsep/ ciri-sifat dan memberikan spesifikasi berbagai indikator yang dalam hal ini Anda akan membuat simpulan bagaimana hubungan indikator tadi dengan ciri-sifat (ini juga berhubungan dengan validitas yang akan dibahas pada bagian berikutnya). Kedua, Anda menggunakan berbagai indikator atau butir (item), dengan alasan makin banyak indikator maka Anda mempunyai simpulan yang lebih baik yang bias didapatkan. Tentu saja Anda memerlukan untuk mengumpulkan respons secara keseluruhan dari berbagai butir tadi untuk memperkuat simpulan yang akan diambil. Ketiga, butir yang setara bagus dalam menjelaskan indikator untuk ciri-sifat tadi.

C. Teknik Pengukuran

Dalam riset kuantitatif teknik pengukuran adalah hal krusial lainnya karena hal ini menentukan bagaimana respons yang didapat menjadi angka. Teknik pengukuran bukanlah hal yang asing bagi seorang peneliti, mengikuti ujian akhir semester di sekolah misalnya, hasil yang didapat (angka) menunjukkan prestasi kita adalah jumlah jawaban benar yang dikerjakan saat menjawab soal-soal yang diberikan. Demikian juga dalam fisika, untuk pengukuran panjang dan suhu misalnya, angka dan satuan didapat dari pengukuran dengan skala yang sudah dikalibrasi.

Namun, dalam ilmu-ilmu sosial kondisinya sedikit berbeda, 'skala' untuk mengukur sikap (measuring attitude) banyak diperdebatkan dan salah satu kritiknya menyatakan bahwa ini bukanlah skala karena tidak dimulai dari angka nol. Kenyataannya teknik pengukuran sikap adalah porsi terbesar dalam penelitian kuantitatif dalam ilmu sosial untuk mendapatkan data yang diinginkan. Bagian ini membahas tiga buah teknik pengukuran yang paling populer digunakan, yaitu peringkat (*rating*) Thurstone, Guttman, dan Likert.

1. Skala Thurstone

Thurstone adalah seorang psikolog yang pada tahun 1928 mengembangkan teknik pengukuran yang disebut '*equal appearing interval scale*'. Awalnya dikembangkan untuk pengukuran sikap, namun kemudian dikembangkan untuk pengukuran atribut nonsikap seperti ciri-sifat kepribadian. Cara singkatnya adalah butir dengan sikap yang berbeda akan berada dalam satu kontinum sikap yang berbeda, yang kemudian dilakukan metode pembobotan untuk setiap butir oleh satu kelompok (penilai), dan menggunakan pembobotan tersebut untuk mengukur orang berdasarkan respons sikap yang dia diberikan. Contohnya sikap tentang hubungan kerja keras dan do'a, dibuatlah beberapa pernyataan yang berhubungan dengan hal itu, baik positif maupun negative. Kemudian dikumpulkan sekelompok orang yang menjadi penilai akan tiap-tiap pernyataan. Para penilai itu harus memberikan respons terhadap setiap pertanyaan dalam rentang tidak setuju (*unfavorable*) dengan nilai 1, netral (nilai 6) sampai ke sikap setuju (*favorable*, nilai =11), seperti contoh berikut.

Beriikanlah tanda X pada kolom sebelah kanan:

Bekerja dan berdo'a haruslah seimbang

Tidak Setuju ↑ Netral X Setuju

Contohnya, dua puluh orang penilai diminta untuk memberikan persetujuannya pada daftar pernyataan tentang hubungan kerja keras dan doa, dua contohnya seperti berikut ini.

1. Rezeki yang saya dapatkan adalah karena kerja keras saya

Tidak Setuju ↑ Netral Setuju

2. Saya yakin bahwa keberhasilan usaha saya karena kerja keras saya

Tidak Setuju ↑ Netral Setuju

Hasil penelitian dari para peneliti tersebut dikumpulkan dan dihitung rata-rata nilai untuk setiap aitemnya. Nilai untuk setiap butir bisa menunjukkan ke sikap tidak setuju (misalnya 2,3), berada di titik netral (6). Atau pun setuju (9,6).

Setiap pertanyaan (aitem) mendapat angka dari tiap-tiap penilai, kemudian dijumlahkan dan didapatkan rata-ratanya, misal seperti ini:

penilai	Rata-rata	dari	20
1. Rezeki yang saya dapatkan adalah karena kerja keras saya.	8,6		
2. Saya yakin bahwa keberhasilan usaha saya karena kerja keras saya	9,4		

Langkah berikutnya adalah mengukur sikap partisipan dengan meminta mereka untuk mengisi jawaban dengan cara mencentang atau mencheck list (ya) atau kosong (tidak), terhadap daftar pertanyaan yang diberikan.

Beriikanlah tanda centang pada pernyataan di bawah yang anda setujui:

_____ 1. Rezeki yang saya dapatkan adalah karena kerja keras saya.

- _____ 2. Saya yakin bahwa keberhasilan usaha saya karena kerja keras saya.
- _____ 3. Selain berusaha keras, saya tidak lupa berdoa untuk keberhasilan usaha saya.
- _____ 4. Usaha tidak perlu maksimal karena Tuhan akan membantu kita.
- _____ 5. Saya lebih banyak berdoa daripada berusaha.
- _____ 6. Saya mengimbangi usaha keras saya dengan terus mendekatkan diri kepada-Nya.
- _____ 7. Hasil kerja keras seseorang tidak ada hubungannya dengan anugerah Tuhan.

Contohnya, peringkat rata-rata dari penilai untuk tujuh pernyataan tersebut adalah $1=8,6; 2=9,4; 3=10,4; 4=6,5; 5=3,3; 6=7,2; 7=2,6$ (informasi pembobotan ini hanya peneliti saja yang tahu).

Seorang responden yang berpartisipasi, memberikan tanda centang untuk pernyataan nomor 2,3, dan 4. Dengan demikian, skor untuk sikapnya adalah: $9,4 + 10 + 6,5 = 25,9$. Kemudian, jumlah skor dibagi banyaknya tanda centang: $25,9/3=8,63$. Dengan rata-rata skor 8,63 maka sikap responden menunjukkan kecenderungan setuju adanya hubungan kerja keras dan berdoa.

2. Skala Guttman

Metode yang digagas oleh Guttman ini juga awalnya dikembangkan untuk mengukur atribut dengan indikator yang memiliki intensitas berbeda-beda. Namun, prosedur ini menggunakan cara yang berbeda, Guttman mengusulkan cara pemeringkatan indikator dari intensitas terendah hingga tertinggi. Dengan cara ini akan mudah diketahui sikap dari partisipan akan suatu isu tertentu dalam satu kontinum hanya cukup berdasar atas satu respons yang diberikan.

Contohnya:

Berikan jawaban tentang peran teknologi informasi (TI) dengan memberikan satu centang saja pada bagian kanan pertanyaan yang anda setuju.

No.	Pernyataan	Respons
1.	TI tidak diperlukan dalam administrasi kantor.	
2.	Penggunaan TI memerlukan pakar dalam pengoperasiannya di kantor.	
3.	TI dapat digunakan di kantor oleh orang yang telah diberikan pelatihan.	
4.	Saya suka apabila ada seseorang yang menggunakan TI untuk saya di kantor.	
5.	Saya akan menggunakan TI jika diberikan pelatihan.	
6.	Saya mau belajar sendiri menggunakan TI untuk kerja di kantor.	

Daftar pernyataan tersebut diurutkan dari sikap yang sangat tidak setuju (1) menuju sikap yang sangat setuju (6) di bagian bawah. Seorang responden cukup mencentang satu saja dari deretan pernyataan tersebut dan kita akan mudah menghitung sikap yang bersangkutan secara keseluruhan. Contoh jika dia mencentang pilihan ketiga (TI dapat digunakan di kantor oleh orang yang telah diberikan pelatihan), Anda bias simpulkan bahwa pernyataan di bawahnya (4,5, dan 6)

Prosedur administrasi dan penskoran model Guttman juga dapat dilakukan dengan cara seperti pada model Thurstone. Responden diminta untuk merespons semua butir, tetapi penskoran dilakukan berdasarkan skor 0 dan 1. Setiap jawaban “ya” akan mendapatkan skor 1, dan setiap jawaban “tidak” akan mendapatkan skor 0. Dengan cara ini, tanggapan yang tidak konsisten dengan urutan butir dapat diketahui.

3. Skala Likert

Rensis Linkert menawarkan model peringkat yang lebih sederhana dibandingkan Thurstone dan Guttman. Hal ini pulalah yang menyebabkan prosedur Linkert sangat populer digunakan di berbagai cabang ilmu social. Responden cukup memberikan jawaban pada setiap butir pernyataan berdasar peringkat sikap yang diberikan. Respons yang diberikan pun bias langsung didapatkan dalam bentuk angka sehingga memudahkan untuk analisis di tahap berikutnya.

Contohnya seperti gambar berikut.

No.	Pernyataan	Sangat Tidak Setuju	Tidak Setuju	Netral	Setuju	Sangat Setuju
1	Saya termasuk orang yang tidak mudah menyerah	[]	[]	[]	[]	[]
2	Dalam segala hal saya selalu gigih untuk meraih sesuatu	[]	[]	[]	[]	[]
3	Meskipun sumber dana sangat minim, saya masih dapat bertahan.	[]	[]	[]	[]	[]
4	Saya tidak dapat belajar dari pengalaman hidup.	[]	[]	[]	[]	[]
5	Dalam menghadapi ber-bagai masalah saya masih dapat berpikir secara positif.	[]	[]	[]	[]	[]

Skor yang diberikan disesuaikan dengan Janis aitem yang diberikan, contoh untuk aitem nomor 1 (pernyataan positif) maka nilai untuk “sangat tidak setuju” = 1 ; “tidak setuju” = 2; terus berlanjut sampai ke pilihan “sangat setuju” = 5. Namun, untuk aitem dengan pernyataan yang negatif, seperti aitem nomor 4 maka skor yang diberikan kebalikannya, yaitu: “sangat tidak setuju” =5, terus ke skor terendah “sangat setuju” =1

Berbagai studi yang dilakukan menunjukkan ketiga metode ini mengindikasikan hasil yang hampir sama. Karena mudahnya prosedur teknik pengukuran dengan Likert rating ini, prosedur ini menjadi pilihan utama dibandingkan model Thurstone yang harus melakukan proses penilaian-pembobotan maupun model Guttman yang relative sulit untuk dibuat. Namun, tentu saja hal ini berdampak lanjutan. Dalam model Thurstone dan Guttman, pernyataan berbeda membawa perbedaan sikap yang diukur, yang berimplikasi adanya skor dari setiap butir. Sebaliknya, dalam model Likert, nilai bobot pada butir menjadi tidak ada. Terlebih lagi jenis data yang dihasilkan adalah skala ordinal-di sini proses analisis yang bisa dilakukannya pun sangat terbatas (yaitu: median dan modus, percentile rank, dan korelasi Spearman). Apalagi, jika dimasukkan syarat sebagai riset kuantitatif yang bagus bahwa butir-butir pernyataan dalam instrument yang dijawab oleh responden haruslah dikalibrasi. Dengan demikian, hanya dengan pemodelan pengukuran Rasch (*Rasch measurement model*) hal ini dapat dilakukan.

4. Tomat Kebahagiaan Thurstone & Likert

Kuesioner dengan model Likert melihat butir sebagai sampel perilaku yang mempresentasikan atribut ukur. Semakin banyak indikator yang melekat, semakin tinggi level atributnya. Oleh karena itu, model ini dinamakan dengan rating sumatif (*summative rating*). Sumatif menunjukkan akumulasi secara total. Semakin banyak indikator kebahagiaan yang melekat dalam diri kita, kebahagiaan kita tergolong tinggi.

Model Thurstone berbeda, indikator dilihat sebagai sampel perilaku yang memiliki intensitas dan relevansi yang berbeda. Butir-butir di dalam kuesioner harus didesain agar memiliki peringkat intensitas yang sama. Skor untuk tiap butir pun disesuaikan dengan intensitas dan relevansinya. Dalam pengukuran kebahagiaan, menjawab “Ya” pada butir “*Saya menikmati kehidupan yang luar biasa ini*” harus mendapatkan skor yang lebih tinggi dibandingkan pada butir “*Saya cukup puas dengan hidup saya*”.

Mari kita analogikan bahwa mendapatkan skor pada satu butir ibarat memetik satu buah tomat dari pohonnya di kebun tomat yang kita tanam. Menurut Likert, level kebahagiaan Anda ditunjukkan dengan jumlah tomat yang Anda petik. Sebaliknya, menurut Thurstone, kebahagiaan Anda ditunjukkan dengan berat untuk semua tomat yang Anda petik.

D. Desain Pembuatan Instrumen

Memiliki jumlah pertanyaan atau pernyataan (item) yang mencukupi dalam kuesioner selalu menjadi tema utama dalam pembuatan kuesioner. Apa yang menjadi hal mendasar tentu saja adalah apakah pernyataan/pertanyaan yang diajukan memang mengukur apa yang seharusnya diukur (validitas)? Peneliti yang membuat kuesioner selalu mendapati pilihan bentuk skala peringkat (*rating*) mana yang akan digunakan serta tantangan bagaimana mengesahkan pilihan yang dibuat. Instrumen apa pun yang dibuat kondisinya adalah invalid sampai hal tersebut diuji dan memenuhi kriteria yang menunjukkan instrumen memang memiliki reliabilitas. Jika hal ini tidak dilakukan, akan menyebabkan semua data yang diperoleh dalam riset tidak berguna dan tidak dapat digunakan untuk interpretasi yang bermakna.

Dengan demikian, pengetahuan dan keterampilan pembuatan kuesioner sangatlah diperlukan oleh peneliti kuantitatif mana pun untuk menghindari kesalahan fatal karena ketidaktahuan. Bagi peneliti pemula (*novice*) adalah hal yang normal ketika dihadapkan pada tantangan pembuatan kuesioner untuk risetnya dengan merasa serba tidak jelas, bingung, atau malah “tersesat”. Kebingungan makin bertambah saat berhadapan dengan ketidakjelasan apa yang seharusnya diukur dan bagaimana mengukurnya.

Bertanya pada pakar pun bukanya menambah kejelasan, berbagai info tambahan yang didapat malah menimbulkan kesangsian, apalagi saat mendiskusikan bahwa kuesioner sesungguhnya adalah instrumen pengukuran. Pembuatan instrumen untuk riset jelas memerlukan pemahaman konsep yang harus didukung keterampilan dan pengetahuan dari bidang lainnya. Hal utama tentunya pemahaman yang jelas tentang ontologi dan epistemology struktur data sebelum menyusun instrumen pengukuran tersebut.

Dalam ilmu-ilmu sosial, khususnya penelitian kuantitatif, urusan pengukuran menjadi hal yang sangat penting karena merupakan inti aktivitas riset itu sendiri. Berbagai peneliti telah melakukan eksplorasi untuk menjadikan instrumen yang memuaskan seperti halnya yang dipersyaratkan dalam metrologi (ilmu pengukuran). Dalam konteks pemodelan Rasch, para pakar seperti Guttman, Mokken, Rasch, Wright menghibiskan waktunya mengembangkan pengukuran dan pengujian instrumen; juga generasi berikutnya, seperti Linacre dan Andrich yang menjadikan Rasch Model lebih konsisten dan dapat diterima. Lebih jauh, Mok dan Wright (2004) menetapkan lima syarat yang harus dipenuhi yang akan menjadikan pengujian pada satu instrumen memang benar-benar valid, yaitu 1) unit Kuantitas terukur; 2) konsep yang terskala; 3) mempunyai interval yang linear; 4) *replicable*; dan 5) dapat melakukan prediksi. Hanya pemodelan Rasch-lah yang dapat memenuhi syarat pengujian instrumen seperti ini.

1) Kuesioner Juga Punya Tingkat Kesulitan

Mudah dan sulit bukan hanya untuk butir pada tes abilitas atau tes prestasi, seperti TOEFL, Tes Potensi Akademik (TPA), atau tes matematika. Mudah dan sulit juga milik butir-butir (item) pernyataan dari kuesioner. Mudah sulit dalam pernyataan di kuesioner tidak terkait dengan pernyataan yang mudah atau sulit di cerna, tetapi pernyataan itu mudah atau sulit disetujui, baik memilih jawaban “Ya” maupun “Setuju”. Jika tingkat kepuasan Anda terhadap sebuah pasta gigi X adalah sedang-sedang saja, Anda akan menjawab:

- “*Setuju*” Pada pernyataan “*Saya suka dengan pasta gigi X*”
- “*Tidak Setuju*” Pada pernyataan “*Saya melakukan Persuasi kepada teman-teman agar memakai pasta gigi X*”

Untuk menjawab “*Setuju*” pada butir kedua dibutuhkan orang yang sangat puas dengan produk pasta gigi X. Hal ini menunjukkan tingkat kesulitan butir itu juga ada pada butir pernyataan pada kuesioner. Butir kedua memiliki intensitas atau kedalaman yang tinggi dibandingkan butir pertama. Perbedaan tingkat intensitas butir ini telah menjadi perhatian dalam pengembangan pengukuran dengan Model Thurstone dan Guttman. Dalam pemodelan Rasch, intensitas butir ini dinamakan dengan lokasi butir (*item locations*).

E. Penyusunan Instrumen

Aturan dasar yang perlu dipertahankan adalah uji validitas yaitu instrumen yang dibuat memang mengukur sesuatu yang seharusnya diukur. Dalam kuesioner hal ini bergantung pada butir-butir konsep dan teknik pengukuran yang digunakan. Pertanyaan berikutnya adalah apakah responden mampu membedakan jawaban dari pertanyaan atau pernyataan yang diberikan sesuatu dengan kategori peringkat yang diberikan.

Analisis dengan pemodelan Rasch mempunyai kemampuan dengan teknik statistiknya untuk memastikan dan memeriksa persepsi setiap responden. Kategori opsi dengan label “Tidak tahu”, “Tidak berlaku”, atau “Lainnya” merupakan contoh pilihan kategori yang salah penempatan sehingga lebih baik tidak digunakan. Contoh lainnya, biasa terdapat pada bagian akhir kuesioner dinyatakan dengan kalimat “Secara keseluruhan/secara umum bagaimana pandangan Anda?” adalah contoh pilihan yang bertentangan dan juga berlebihan. Hal ini tidak diperlukan sehingga dihilangkan saja.

Disarankan bahwa setiap peneliti (khususnya pemula) menyusun instrumen risetnya sendiri. Hal ini akan memberikan nilai tambah bagi si peneliti dalam mengembangkan kreatifitas dan kemandirian dalam pembuatan instrumen. Pada saat yang sama, riset seharusnya memberikan sumbangan yang berharga dan unik bagi ilmu pengetahuan. Hal ini tidak akan terlalu terlihat apabila hanya mengandalkan “instrumen standar” yang sudah dibuat orang lain/pakar lain sehingga tidak ada kebaruan dihasilkan. Untuk memulainya, tidak lain tentu harus merujuk kepa model yang sukses dalam bidang kepakaran tersebut. Biasanya, model yang telah dikembangkan sudah terstruktur dengan baik dan logis, serta telah diuji berulang kali dan mendapatkan hasil yang konsisten. Model inhi akan menjadi kerangka untuk penyusunan instrumen. Dimensi-dimensi yang telah teridentifikasi dan fakti-faktor yang terbukti berkontribusi, dapat menjadi masukan model yang akan kita kembangkan. Hal ini tentu memudahkan kita dalam menyusun butir. Namun, akan menjadi lebih bagus apabila semua ini didukung oleh kajian pustaka yang komprehensif yang akan menyelesaikan masalah ini dengan mudah.

a) Hal-Hal yang Perlu Dipertimbangkan

Berikut ini beberapa saran yang perlu dilakukan saat penyusunan instrumen.

1. Buatlah kuesioner yang singkat dan padat. Seorang peneliti harus mendefinisikan secara tepat apa informasi yang ingin digali. Kuesioner yang dirancang untuk bertanya banyak hal akan menghasilkan kuesioner yang panjang dan bertele-tele. Hal ini berpotensi menyebabkan responden merasa tidak nyaman saat mengerjakannya, responnya menjadi tidak akurat, tidak menyelesaikan secara tuntas, serta memperkecil tingkat pengembaliannya.
2. Lakukan uji validitas isi dan uji coba (*pilot test*) kuesioner. Dapatkan umpan balik dari rancangan awal dengan jumlah sampel yang terbatas, namun representatif-untuk mendapatkan umpan balik dari uji coba lapangan.
3. Berikan pilihan respon pada partisipan yang jelas dari segi bahasa dan logika. Respon setuju terhadap pernyataan atau pernyataan dapat diringkas dalam dua pilihan yang sederhana:
1) *Tidak setuju* 2) *Setuju*
4. Lakukan pengerutan kategorisasi pilihan yang sesuai dengan kebiasaan responden. Ketika responden memberikan pilihan dari sikap terendah ke yang tertinggi, lebih baik mengurutkan dari kiri ke kanan yang makin meningkat:
1) *Tidak pernah* 2) *Jarang* 3) *Kadang-kadang* 4) *Selalu*
5. Berikan pilihan yang jelas apabila menggunakan peringkat positif dan negatif sekaligus. Responden akan mendapat kesulitan untuk menentukan peringkat pada satu butir yang memiliki opini yang lebih kuat; kecenderungannya adalah mereka cenderung memilih peringkat paling ujung yang ditampilkan. Oleh karena itu, pilihan opini yang lebih cocok bagi responden adalah
1) *Tidak setuju* 2) *Cenderung tidak setuju*
3) *Cenderung setuju* 4) *Setuju*
6. Poin tersebut adalah gagasan yang berawal dari usaha untuk memperluas spektrum kontinum tanpa mengubah batas bawah-atas opsi yang sudah ada. Terlihat pada keempat opsi tersebut, batas atas-bawah kontinum tidak berubah dari opsi dua pilihan yang telah dibahas dipoin ketiga, yaitu “tidak setuju” dan “setuju”. Penggunaan jenis ini alternatif untuk mengekspresikan kesetujuan dengan pilihan opini yang bisa digunakan, contohnya,
1) *Sangat tidak setuju* 2) *Tidak setuju* 3) *Setuju* 4) *Sangat setuju*
Dari aspek bahasa kata “Sangat tidak setuju” atau “Sangat setuju” dapat dipersepsi memiliki konotasi yang menunjukkan sesuatu yang terlalu berlebihan. Disisi lain, pada jenis ini spektrum kontinum menjadi melebar dari sifat asalnya yaitu tidak setuju-setuju. Hal ini terlihat dari batas bawah-atas kontinum tidak lagi tidak setuju-setuju, tapi menjadi sangat tidak setuju-sangat setuju.
7. Kuesioner yang mengukur sikap ada baiknya menggunakan butir-butir yang tidak hanya berisi pernyataan positif (*favorable*). Pernyataan yang sifatnya negative juga diperlukan yang ditempatkan secara acak (misal satu pertanyaan negative dalam setiap tujuh pernyataan positif). Tujuannya adalah untuk meningkatkan keseriusan responden dan membuat mereka waspada dengan isi pernyataan dalam mengerjakan kuesioner yang kita berikan. Tentu saja hal yang tidak boleh dilupakan, peringkat yang diberikan pun akan terbalik dibandingkan pernyataan positif.

b) Hal-Hal yang Perlu Diperhatikan

Berikut ini saran mengenai hal-hal yang perlu dihindari dalam penyusunan kuesioner.

1. Hindari pernyataan terbuka, biasanya responden tidak ingin atau tidak mampu untuk memberikan respons secara tertulis.
2. Hindari alternatif pilihan jawaban “lainnya”. Responden yang tidak teliti akan cenderung memilih pilihan “lainnya” apabila dirasa banyak ketidakpastian atau karena alasan sepele.
3. Hindari memberikan pilihan yang sangat beragam, contohnya:
1) Tidak Pernah 2) Jarang 3) Kadang-kadang 4) Cukup sering
5) Sering 6) Sering kali 7) Hampir selalu 8) Selalu dilakukan

Pilihan tersebut potensial membingungkan dan mengganggu responden untuk menentukan pilihan mana yang paling tepat dengan kondisinya. Berbagai riset psikometri menunjukkan kebanyakan responden tidak dapat melihat perbedaan lebih dari enam pilihan, beberapa lainnya menemukan tidak lebih dari tujuh pilihan. Hal ini menunjukkan lima pilihan yang diberikan adalah ideal. Analisa dengan pemodelan Rasch menunjukkan bahwa sensitivitas pengukuran tidak hilang walau menggunakan peringkat yang tidak beragam. Bahkan, hasil analisis bias menunjukkan peringkat pilihan jika perlu digabungkan atau dipecah.

4. Sehubungan dengan poin tersebut, terdapat berbagai jenis butir yang memang diperlukan posisi netral, contohnya untuk pernyataan

“Bagaimana dengan tugas pekerjaan rumah yang diberikan:”

- 1) Terlalu sedikit 2) Layak 3) terlalu banyak

Sesuatu yang tidak realistis meminta responden untuk menilai sesuatu yang sifatnya bias dibandingkan, namun diberikan pilihan netral, seperti berikut.

“Penilai yang dilakukan oleh dosen metodologi riset,
Menurut saya adalah”

- 1) Setuju 2) Cenderung setuju 3) Tidak menjawab
4) Cenderung tidak setuju 5) Tidak setuju

5. Hindari perintah yang meminta responden untuk memeringkat. Responden tidak akan bisa memberikan respons ranking lebih dari enam buah dalam satu waktu tertentu, kebanyakannya salah interpretasi atau membuat respons yang keliru. Untuk menghindarinya, Contoh berikut bisa membantu untuk mereka memberikan respons yang tepat:

Ini adalah pilihan warna untuk dinding kantor:

- 1) putih 2) krem 3) abu-abu

- a) Warna yang paling tepat menurut saya adalah? _____
- b) Pilihan warna kedua bagi saya adalah? _____
- c) Pilihan warna terakhir yang saya pilih? _____

F. Penggunaan Opsi Tengah

Upaya untuk melibatkan opsi tengah (*middle category*) adalah upaya untuk memfasilitas responden yang memiliki sikap yang moderat. Apa yang dimaksud dengan opsi tengah disini adalah respons yang menunjukkan sikap antara negatif (tidak setuju) dan positif (setuju), bukan opsi “tidak tahu”, “tidak menjawab”, atau “tidak yakin” (*unsure*). Permasalahan yang muncul adalah dengan disediakannya opsi tengah maka akan banyak responden yang memilikinya. Mereka adalah orang-orang yang kurang kooperatif, kurang termotivasi, tidak memahami butir dan instruksi atau cenderung menyembunyikan diri. Sebaliknya, tidak disediakannya opsi tengah menyebabkan responden merasa dipaksa untuk memilih alternatif yang tidak mencerminkan

sikapnya. Keterpaksaan ini akan memberikan kontribusi terhadap besarnya *error* sistematis dalam pengukuran.

Peneliti tentang opsi tengah menunjukkan hasil yang sangat beragam karena penelitian yang dilakukan dilatarbelakangi oleh konteks yang berbeda-beda. *Brace* (2004) mengatakan bahwa penelitalah yang perlu harus memutuskan apakah memasukkan opsi tengah atau tidak karena peneliti mengetahui konteks instrumen yang dikembangkannya. Berikut ini beberapa pertimbangan yang dapat dipakai untuk melibatkan atau tidak melibatkan opsi tengah.

1. *Fungsi pengukuran.* Penggunaan opsi tengah bergantung pada spesifikasi isi dan fungsi pengukuran (Friedman et al., 1981). Jika kita ingin menggali informasi yang lebih eksploratif, opsi tengah kita digunakan. Sebaliknya, ketika ketegasan respons dianggap penting, Anda mengondisikannya dengan cara tidak menyediakan opsi tengah.
2. *Konteks Pengukuran.* Dalam konteks pengukuran untuk keperluan penelitian secara umum ketika responden tidak mendapatkan keuntungan (misalnya, ekonomis) atas respons yang diberikan, respons tengah perlu dipakai. Sebaliknya, ketika keuntungan tersebut ada, perlu dihindari. Contohnya, dalam konteks seleksi karyawan, penggunaan skala tanpa opsi tengah lebih mampu mereduksi kepatutan sosial (*social desirability*) dibandingkan yang menggunakan kategori tengah (Garland, 1991). Termasuk disini adalah jenis atribut yang diukur. Butir-butir pada atribut yang sensitif, seperti perilaku seksual. Konsumsi alkohol, atau gaya hidup, akan menstimulasi responden untuk memilih opsi tengah.
3. *Jenis Pernyataan.* Untuk pernyataan yang terkait dengan perilaku pada masa lalu, kita tidak mungkin menyediakan opsi kategori tengah. Opsi tengah tidak dapat disediakan untuk pernyataan yang terkait perilaku atau pengalaman masa lalu. Contohnya, “*Saya pernah berurusan dengan polisi*” atau yang memfokuskan pada perilaku, seperti “*Saya akan menegur orang yang merokok di depan saya*”.
4. *Kejelasan Pernyataan.* Pada beberapa instrumen, butir yang memiliki kejelasan yang rendah (*subtle item*) dipakai untuk mengurangi responden memberikan respons distoratif. Menghadapi butir ini, responden akan kesulitan memahaminya sehingga akan memilih opsi tengah. Jadi, kejelasan *statement* butir dapat menjadi pertimbangan untuk melibatkan atau tidak melibatkan opsi tengah.
5. *Karakteristik Responden.* Telah dikemukakan sebelumnya bahwa faktor yang menyebabkan banyaknya memilih opsi tengah adalah butir dan responden. Masalah butir relatif bisa dikendalikan sehingga dapat diatasi dengan cara menuliskan butir yang sederhana dan mudah dipahami. Pada masalah yang bersumber pada responden, banyaknya memilih opsi tengah diakibatkan oleh ketidakmauan mereka untuk bekerja sama dan menyembunyikan diri. Jika kita yakin, bahwa responden kita dapat diajak bekerja sama dan terbuka, opsi tengah dapat kita libatkan.

PENGEMBANGAN INSTRUMEN RANAH KOGNITIF

BAB IV

A. Pengantar

Instrumen penelitian merupakan bagian penting dari suatu proses penilaian dan penelitian secara keseluruhan, sedangkan bahan ajar merupakan bagian penting dari suatu proses pembelajaran secara keseluruhan.

Dengan demikian, tujuan yang ingin dicapai dari pengembangan instrumen penelitian dan penilaian dalam pendidikan ini adalah tersusunnya bahan ajar dan instrumen untuk mengukur kemampuan secara menyeluruh (*holistic*) kepada peserta didik, tervalidasi, mempunyai reliabilitas, daya pembeda (DP), dan indeks kesukaran (IK) yang memadai, dan keberfungsian pengecoh (*distractor*).

B. Penerapan Pengukuran dalam Domain Kognitif

1. Karakteristik Ranah Penilaian Kognitif

Aspek kognitif berhubungan dengan kemampuan berfikir termasuk didalamnya kemampuan memahami, menghafal, mengaplikasi, menganalisis, mensistesis dan kemampuan mengevaluasi. Menurut Taksonomi Bloom (Sax, 1980), kemampuan kognitif adalah kemampuan berfikir secara hirarki yang terdiri dari pengetahuan, pemahaman, aplikasi, analisis, sintesis dan evaluasi.

Karakteristik kognitif untuk pengukuran sebagaimana Bloom telah menggambarkan sebagai berikut :

1. Mendefinisikan istilah teknis dengan memberikan atribut, sifat, atau relasi;
2. Kemampuan untuk membedakan referensi untuk kata-kata dan membangun batasan agar istilah biologis memiliki arti;
3. Keakraban dengan sejumlah besar kata-kata dalam rentangan maknanya;
4. Pengetahuan tentang perbendaharaan kata tentang seni yang bisa dibaca dan dikonversikan dengan cermat ;
5. Mengakui pengertian perbendaharaan kata dalam pemikiran kuantitatif;
6. Pengetahuan tentang istilah-istilah akuntansi yang penting;
7. Penguasaan tentang istilah-istilah untuk bekerja dalam bidang ilmu pengetahuan ;
8. Memahami pengertian terminologi berkaitan dengan bangun-bangun geometrik. (Bloom B., 1956:64-65)

Pada tingkat *pengetahuan*, peserta didik menjawab pertanyaan berdasarkan hafalan saja. Pada tingkat *pemahaman* peserta didik dituntut untuk menyatakan masalah dengan kata-katanya sendiri, memberi contoh suatu konsep atau prinsip. Pada tingkat *aplikasi*, peserta didik dituntut untuk menerapkan prinsip dan konsep dalam situasi yang baru. Pada tingkat *analisis*, peserta didik diminta untuk untuk menguraikan informasi ke dalam beberapa bagian, menemukan asumsi,

membedakan fakta dan pendapat serta menemukan hubungan sebab-akibat. Pada tingkat *sintesis*, peserta didik dituntut untuk menghasilkan suatu cerita, komposisi, hipotesis atau teorinya sendiri dan mensintesis pengetahuan. Pada tingkat *evaluasi*, peserta didik mengevaluasi informasi seperti bukti, sejarah, editorial, teori-teori yang termasuk di dalamnya *judgement* terhadap hasil analisis untuk membuat kebijakan.

Tujuan aspek kognitif berorientasi pada kemampuan berfikir yang mencakup kemampuan intelektual yang lebih sederhana, yaitu mengingat, sampai pada kemampuan memecahkan masalah yang menuntut siswa untuk menghubungkan dan menggabungkan beberapa ide, gagasan, metode atau prosedur yang dipelajari untuk memecahkan masalah tersebut.

Dengan demikian aspek kognitif adalah sub-taksonomi yang mengungkapkan tentang kegiatan mental yang sering berawal dari tingkat pengetahuan sampai ke tingkat yang paling tinggi yaitu evaluasi. Aspek kognitif terdiri atas enam tingkatan dengan aspek belajar yang berbeda-beda. Keenam tingkat tersebut yaitu:

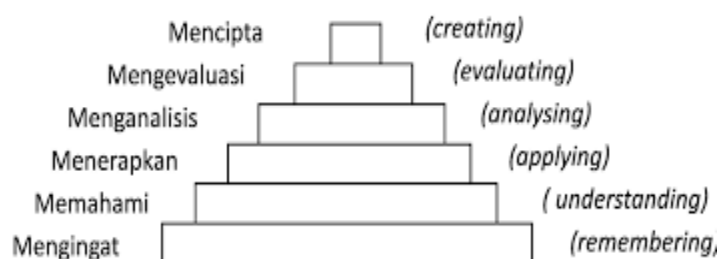
1. Tingkat pengetahuan (*knowledge*), pada tahap ini menuntut siswa untuk mampu mengingat (*recall*) berbagai informasi yang telah diterima sebelumnya, misalnya fakta, rumus, terminologi strategi problem solving dan lain sebagainya.
2. Tingkat pemahaman (*comprehension*), pada tahap ini kategori pemahaman dihubungkan dengan kemampuan untuk menjelaskan pengetahuan, informasi yang telah diketahui dengan kata-kata sendiri. Pada tahap ini peserta didik diharapkan menerjemahkan atau menyebutkan kembali yang telah didengar dengan kata-kata sendiri.
3. Tingkat penerapan (*application*), penerapan merupakan kemampuan untuk menggunakan atau menerapkan informasi yang telah dipelajari kedalam situasi yang baru, serta memecahkan berbagai masalah yang timbul dalam kehidupan sehari-hari.
4. Tingkat Analisis (*analysis*), analisis merupakan kemampuan mengidentifikasi, memisahkan dan membedakan komponen-komponen atau elemen suatu fakta, konsep, pendapat, asumsi, hipotesa atau kesimpulan, dan memeriksa setiap komponen tersebut untuk melihat ada atau tidaknya kontradiksi. Dalam tingkat ini peserta didik diharapkan menunjukkan hubungan di antara berbagai gagasan dengan cara membandingkan gagasan tersebut dengan standar, prinsip atau prosedur yang telah dipelajari.
5. Tingkat sintesis (*synthesis*), sintesis merupakan kemampuan seseorang dalam mengaitkan dan menyatukan berbagai elemen dan unsur pengetahuan yang ada sehingga terbentuk pola baru yang lebih menyeluruh.
6. Tingkat evaluasi (*evaluation*), evaluasi merupakan level tertinggi yang mengharuskan peserta didik mampu membuat penilaian dan keputusan tentang nilai suatu gagasan, metode, produk atau benda dengan menggunakan kriteria tertentu.

Apabila melihat kenyataan yang ada dalam sistem pendidikan yang diselenggarakan, pada umumnya baru menerapkan beberapa aspek kognitif tingkat rendah, seperti pengetahuan, pemahaman dan sedikit penerapan. Sedangkan tingkat analisis, sintesis dan evaluasi jarang sekali diterapkan. Apabila semua tingkat kognitif diterapkan secara merata dan terus-menerus maka hasil pendidikan akan lebih baik.

Tabel 5 Kaitan Antara Kegiatan Pembelajaran dengan Domain Tingkatan Aspek Kognitif

No	Tingkatan	Deskripsi
1	Pengetahuan	Arti: Pengetahuan terhadap fakta, konsep, definisi, nama, peristiwa, tahun, daftar, teori, prosedur, dll. Contoh kegiatan belajar: • Mengemukakan arti • Menentukan lokasi • Mendriskripsikan sesuatu • Menceritakan apa yang terjadi • Menguraikan apa yang terjadi
2	Pemahaman	Arti: pengertian terhadap hubungan antar-faktor, antar konsep, dan antar data hubungan sebab akibat penarikan kesimpulan. Contoh kegiatan belajar: • Mengungkapkan gagasan dan pendapat dengan kata-kata sendiri • Membedakan atau membandingkan • Mengintepretasi data • Mendriskripsikan dengan kata-kata sendiri • Menjelaskan gagasan pokok • Menceritakan kembali dengan kata-kata sendiri.
3	Aplikasi	Arti: Menggunakan pengetahuan untuk memecahkan masalah atau menerapkan pengetahuan dalam kehidupan sehari-hari. Contoh kegiatan: • Menghitung kebutuhan • Melakukan percobaan • Membuat peta • Membuat model • Merancang strategi
4	Analisis	Artinya: menentukan bagian-bagian dari suatu masalah, penyelesaian, atau gagasan dan menunjukkan hubungan antar bagian tersebut Contoh kegiatan belajar: • Mengidentifikasi faktor penyebab • Merumuskan masalah • Mengajukan pertanyaan untuk mencari informasi • Membuat grafik • Mengkaji ulang
5	Sintesis	Artinya: menggabungkan berbagai informasi menjadi satu

		kesimpulan/konsep atau meramu/merangkai berbagai gagasan menjadi suatu hal yang baru Contoh kegiatan belajar: • Membuat desain • Menemukan solusi masalah • Menciptakan produksi baru,dst.
6	Evaluasi	Arti: mempertimbangkan dan menilai benar-salah, baik-buruk, bermanfaat-tidak bermanfaat Contoh kegiatan belajar: • Mempertahankan pendapat • Membahas suatu kasus • Memilih solusi yang lebih baik • Menulis laporan,dst.



Gambar 10. Hierarchi Taksonomi Bloom Versi Revisi

1. Mengingat (*remembering*)

Mengingat merupakan proses kognitif paling rendah tingkatannya. Untuk mengkondisikan agar “mengingat” bisa menjadi bagian belajar bermakna, tugas mengingat hendaknya selalu dikaitkan dengan aspek pengetahuan yang lebih luas dan bukan sebagai suatu yang lepas dan terisolasi. Kategori ini mencakup dua macam proses kognitif yaitu mengenali (*recognizing*) dan mengingat. Kata operasional mengetahui yaitu mengutip, menjelaskan, menggambar, menyebutkan, membilang, mengidentifikasi, memasang, menandai, menamai.

2. Memahami (*understanding*).

Pertanyaan pemahaman menuntut siswa menunjukkan bahwa mereka telah mempunyai pengertian yang memadai untuk mengorganisasikan dan menyusun materi-materi yang telah diketahui. Siswa harus memilih fakta-fakta yang cocok untuk menjawab pertanyaan. Jawaban siswa tidak sekedar mengingat kembali informasi, namun harus menunjukkan pengertian terhadap materi yang diketahuinya. Kata operasional memahami yaitu menafsirkan, meringkas, mengklasifikasikan, membandingkan, menjelaskan, membeberkan.

3. Menerapkan (*applying*).

Pertanyaan penerapan mencakup penggunaan suatu prosedur guna menyelesaikan masalah atau mengerjakan tugas. Oleh karena itu, mengaplikasikan berkaitan erat dengan pengetahuan prosedural. Namun tidak berarti bahwa kategori ini hanya sesuai untuk pengetahuan prosedural saja. Kategori ini mencakup dua macam proses kognitif yaitu menjalankan dan mengimplementasikan. Kata operasionalnya melaksanakan, menggunakan, menjalankan, melakukan, mempraktekan, memilih, menyusun, memulai, menyelesaikan, mendeteksi.

4. Menganalisis (*analyzing*).

Pertanyaan analisis menguraikan suatu permasalahan atau obyek ke unsur-unsur-unsurnya dan menentukan bagaimana saling keterkaitan antar unsur-unsur tersebut. Kata operasionalnya yaitu menguraikan, membandingkan, mengorganisir, menyusun ulang, mengubah struktur, mengkerangkakan, menyusun outline, mengintegrasikan, membedakan, menyamakan, membandingkan, mengintegrasikan.

5. Mengevaluasi (*evaluating*).

Mengevaluasi membuat suatu pertimbangan berdasarkan kriteria dan standar yang ada. Ada dua macam proses kognitif yang tercakup dalam kategori ini adalah memeriksa dan mengkritik. Kata operasionalnya yaitu menyusun hipotesis, mengkritik, memprediksi, menilai, menguji, membenarkan, menyalahkan.

6. Mencipta (*creating*).

Membuat adalah menggabungkan beberapa unsur menjadi suatu bentuk kesatuan. Ada tiga macam proses kognitif yang tergolong dalam kategori ini yaitu membuat, merencanakan, dan memproduksi. Kata operasionalnya yaitu merancang, membangun, merencanakan, memproduksi, menemukan, membarui, menyempurnakan, memperkuat, memperindah, mengubah.

Tabel 6.
Kata Operasional Pada Ranah Kognitif

Mengingat (C1)	Memahami (C2)	Menerapkan (C3)	Menganalisis (C4)	Mengevaluasi (C5)	Menciptakan (C6)
1	2	3	4	5	6
Mengenali Mengingat kembali Membaca Menyebutkan Melafalkan/ melafazkan Menuliskan Menghafal	Menjelaskan Mengartikan Menginterpretasikan Menceritakan Menampilkan Memberi contoh Merangkum Menyimpulkan Membandingkan Mengklasifikasikan Menunjukkan Menguraikan Membedakan Mengidentifikasi	Melaksanakan Mengimplementasikan Menggunakan Mengonsepan Menentukan Memproseskan	Mendiferensiasi -kan Mengorganisasi -kan Mengatribusi- kan Mendiagnosis Memerinci Menelaah Mendeteksi Mengaitkan Memecahkan Menguraikan	Mengecek Mengkritik Membuktikan Mempertahankan Memvalidasi Mendukung Memproyeksikan	Membangun Merencanakan Memproduksi Mengkombinasi -kan Merancang Merekonstruksi Membuat Menciptakan Mengabstraksi

2. Pengukuran Ranah Penilaian Kognitif

Apabila melihat kenyataan yang ada dalam sistem pendidikan yang diselenggarakan, pada umumnya baru menerapkan beberapa aspek kognitif tingkat rendah, seperti pengetahuan, pemahaman dan sedikit penerapan. Sedangkan tingkat analisis, sintesis dan evaluasi jarang sekali diterapkan. Apabila semua tingkat kognitif diterapkan secara merata dan terus-menerus maka hasil pendidikan akan lebih baik. Pengukuran hasil belajar ranah kognitif dilakukan dengan tes tertulis.

Bentuk tes kognitif diantaranya: (1) tes atau pertanyaan lisan di kelas, (2) pilihan ganda, (3) uraian obyektif, (4) uraian non obyektif atau uraian bebas, (5) jawaban atau isian singkat, (6) menjodohkan, (7) portopolio dan (8) performans.

Cakupan yang diukur dalam ranah Kognitif adalah:

- 1) Ingatan (C1) yaitu kemampuan seseorang untuk mengingat. Ditandai dengan kemampuan menyebutkan simbol, istilah, definisi, fakta, aturan, urutan, metode.
- 2) Pemahaman (C2) yaitu kemampuan seseorang untuk memahami tentang sesuatu hal. Ditandai dengan kemampuan menerjemahkan, menafsirkan, memperkirakan, menentukan, menginterpretasikan.
- 3) Penerapan (C3), yaitu kemampuan berpikir untuk menjangkau & menerapkan dengan tepat tentang teori, prinsip, simbol pada situasi baru/nyata. Ditandai dengan kemampuan menghubungkan, memilih, mengorganisasikan, memindahkan, menyusun, menggunakan, menerapkan, mengklasifikasikan, mengubah struktur.
- 4) Analisis (C4), Kemampuan berfikir secara logis dalam meninjau suatu fakta/ objek menjadi lebih rinci. Ditandai dengan kemampuan membandingkan, menganalisis, menemukan, mengalkasikan, membedakan, mengkategorikan.
- 5) Sintesis (C5), Kemampuan berpikir untuk memadukan konsep-konsep secara logis sehingga menjadi suatu pola yang baru. Ditandai dengan kemampuan mensintesis, menyimpulkan, menghasilkan, mengembangkan, menghubungkan, mengkhususkan.
- 6) Evaluasi (C6), Kemampuan berpikir untuk dapat memberikan pertimbangan terhadap suatu situasi, sistem nilai, metoda, persoalan dan pemecahannya dengan menggunakan tolak ukur tertentu sebagai patokan. Ditandai dengan kemampuan menilai, menafsirkan, mempertimbangkan dan menentukan.

Contohnya siswa dibina kompetensinya menyangkut kemampuan melukis jaring-jaring kubus. Namun, untuk dapat melukis jaring-jaring kubus setidaknya diperlukan pengetahuan (kognitif) tentang bentuk-bentuk jaring kubus dan cara-cara melukis garis-garis tegak lurus.

Tabel 7. Kompetensi dan Kata Operasional pada Ranah Kognitif

Domain	Kategori Jenis Perilaku	Kemampuan Internal	Kata Kerja Operasional
Kognitif	Pengetahuan	Mengetahui Misalnya : -Istilah -Fakta -Aturan -Urutan -Metode	Menyusun Menata Mendefinisikan Menyalin Menunjuk (nama benda) Mendaftar Menghafalkan Menyebutkan Mengurutkan

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

			Mengenal Menghubungkan Mengingat kembali Memproduksi
	Pemahaman	Menterjemahkan Menafsikan Memperkirakan Menentukan..... Misalnya : -metode -prosedur Memahami... Misalnya: -konsep -kaidah -prinsip -kaitan antar fakta -isi pokok Mengartikan/Menginterpre- tasikan... Misalnya : -Tabel -Grafik -Bagan	Mengklasifikasikan Mengggambarkan Mendiskusikan Menjelaskan Mengungkapkan Mendefinisikan Menunjukkan Mengalokasikan Melaporkan Mengakui Menjatuhkan Mengkajiulang Memilih Menyatakan Menterjemahkan
	Penerapan	Memecahkan masalah Membuat bagan dan grafik Menggunakan..... Misalnya : -Metode/Prosedur -Konsep -Kaidah -Prinsip	Menerapkan Memilih Mendemonstrasikan Mendramatisir Mengerjakan Membuat ilusi Menginterpretasikan Mengoperasikan Melatih Menyusun jadual Membuat sketsa Memcahkan Mengakui
	Analisis	Mengenali kesalahan Membedakan Misalnya: -fakta dari interpretasi data dari kesimpulan	Mengenali Mengira-ira Menghitung Mengkategorikan Membandingkan Melawankan Mengkritik Membuat diagram Membedakan Memperlakukan lain Menguji Mencoba Menginventaris Menanyakan Mengetes Membuat lain (dari yang lain
	Sintesis	Menghasilkan Misalnya : -klasifikasi -karangan	Mengatur (sesuai dengan) Merangkum Mengumpulkan Mengatur komposisi

		-kerangka teoritis Menyusun: Misalnya : -Rencana -Skema -Program kerja	Membangun Menciptakan Merancang Merumuskan Mengatur Mengorganisasikan Merencanakan Menyiapkan Mengusulkan Menyusun Menulis
	Evaluasi	Menilai berdasarkan norma internal..... Misalnya : Hasil karya seni Mutu karangan Mutu ceramah Program penataran Menilai berdasarkan norma eksternal..... Misalnya : Hasil karya seni Mutu karangan Mutu pekerjaan Mutu ceramah Program penataran Mempertimbangkan..... Misalnya : Baik buruknya Pro-kontrannya Untung-ruginya	Menduga-duga Membuat argumentasi Mengkoreksi Melampirkan Memilih Membandingkan Mempertahankan Mengestimasi Memutuskan Mengira-ira Menganggap Memberi nilai (score) Memilih Mendukung Menilai Mengevaluasi

3. Instrumen Tes Hasil Belajar (*Achievement Test*)

Komponen penting dalam silabus maupun rencana pelaksanaan pembelajaran (RPP) adalah indikator pencapaian kompetensi. Menurut Permendikbud No. 65 Tahun 2013 tentang Standar Proses, Rencana Pelaksanaan Pembelajaran (RPP) adalah rencana kegiatan pembelajaran tatap muka untuk satu pertemuan atau lebih. RPP dikembangkan dari silabus untuk mengarahkan kegiatan pembelajaran peserta didik dalam upaya mencapai Kompetensi Dasar. Selanjutnya menurut Permendikbud Nomor 81A Tahun 2013 Lampiran IV tentang Implementasi Kurikulum Pedoman Umum Pembelajaran (Kemdikbud, 2013: 37) tahapan pertama dalam pembelajaran menurut Standar Proses adalah perencanaan pembelajaran yang diwujudkan dengan kegiatan penyusunan Rencana Pelaksanaan Pembelajaran (RPP). RPP adalah rencana pembelajaran yang dikembangkan secara rinci dari suatu materi pokok atau tema tertentu yang mengacu pada silabus. RPP adalah rencana kegiatan pembelajaran tatap muka untuk satu pertemuan atau lebih. RPP dikembangkan secara rinci dari suatu materi pokok atau tema tertentu yang mengacu pada silabus untuk mengarahkan kegiatan pembelajaran siswa dalam upaya mencapai Kompetensi Dasar (KD). Komponen ini penting karena menjadi dasar untuk menyusun indikator penilaian. Indikator penilaian perlu dirumuskan untuk dijadikan dasar pedoman penilaian bagi guru, peserta didik, maupun pengawas di sekolah. Setiap penilaian yang dilakukan melalui tes dan non tes harus sesuai

dengan indikator penilaian. Indikator penilaian ini menggunakan kata kerja lebih terukur dibandingkan dengan indikator pencapaian kompetensi.

Rumusan indikator penilaian memiliki batasan-batasan tertentu sehingga dapat dikembangkan menjadi instrumen penilaian dalam bentuk soal, lembar pengamatan, dan atau penilaian hasil karya atau produk, termasuk penilaian diri.

Kata Kerja Operasional (KKO) untuk pengembangan Indikator Silabus dan rencana pelaksanaan pembelajaran (RPP) berdasarkan taksonomi Bloom dibagi dalam beberapa pencapaian kompetensi dasar, yang ditandai oleh perubahan perilaku yang dapat diukur yang mencakup sikap, pengetahuan, dan keterampilan. Indikator dikembangkan sesuai dengan karakteristik peserta didik, mata pelajaran, satuan pendidikan, potensi daerah dan dirumuskan dalam kata kerja operasional yang terukur dan/atau dapat diobservasi.

Instrument penelitian merupakan salah satu komponen penting yang diperlukan dalam kegiatan penilaian dan penelitian. Dalam konteks pembelajaran, instrumen penelitian jenis tes dijadikan alat untuk mengukur hasil belajar. Hal penting dilakukan bagi peserta didik untuk mengetahui tingkat kemajuan siswa dalam kegiatan pembelajaran. Walaupun pada realitasnya kadang kala dalam proses belajar mengajar (PBM), aspek evaluasi hasil belajar ini diabaikan. Artinya, dosen, guru atau instruktur terlalu memperhatikan penyajian pelajaran saja. Perkuliahan atau pelajaran berjalan baik, praktikum berjalan rapi, namun saat membuat tes atau soal praktikum, tidak lagi melihat tujuan pembelajaran yang pernah dibuatnya. Akibatnya, tes hasil belajar yang dibuat terkesan seperti jatuh dari langit saja. Artinya dosen atau guru dalam membuat soal tes menjadi seadanya atau seingatnya saja, serta merta tanpa ada perencanaan yang baik, tidak harus memenuhi kriteria pembuatan tes yang baik dan benar. Misalnya apakah soal ujian tersebut sudah sesuai dengan tujuan yang telah ditetapkan?, apakah dalam pembuatan soal sudah memperhatikan aspek kognitif, afektif dan psikomotorik, dan sebagainya.

Hal penting yang harus diperhatikan dalam penyusunan tes hasil belajar (*Achievement Test*) untuk keperluan penilaian dan penelitian, adalah:

- 1) Tes berfungsi untuk memperoleh informasi tentang kemampuan subjek penelitian.
- 2) Fungsi tes. Proses kegiatan yang berupa mendiskusikan tentang fungsi penilaian untuk memperoleh pemahaman tentang hal-hal apa saja yang dapat dinilai melalui pelaksanaan suatu tes. Apakah sekedar memberi nilai untuk menentukan lulus tidaknya mahasiswa atau siswa. Ataukah ada fungsi-fungsi lain yang ingin dicapai melalui penilaian tersebut misalnya data yang diperoleh digunakan untuk penelitian.
- 3) Penentuan kriteria. Menentukan kriteria penilaian untuk kepentingan penelitian ini berarti untuk melakukan penilaian yang baik dibutuhkan mutu soal tes yang baik pula. Dalam praktek pengajaran, tes dilaksanakan dengan memberikan serangkaian soal. Tes hasil belajar dengan demikian akan sangat tergantung pada mutu tes. Semakin bermutu soal yang diberikan, dan semakin bermutu suatu tes semakin terandalkan pula penilaian yang diperoleh dan hal ini berdampak pada makin baik data yang diperoleh untuk keperluan penilaian dan penelitian yang berkualitas.
- 4) Merancang soal-soal yang memberikan kepada subjek penelitian dalam suatu struktur soal sedemikian rupa sehingga jumlah maupun derajat kesukaran soal yang tetap relevan dengan pencapaian tujuan pembelajaran yang telah ditetapkan dalam Rancangan Kegiatan Belajar Mengajar (RKBM).

- 5) Mengingat karakteristik soal terkait derajat kesukaran soal yang terjadi dapat berbeda-beda satu dengan lainnya, tiap-tiap soal perlu mendapat bobot soal menurut relevansinya dengan tujuan belajar yang dilakukan.
- 6) Sesudah proses membuat, menstrukturkan, dan menentukan bobot soal, maka soal-soal tersebut disajikan melalui ujian dengan memperhatikan prosedur pelaksanaan ujian yang baik.
- 7) Setelah itu dilakukan pengukuran dan penilaian hasil ujian, maka dilakukan penskoran dan pengolahan nilai hasil ujian. Nilai kasar (*raw score*) menjadi nilai terjabar (*drive score*) dengan menetapkan batas kelulusan (*passing grade*) merupakan kegiatan tindak lanjut untuk keperluan interpretasi hasil penilaian dan penelitian dalam meningkatkan mutu pembelajaran di sekolah.

Memperhatikan serangkaian alur pikir di atas, maka aktivitas yang dilaksanakan pertama kali oleh penilai dan peneliti adalah mendiskusikan materi bahasan secara urut, sejak dari persiapan sampai pada tes itu diadakan dengan mengacu pada pokok bahasan yang telah diberikan guru berdasarkan kurikulum yang berlaku.

Peran instrumen berupa tes dalam kegiatan penilaian dan penelitian sangatlah berharga. Tes merupakan seperangkat rangsangan (*stimuli*) yang diberikan kepada seseorang dengan maksud untuk mendapatkan jawaban-jawaban yang dapat dijadikan dasar bagi penetapan skor angka. Skor yang didasarkan pada sampel yang *representative* dari tingkah laku pengikut tes merupakan indikator tentang seberapa jauh orang yang dites itu memiliki karakteristik yang sedang diukur. Perlu diingat ada dua persyaratan pokok dari tes yang digunakan untuk pengumpulan data penelitian yakni validitas dan reliabilitas, disamping daya pembeda (DP), dan indeks kesukaran (IK) yang memadai, dan keberfungsian pengecoh (*distractor*).

3.1. Jenis Tes

- a) *Achievement test* (tes hasil belajar) adalah tes yang digunakan untuk mengukur kepuasan dan kecakapan individu dari berbagai bidang pengetahuan.
 - b) Jenis *achievement test* terdiri atas (1) test standar atau test baku, yaitu test yang dipublikasikan dan yang telah disiapkan oleh para ahli secara cermat serta mencakup tujuan akademis yang lazim bagi sekolah-sekolah pada umumnya, (2) test buatan guru adalah test yang disusun oleh guru berdasarkan taksonomi yang ingin diukur dari unjuk kerja seseorang setelah mempelajari sesuatu.
- 1) Fungsi *achievement test* untuk mengukur kemampuan seseorang dalam satu bidang atau bidang tertentu.
 - 2) Bentuk tes terdiri dari tes (a) objektif dan (b) isian (*essay*)

3.1.1 Bentuk tes obyektif

a) Tes Benar Salah (*True-False*)

Tes benar salah adalah tes yang memuat pernyataan-pernyataan (*statement*). Pernyataan tersebut ada yang benar dan ada yang salah. Orang yang ditanyakan tugasnya hanya menandai masing-masing pernyataan itu dengan melingkari huruf “B” jika pernyataan benar, dan “S” jika pernyataan salah.

Contoh :

B – S 1. Jakarta adalah ibukota Negara Indonesia
B – S 2. Semua binatang tidak berdaun telinga berkembang
biak dengan bertelur

Petunjuk Penyusunan Tes Benar-Salah

- Tulislah huruf B-S pada permulaan masing-masing item dengan maksud untuk mempermudah mengerjakan dan menskorning.
- Disarankan jumlah butir yang harus dijawab B sama dengan yang dijawab S. dengan demikian perlu dihindari menyusun pola jawaban dalam bentuk B-S-B-S-B-S atau SS-BB-SS-BB-SS.
- Perlu dihindari item yang masih diperdebatkan.
- Perlu dihindari pernyataan-pernyataan yang hampir sama dengan yang ada di buku.

b) Tes Pilihan Ganda (*Multiple Choice Test*)

Dalam pengembangan instrument tes pilihan ganda. Penilai dan peneliti harus memperhatikan beberapa konseptual tentang tes pilihan ganda. Tes pilihan ganda adalah tes yang memuat serangkaian informasi yang belum lengkap, dan untuk melengkapinya adalah dengan jalan memilih dari berbagai alternatif pilihan yang sudah disediakan. Bentuk tes pilihan ganda biasa merupakan bentuk tes yang paling umum digunakan dalam penelitian. Ada empat variasi yang masuk dalam pilihan ganda tersebut yakni (1) pilihan ganda biasa, (2) asosiasi, (3) hubungan antar hal, (4) dan (5) menjodohkan. berikut adalah penjelasan tentang semua bentuk tersebut beserta contohnya.

1) Bentuk Tes pilihan ganda biasa

Contoh soal:1

Diantara bangun-bangun berikut sebutkan bangun mana yang merupakan bangun geometri berdimensi tiga

- a. Lingkaran
- b. Segitiga
- c. Kubus
- d. Trapezium
- e. Persegi panjang

Contoh soal 2

Indonesia selama ini sudah memiliki empat presiden. Urutan dari presiden tersebut dari presiden pertama sampai presiden yang keempat adalah

- a. Soekarno, Soeharto, Habbibie, Gus Dur
- b. Soekarno, Habbibie, Soeharto, Gus Dur
- c. Soekarno, Habbibie, Soeharto, Gus Dur
- d. Soekarno, Gus Dur, Soehato, Habbibie

Petunjuk penyusunan tes pilihan ganda

- Tuliskan dengan jelas petunjuk pengerjaannya
- Tetapkan waktu untuk mengerjakan
- Tuliskan soalnya dan jawabannya dipilih salah satunya di bawah dari soal
- Berikan satu contoh soal untuk cara pengerjaannya
- Kalimat dalam soal diusahakan singkat, jelas dan padat.
- Hindari kalimat negatif dalam kalimat pokoknya
- Setiap satu soal hanya memiliki satu ide
- Usahakan jumlah alternatif jawaban setiap item sama banyaknya
- Perlu dihindari kalimat yang sama dengan kalimat yang ada pada buku teks

2) *Bentuk Tes Hubungan antara Hal*

Tes hubungan antar hal adalah soal yang memuat pernyataan dan alasan, dengan pola memuat pernyataan dan memuat alasan.

Petunjuk Pilihan

- (a) Jika pernyataan benar, alasan benar dan ada hubungan sebab akibat
- (b) Jika pernyataan benar, alasan benar dan tidak ada hubungan sebab akibat
- (c) Jika pernyataan benar, alasan salah
- (d) Jika pernyataan salah, alasan benar
- (e) Baik pernyataan maupun alasan salah

Contoh Soal :

Adanya taufan di kepulauan Filipina selalu diikuti oleh curah hujan cukup besar di Pulau Jawa.

SEBAB

Angin pasat Tenggara tertarik ke Utara Katulistiwa melalui Pulau Jawa, yang menambah banyaknya hujan.

3) *Bentuk Tes Menjodohkan (Completion test)*

Dalam bentuk tradisional item tes menjodohkan terdiri dari dua kolom yang parallel. Tiap kata, bilangan, atau simbol dijodohkan dengan kalimat, frase, atau kata dalam kolom yang lain. Item pada kolom di mana penjodohan dicari disebut premis, sedangkan kolom di mana pilihan dicari disebut respon. Tugas siswa adalah memasangkan antara premis dengan respon berdasarkan aturan yang ditentukan.

Penggunaan item tes menjodohkan terbatas pada mengukur informasi faktual berdasarkan asosiasi sederhana. Apabila hasil belajar menekankan pada pemahaman untuk mengidentifikasi hubungan antara dua hal, dan sejumlah premis dan respon homogen bisa diperoleh maka item tes menjodohkan sangat tepat untuk digunakan.

Contoh :

Premis	Respon
1. Ibu kota Indonesia	a. Athena
2. Tempat penyelenggaraan Olimpiade pertama	b. Jakarta
	c. New York

3. Kota terbesar di dunia	d. Paris
4. Disebut kota mode Dunia	e. New Delhi
	f. Manila

4) *Bentuk Tes Asosiasi*

Bentuk asosiasi merupakan modifikasi dari tes pilihan ganda biasa. Bentuk asosiasi juga terdiri dari satu pernyataan dan beberapa alternatif jawaban, hanya saja terdapat lebih dari satu jawaban yang benar. Salah satu bentuknya adalah dengan mengikuti petunjuk sebagai berikut.

Petunjuk pengerjaan soal :

- Pilihlah a bila jawaban 1,2, dan 3 benar
- Pilihlah b bila jawaban 1 dan 3 benar
- Pilihlah c bila jawaban 2 dan 4 benar
- Pilihlah d bila jawaban 4 saja yang benar

Contoh soal :

Pada kalimat manakah dari kalimat-kalimat berikut subyek dan predikat tidak setuju

1. Bila ia menang maka ia suka
2. Politik tidak bias diprediksi
3. Mayoritas selalu benar
4. Penyebab krisis ekonomi sulit untuk difahami

c) **Tes Essay (Tes Subjectif)**

Tes essay adalah salah satu bentuk tes yang digunakan juga dalam instrument penelitian. Tes ini digunakan biasanya pada subjek penelitian yang sifatnya kecil. Tes ini juga dimaksudkan untuk melihat berbagai kemampuan yang dimiliki subjek dalam bentuk tertulis. Item tes essay menuntut siswa untuk mengorganisasikan atau menyajikan jawaban dalam bentuk uraian (*essay*).

Ada beberapa petunjuk yang perlu diperhatikan dalam membuat tes essay, antara lain :

- 1) Item tes disusun dari yang mudah ke yang sukar
- 2) Pada tiap item dibuat penjenjangan skor menurut tingkatan pengerjaan siswa
- 3) Pengoreksian dilakukan per nomor soal untuk semua siswa.

Tes essay masih diklasifikasikan lagi menjadi beberapa bentuk, yaitu uraian bebas, uraian terstruktur, jawaban singkat, dan isian (melengkapi). Klasifikasi tersebut didasarkan atas kebebasan siswa untuk memberikan jawaban terhadap soal.

1) **Bentuk Uraian Bebas**

Bentuk uraian bebas memberikan kebebasan kepada siswa untuk memberikan opini serta alasan yang diperlukan. Jawaban siswa tidak dibatasi oleh persyaratan tertentu.

Contoh Soal :

Jelaskan pendapat anda tentang apa yang harus dilakukan oleh Bank Indonesia untuk menjaga kestabilan nilai rupiah. Berikan program dan kebijakan yang tepat menurut anda, dan sertakan alasan untuk usulan anda.

2) Bentuk Uraian Terstruktur atau Terbatas

Bentuk uraian terstruktur atau uraian terbatas meminta siswa untuk memberikan jawaban terhadap soal dengan persyaratan tertentu.

Contoh Soal :

Sebutkan dua manfaat media internet bagi pendidikan.

3) Bentuk Jawaban Singkat

Tes jawaban singkat merupakan tipe item tes yang bias dijawab dengan kata, frase, bilangan, atau simbol. Item tes jawaban singkat menggunakan pertanyaan langsung, dan siswa diminta untuk memberi jawaban secara singkat, tepat dan jelas. Item jawaban singkat cocok untuk mengukur bebabagai hasil belajar yang relatif sederhana.

Contoh Soal :

1. Siapakah yang menemukan mesin uap?
2. Dua bilangan ganjil jika dijumlahkan hasilnya bilangan apa ?
3. Apa nama ibukota Negara Republik Indonesia?

Keuntungan item tes jawaban singkat adalah merupakan item tes yang paling mudah dalam penyusunan karena mengukur hasil belajar yang relatif sederhana. Kecuali pada pemecahan masalah-masalah matematika atau ilmu pengetahuan alam. Item-tes jawaban singkat hampir selalu mengukur ingatan.

Petunjuk Penyusunan

- 1) Susun item sedemikian rupa sehingga jawaban yang diinginkan benar-benar singkat dan spesifik.
- 2) Pernyataan item tes tidak diambil langsung dari buku karena pernyataan dalam buku teks sering masih terlalu umum dan ambigu, sehingga memerlukan penjelasan tambahan.
- 3) Bila jawaban menghendaki satuan bilangan maka tipe satuan tersebut harus dijelaskan. Apakah kg, cm, lusin, dan sebagainya.

4) Bentuk Melengkapi (Isian)

Item tes melengkapi hampir sama dengan jawaban singkat, yaitu merupakan tipe item tes yang bias dijawab dengan kata, frase, bilangan, atau simbol. Bedanya, item tes melengkapi merupakan pernyataan yang tidak lengkap, dan siswa diminta untuk melengkapi pernyataan tersebut. Item tes melengkapi juga cocok untuk mengukur bebabagai hasil belajar yang relatif sederhana.

Contoh Soal :

1. Mesin uap ditemukan oleh.....
2. Dua bilangan ganjil jika dijumlahkan hasilnya adalah bilangan.....
3. Ibukota Negara Republik Indonesia adalah.....

Item tes melengkapi adalah merupakan item tes yang paling mudah dalam penyusunan karena mengukur hasil belajar yang relatif sederhana. Kecuali pada pemecahan masalah-masalah matematika atau ilmu pengetahuan alam, item tes jawaban singkat hampir selalu mengukur ingatan. Keuntungan item tes jawaban singkat adalah siswa harus memberikan jawaban sehingga menghindari tindakan siswa menebak jawaban.

Petunjuk Penyusunan

- 1) Susun item sedemikian rupa sehingga jawaban yang diinginkan benar-benar singkat dan spesifik.
- 2) Pernyataan item tes tidak diambil langsung dari buku karena pernyataan dalam buku teks sering masih terlalu umum dan ambigu, sehingga memerlukan penjelasan tambahan.
- 3) Bila jawaban menghendaki satuan bilangan maka tipe satuan tersebut harus dijelaskan. Apakah kg, cm, lusin, dan sebagainya.
- 4) Hindari pernyataan soal yang memuat lebih dari satu bagian kosong

Prosedur Penyusunan Instrumen Tes

Prosedur penyusunan instrument tes menurut Norman E. Gronlund & Robert L. Linn (1990: 109) adalah sebagai berikut :

- a. Mengembangkan spesifikasi tes, yang meliputi :
 - Menyusun tujuan khusus pembelajaran
 - Mengadakan pembatasan terhadap bahan yang akan diteskan
 - Menyusun kisi-kisi tes, yang memuat pokok materi, tujuan instruksional khusus, dan aspek berpikir yang diukur. Selanjutnya ditentukan banyak item tes untuk masing-masing tujuan instruksional khusus pada masing-masing domain.
- b. Memilih bentuk tes yang tepat
- c. Menulis item-item tes berdasarkan kisi-kisi yang sudah disusun.

Agar lebih jelas maka berikut ini disajikan contoh penyusunan instrument pada bab berikut:

Revisi Domain Kognitif.

1. Soal Mengingat (*Remembering item*)

Buatlah daftar jenis makanan dan minuman yang dapat kamu beli dengan harga Rp 5000,-, Rp 50000,- dan Rp. 100.000,-!

2. Soal Memahami (*Understanding item*)

Jelaskan besaran uang rupiah yang dapat digunakan untuk membayar barang-barang tersebut!

3. Soal Menerapkan (*Applying item*)

Hitunglah kembalian yang kamu terima jika uangmu Rp 5000,-, Rp 50.000,- atau Rp 100.000,- untuk makanan/minuman yang kamu beli!.

4. Soal Menganalisa (*Analysing item*)

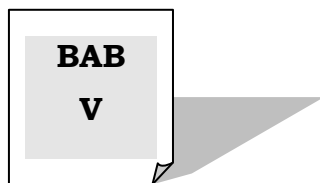
Tentukan dan catat operasi hitung apa yang kamu gunakan untuk menghitung kembalian tersebut!

5. Soal Mengevaluasi (*Evaluating item*)

Kriteria apa yang kamu gunakan untuk mengetahui apakah jawabanmu benar atau salah?.

6. Soal Menciptakan (*Creating item*)

Buatlah daftar pesanan makanan yang terdiri dari 3 macam makanan yang harganya mendekati atau seharga Rp 2.500,-, Rp 7.500,-, dan Rp 25.000,-. Hitung harga total pesananmu!, jika kamu diberikan uang sebesar Rp 100.000,- hitung uang kembalinya!.



PENGEMBANGAN INSTRUMEN RANAH AFEKTIF

A. Pengantar

Menurut Popham (1995), ranah afektif menentukan keberhasilan belajar seseorang. Orang yang tidak memiliki minat pada pelajaran tertentu sulit untuk mencapai keberhasilan belajar secara optimal. Oleh sebab itu, semua pendidik harus mampu membangkitkan minat semua peserta didik untuk mencapai kompetensi yang telah ditentukan. Kebiasaan guru pada awal pembelajaran untuk mengingatkan siswa tentang kompetensi apa sebenarnya yang diharapkan dikuasai siswa setelah penyampaian bahan ajar adalah kebiasaan yang baik. Sambil menyampaikan kepada siswa kompetensi apa yang diharapkan dikuasai siswa terkait bahan ajar tertentu, guru harus berupaya membangkitkan minat siswa untuk meraih kompetensi tersebut. Syarat awal untuk mampu membangkitkan minat adalah adanya kedekatan antara guru dengan siswa. Oleh sebab itu, menjalin kedekatan antara guru dengan para pebelajar secara profesional merupakan syarat utama keberhasilan belajar. Antara guru dengan peserta didik harus dijalin ikatan emosional sehingga kebersamaan sebagai pebelajar kembar (pembelajar-pebelajar), guru sebagai pembelajar, siswa sebagai pebelajar yang belajar bersama dapat terjalin dengan baik.

Keberhasilan pembelajaran pada ranah kognitif dan psikomotor dipengaruhi oleh kondisi afektif peserta didik. Peserta didik yang memiliki minat belajar dan sejumlah sikap positif sesuai yang diungkapkan oleh Popham di atas, akan merasa senang bahkan tertantang untuk mempelajari mata pelajaran tertentu sehingga dapat mencapai keberhasilan pembelajaran yang optimal. Berkenaan dengan itu untuk mencapai hasil belajar yang optimal, dalam merencanakan program pembelajaran (misalnya dalam menyusun silabus dan RPP) dan kegiatan pembelajaran, pendidik harus memperhatikan karakteristik afektif peserta didik.

Anderson (1981:4) menyatakan bahwa pemikiran atau perilaku harus memiliki dua kriteria agar dapat diklasifikasikan sebagai ranah afektif. Pertama, perilaku melibatkan perasaan dan emosi seseorang. Kedua, perilaku harus merupakan tipikal seseorang. Kriteria lain yang berkenaan dengan ranah afektif adalah intensitas, arah, dan sasaran. Intensitas menyatakan derajat atau kekuatan perasaan. Arah perasaan berkaitan dengan orientasi positif atau negatif dari perasaan yang menunjukkan apakah perasaan itu baik atau buruk, sedangkan sasaran mengacu kepada objek, aktivitas, atau ide sebagai arah dari perasaan.

B. Penerapan Pengukuran dalam Domain

1. Karakteristik Domain Afektif yang Penting

Paling tidak ada lima karakteristik afektif yang penting, yaitu sikap, minat, konsep diri, nilai, dan moral. Sejumlah ahli menambahkan beberapa aspek lagi terkait karakteristik afektif yang juga layak diperhatikan, antara lain adalah preferensi (pertimbangan baik dan buruk), kontrol diri, pengembangan emosi, lingkungan kelas, opini, motivasi, hubungan sosial, dan altruisme.

1. Sikap (*attitude*)

Menurut definisi konseptual, sikap merupakan kecenderungan merespons secara konsisten tentang menyukai atau tidak menyukai suatu objek. Sikap bisa positif atau negatif. Definisi operasional; sikap adalah perasaan positif atau negatif terhadap suatu objek. Dalam hal ini yang dimaksud dengan definisi konseptual adalah definisi yang mengacu pada prinsip atau konsep dari objek kajian yang bersangkutan, sedangkan definisi operasional merupakan penerapan definisi konseptual dalam pembelajaran.

Definisi lain seperti dinyatakan oleh Secord dan Beckman (1964, dalam Azwar 2012:5) yang menyatakan sikap adalah keteraturan tertentu dalam hal perasaan (afeksi), pemikiran (kognisi), dan predisposisi tindakan (konasi) seseorang terhadap suatu aspek di lingkungan sekitarnya. Konasi merupakan predisposisi dari tindakan karena konasi identik dengan hasrat, kehendak, dan kemauan. Definisi Secord dan Beckman ini berorientasi kepada skema triadik (*triadic scheme*) dengan asumsi sikap merupakan konstelasi komponen-komponen kognitif, afektif, dan konatif yang saling berinteraksi dalam memahami, merasakan, dan berperilaku terhadap suatu objek. Pengertian triadik adalah suatu kelompok yang terdiri dari tiga hal/aspek.

Perubahan sikap dapat diamati dalam proses pembelajaran, keteguhan, dan konsistensi terhadap sesuatu. Penilaian sikap adalah penilaian yang dilakukan untuk mengetahui sikap peserta didik terhadap mata pelajaran, kondisi pembelajaran, pendidik, dan sebagainya.

2. Minat (*interest*)

Menurut definisi konseptual minat adalah keinginan yang terbentuk melalui pengalaman yang mendorong individu mencari objek, aktivitas, konsep dan keterampilan, untuk tujuan mendapatkan perhatian atau penguasaan. Menurut definisi operasional minat adalah keingintahuan seseorang tentang keadaan suatu objek.

Definisi lain seperti dinyatakan oleh Getzel (1996) yang mengatakan minat adalah suatu disposisi yang terorganisir melalui pengalaman yang mendorong seseorang untuk memperoleh objek khusus, aktivitas, pemahaman, dan keterampilan untuk tujuan perhatian atau pencapaian. Silvia (2006) mendefinisikan minat sebagai suatu perasaan atau emosi yang menimbulkan perhatian kepada sesuatu objek, kejadian, atau proses. Menurut Kamus Besar Bahasa Indonesia (1989: 583), minat atau keinginan adalah kecenderungan hati yang tinggi terhadap sesuatu. Minat juga didefinisikan sebagai perasaan seseorang yang perhatiannya, kepeduliannya, dan rasa ingin tahunya terikat secara khusus pada sesuatu. Secara umum minat termasuk karakteristik afektif yang memiliki intensitas tinggi.

Terkait dengan implementasi Kurikulum 2013, kita tahu bahwa sejak kelas X siswa SMA sudah boleh menentukan peminatannya sendiri. Mereka akan masuk dalam jurusan matematika dan sains, jurusan IPS, atau jurusan bahasa, bergantung pada pilihannya masing-masing. Sebaliknya pihak sekolah jangan melepaskan begitu saja kepada siswa pemilihan minat ini. Berdasarkan evaluasi terhadap kuesioner tentang peminatan yang diisi oleh siswa, pihak sekolah dapat memberikan saran, sebaiknya jurusan apa yang harus dipilih oleh siswa yang bersangkutan. Hasil kuesioner ini dapat memperkuat hasil tes bakat dan minat yang dilaksanakan oleh psikolog yang diminta pihak sekolah. Siswa lulusan SMP belum mampu mengamati dengan baik minat dan bakatnya sendiri.

3. Nilai (value)

Menurut definisi konseptual, nilai adalah keyakinan terhadap suatu pendapat, kegiatan, atau objek. Menurut definisi operasional, nilai adalah keyakinan seseorang tentang keadaan suatu objek atau kegiatan. Definisi nilai dalam dunia psikologi adalah pertimbangan tentang seberapa penting sesuatu itu bagi kita.

Nilai amat penting dalam konstruksi moralitas personal dan sebagai landasan untuk hidup dalam kehidupan. Nilai seseorang pada dasarnya terungkap melalui bagaimana ia berbuat. Nilai berkaitan dengan keyakinan, sikap, dan aktivitas atau tindakan seseorang. Tindakan seseorang terhadap sesuatu merupakan refleksi dari nilai yang dianutnya.

Dalam kaitannya dengan pembelajaran, nilai merupakan konsep penting bagi pembentukan kompetensi peserta didik. Aktivitas yang disukai peserta didik di sekolah dipengaruhi oleh penilaian peserta didik terhadap aktivitas tersebut. Hal ini dipengaruhi oleh sistem nilai yang dimiliki peserta didik, berkaitan dengan penilaian baik dan buruk.

Menurut Rokeach (1968) dalam publikasinya berjudul *Beliefs, Attitudes, and Values: A Theory of Organization and Change*, nilai merupakan suatu keyakinan tentang perbuatan, tindakan, atau perilaku yang dianggap baik dan yang dianggap buruk. Berdasarkan pendapatnya tersebut, Rokeach kemudian membuat apa yang disebutnya sebagai *Rokeach Value Survey (RVS)*. Disini Rokeach mengelompokkan sistem nilai menjadi dua, yaitu nilai-nilai terminal (*terminal values*) dan nilai-nilai instrumental (*instrumental values*). Nilai-nilai terminal mengacu kepada keadaan akhir yang diinginkan. Ini adalah suatu tujuan yang dikehendaki seseorang untuk dicapai sepanjang masa hidupnya. Nilai-nilai ini amat bervariasi di antara kelompok manusia yang memiliki budaya yang berbeda. Termasuk dalam nilai-nilai terminal antara lain harga diri, kebahagiaan, kebebasan, kesenangan, kebijakan, dan harmoni.

Nilai-nilai instrumental mengacu pada moda perilaku yang lebih disukai. Moda perilaku adalah cara untuk meraih nilai-nilai terminal. Termasuk ke dalam nilai-nilai instrumental antara lain kasih sayang, kejujuran, kesantunan, kecerdasan, kepatuhan, logika, kerajinan, kemauan memberikan maaf, pertanggungjawaban, imajinasi, kontrol diri, berpandangan luas, dan kemauan untuk membantu.

Dikenal ada tiga atribut dasar dari nilai. Pertama, nilai adalah pertimbangan subjektif. Kedua, kita mungkin dapat membuat suatu pertimbangan tentang seberapa penting sesuatu relatif terhadap sesuatu yang lain. Misalnya, seseorang mungkin memertimbangkan bahwa membantu orang lain itu lebih penting daripada menjadi terkenal. Ketiga, nilai dapat merupakan prinsip yang membantu kita membuat keputusan pribadi yang penting. Ketiga hal tersebut dapat dipandang sebagai nilai-nilai personal. Contohnya, jika seseorang secara personal menilai kreativitas itu tinggi, mungkin ia akan berusaha sekuat tenaga untuk menemukan karier yang memungkinkan dia menjadi kreatif, atau bebas mengembangkan kreativitasnya.

Berkenaan dengan itu, Hungerford, Volk, dan Ramsey (1994) memberikan deskripsi terhadap sejumlah nilai menurut wilayah kajian tempat nilai tersebut didefinisikan. Hal itu dapat dilihat pada tabel di bawah ini.

Tabel 8. Deskripsi Nilai Menurut Hungerford, Volk, dan Ramsey

No	Wilayah Pendefinisian	Deskripsi
1	Estetika	Nilai adalah apresiasi terhadap bentuk, komposisi, warna dan suara melalui kegiatan alat indra
2	Pendidikan	Terkait dengan akumulasi, penggunaan, dan komunikasi dari pengetahuan.
3	Budaya	Berkenaan dengan kelangsungan/pelestarian pengetahuan manusia, kepercayaannya, seni, tradisi dan lain-lain.
4	Etika/Moral	Berkenaan dengan pertanggungjawaban pada saat ini dan pada masa mendatang, terkait hal yang benar dan salah, dan standar etika.
5	Sosial	Berkenaan dengan saling berbagi empati dan perasaan manusia, dan status, serta interaksi antar manusia.
6	Sains	Berkenaan dengan atribut yang terkait dengan empirisme dan riset empiris.
7	Agama	Penerapan sistem kepercayaan berlandaskan iman dan dogma.

Sumber: Hungerford, Volk, dan Ramsey, 1994.

Sasaran nilai cenderung menjadi ide, sasaran nilai dapat juga berupa sesuatu seperti sikap dan perilaku. Arah nilai dapat positif dan dapat pula negatif. Intensitas nilai dikatakan tinggi atau rendah tergantung pada situasi dan nilai yang diacu.

Sumber lain menyatakan bahwa nilai-nilai yang juga harus diajarkan kepada siswa adalah nilai-nilai kehidupan yang penting. Nilai-nilai kehidupan (*living Values*) yang penting itu antara lain adalah:

- Kedamaian (*peace*); suatu nilai yang menjunjung tinggi perdamaian dan menjauhi permusuhan dalam rangka menegakkan salah satu pilar pembelajaran UNESCO, yaitu belajar untuk hidup bersama-sama (*learning to live together*).
- Penghormatan (*respect*) atau menghargai; nilai yang menjunjung tinggi hak dan kewajiban orang lain di samping memperhatikan hak dan kewajiban diri sendiri.
- Kerja sama (*cooperation*); bekerja sama dengan orang lain untuk mencapai suatu tujuan dan manfaat bersama. Nilai ini amat penting ditanamkan dalam diri siswa-siswa Indonesia yang amat heterogen dalam ras, suku, bahasa, dan agama.
- Kebebasan (*freedom*); melakukan aksentuasi kebebasan diri sendiri dengan menghargai kebebasan orang lain, kebebasan yang bertanggung jawab dengan memedulikan aspek nilai norma, dan tradisi yang berlaku di masyarakat.
- Kebahagiaan (*happiness*); suatu nilai yang menjunjung tinggi tercapainya rasa bahagia karena merasa kehadirannya bermakna, baik bagi diri sendiri maupun orang lain.
- Kejujuran (*honesty*); nilai yang menjunjung tinggi sikap yang terbiasa mengatakan apa yang benar-benar dilakukan dan benar-benar terjadi, menghargai milik dan hak orang lain.
- Amanah (*trustworthy*); nilai yang menjunjung tinggi sifat yang dapat dipercaya dan mampu memercayai orang lain.
- Kerendahan hati (*humility*) nilai yang menjunjung tinggi sikap tidak merasa angkuh dan sombong, dapat menempatkan orang lain apa adanya.

- Kasih sayang (*love*); suatu nilai yang mendorong orang untuk menyayangi dan memedulikan orang lain.
- Tanggung jawab (*responsibility*) nilai yang mendorong orang untuk berani bertanggung jawab terhadap semua perbuatannya.
- Kesederhanaan (*simplicity*); nilai yang mendorong orang untuk tidak bersikap berlebihan dalam penampilan diri, tampil apa adanya dengan penuh percaya diri.
- Toleransi (*tolerance*); nilai yang membuat orang terbiasa bertenggang rasa, menghormati dan tidak mau menyakiti hati orang lain, merasa senasib sepenanggungan.
- Kesatuan (*unity*); merasa sebagai satu bangsa, sebagai orang-orang yang hidup dalam tanah air yang sama.
- Empati (*empathy*); nilai yang mendorong orang untuk mampu merasakan penderitaan orang lain sehingga tergerak hatinya untuk meringankan beban penderitaannya.
- Serta sejumlah nilai lainnya, misalnya kebersamaan (*togetherness*) dan lainnya.

4. Moral

Dalam bidang psikologi, moral didefinisikan sebagai kemampuan untuk membedakan apakah suatu tindakan atau kejadian itu baik atau buruk, dan atau benar atau salah. Penalaran moral adalah suatu proses untuk menentukan benar atau salah dari suatu situasi tertentu.

Moral berkaitan dengan perasaan salah atau benar terhadap kebahagiaan orang lain atau perasaan terhadap tindakan yang dilakukan diri sendiri. Moral juga sering dikaitkan dengan keyakinan agama seseorang, yaitu keyakinan akan perbuatan berdosa dan berpahala. Moral terutama berkaitan dengan prinsip, nilai, dan keyakinan seseorang. Moral berbeda dari moralitas. Moralitas berkenaan dengan suatu tindakan berlandaskan moral. Moralitas adalah suatu sistem atau kepercayaan yang menyakini bahwa suatu tindakan itu harus berlandaskan moral dan kemudian dinilai apakah tindakan tersebut benar atau salah, baik atau buruk.

Dalam pembelajaran, moral dan moralitas berkenaan dengan perilaku siswa dalam memaknai kejujuran. Melalui perangkat moral atau lebih tepatnya karakter, seorang siswa akan menilai baik dan buruknya perbuatan curang itu, dan kemudian meyakini untuk tidak berbuat curang dalam suatu ujian, misalnya mencontek buku catatan/diktat/buku teks, atau mencontoh pekerjaan temannya (*cheating*). Mulai tahun 2010 pendidikan karakter digalakkan penerapannya di Indonesia. Perihal pendidikan karakter, Anda dapat melihat selengkapnya dalam buku berjudul Pendidikan Karakter, Konsep Dan Implementasi yang ditulis oleh Prof. dr. Muchlas Samani dan Drs. Hariyanto, M.S., yang diterbitkan oleh PT. Remaja Rosdakarya.

5. Konsep diri

Menurut definisi konseptual, konsep diri merupakan persepsi seseorang terhadap dirinya sendiri yang menyangkut keunggulan dan kelemahannya. Menurut definisi operasional, konsep diri adalah pernyataan tentang kemampuan diri sendiri yang menyangkut mata pelajaran. Baumeister (1999) mendefinisikan konsep diri sebagai kepercayaan individu tentang dirinya termasuk atribut personal tentang siapa dan apa dirinya. Carl Rogers (1959) memercayai bahwa konsep diri memiliki seseorang tentang dirinya (*self image*, citra diri), seberapa banyak nilai yang ada dalam diri (*self-esteem*, harga diri atau nilai diri), dan seperti apa harapan kita tentang diri kita nantinya (*ideal self*, diri ideal). Konsep diri pada hakikatnya merupakan evaluasi yang dilakukan individu terhadap kemampuan dan kelemahan yang dimilikinya. Sasaran, arah, dan intensitas konsep diri pada dasarnya seperti ranah afektif yang lain. Sasaran konsep diri biasanya orang,

tetapi bisa saja sebuah lembaga seperti sekolah. Arah konsep diri bisa positif bisa negative, intensitasnya dalam kontinum mulai dari rendah sampai tinggi.

Konsep diri penting untuk menentukan jenjang karier peserta didik karena dengan mengetahui kekuatan dan kelemahan dirinya dapat dipilih alternative karier yang tepat baginya. Kecuali itu, informasi tentang konsep diri penting bagi sekolah untuk memberikan motivasi belajar kepada peserta didik dengan tepat.

Tabel 9. Kata Operasional pada Ranah Afektif

Domain	Kategori Jenis Perilaku	Kemampuan Internal	Kata Kerja Operasional
Afektif	Penerimaan	Menunjukkan Misalnya : -kesadaran -kemauan -Perhatian Mengakui... Misalnya : Kepentingan Perbedaan	Menerima Menantang/Tantangan Mendengar Menanyakan Memilih Mengikuti Menjawab Melanjutkan Memberi Menyatakan Menempatkan
	Partisipasi	Mematuhi..... Misalnya: Peraturan Tuntutan Perintah Ikut serta secara aktif Misalnya : Di laboratorium Dalam diskusi Dalam kelompok Tentir	Mempertahankan Memperdebatkan Bergabung Melaksanakan Membantu Menawarkan diri Menyambut Menolong Mendatangi Melaporkan Menyumbangkan Menyesuaikan diri Berlatih Menampilkan Membawakan Mendiskusikan Menyelesaikan Menyatakan persetujuan Mempraktekkan
	Penilaian/Penentuan sikap	Menerima suatu hal Menyukai Menyepakati Menghargai Misalnya : Karya seni Sumbangan ilmu Pendapat Bersikap (positif atau negative) Mengakui	Memutuskan Menawarkan Memuji Berpendapat Menunjukkan Melaksanakan Menyatakan pendapat Mengikuti Mengambil prakarsa Memilih Ikut serta Menggabungkan diri Mengundang Mengusulkan

			Membela Menuntun Membenarkan Menolak Mengajak
	Organisasi	Membentuk system nilai Menangkap relasi antara nilai Bertanggungjawab Mengintegrasi nilai	Merumuskan Membagi Mendukung Mengintegrasikan Menghubungkan Mengaitkan Menyusun Mengubah Melengkapi Menyempurnakan Menyesuaikan Menyamakan Mengatur Memperbandingkan Mempertahankan Memodifikasikan
	Pembentukan pola	Menunjukkan Misalnya : Kepercayaan diri Disiplin pribadi Kesadaran Mempertimbangkan Melibatkan diri	Mengunjungi Berbuat sukarela Bersikap konstan Bertindak Menyatakan Memperlihatkan Mempratekkan Melayani Mengundurkan diri Membuktikan Menunjukkan Bertahan Mempertimbangkan Mempersoalkan

Tabel 10. Kata Operasional Ranah Afektif

Menerima (A1)	Merespon (A2)	Menghargai (A3)	Mengorganisasikan (A4)	Karakterisasi Menurut Nilai (A5)
Mengikuti Menganut Mematuhi Meminati	Mengompromikan Menyenangi Menyambut Mendukung Menyetujui Menampilkan Melaporkan Memilih Mengatakan Memilah Menolak	Mengasumsikan Meyakini Meyakinkan Memperjelas Memprakarsai Mengimani Menekankan Menyumbang	Mengubah Menata Mengklasifikasikan Mengombinasikan Mempertahankan Membangun Membentuk pendapat Memadukan Mengelola Menegosiasi Merembuk	Membiasakan Mengubah perilaku Berakhlak mulia Mempengaruhi Mengkualifikasi Melayani Membuktikan Memecahkan

C. Pengembangan Instrumen Ranah Afektif

Instrumen penilaian afektif dapat berupa kuesioner atau lembar hasil observasi. Instrumen penilaian afektif meliputi lembar pengamatan sikap, minat, nilai, moral, dan konsep diri. Ada 11 (sebelas) langkah dalam mengembangkan instrumen penilaian afektif, yaitu: (1) menentukan spesifikasi instrumen; (2) menulis instrumen, (3) menentukan skala instrumen, (4) menentukan pedoman pemberian skor, (5) menelaah instrumen, (6) merakit instrumen, (7) melaksanakan uji coba, (8) menganalisis hasil uji coba, (9) memperbaiki instrumen, (10) melaksanakan pengukuran, dan (11) menafsirkan hasil pengukuran.

Sesuai dengan uraian tentang ranah afektif, ada lima instrumen pengukuran ranah afektif. Kita akan membahasnya satu per satu.

a. Instrumen sikap

Instrumen ini bertujuan untuk mengetahui sikap peserta didik terhadap suatu obyek, mata pelajaran, metode pembelajaran, pendidik, bahan ajar, dan sebagainya. Sikap terhadap mata pelajaran bisa positif maupun negatif. Hasil pengukuran sikap berguna untuk menentukan strategi pembelajaran yang sesuai.

b. Instrumen minat

Instrumen ini bertujuan untuk memperoleh informasi tentang minat peserta didik terhadap mata pelajaran, yang selanjutnya digunakan untuk meningkatkan minat peserta didik terhadap mata pelajaran.

c. Instrumen nilai

Instrumen ini bertujuan untuk mengungkap nilai dan keyakinan peserta didik. Informasi yang diperoleh berupa nilai dan keyakinan, baik yang positif maupun yang negatif.

d. Instrumen moral

Instrumen ini bertujuan untuk mengungkap moral. Informasi tentang moral seseorang diperoleh melalui pengamatan terhadap perbuatan yang ditampilkan, maupun hasil laporan evaluasi diri melalui pengisian kuesioner. Hasil pengamatan dan hasil kuesioner merupakan informasi tentang moral seseorang.

e. Instrumen konsep diri

Instrumen ini bertujuan untuk mengetahui kekuatan dan kelemahan diri siswa sendiri. Peserta didik melakukan evaluasi secara objektif terhadap potensi yang dimilikinya. Karakteristik potensi peserta didik amat penting untuk menentukan jenjang kariernya. Dalam pembelajaran, informasi kekuatan dan kelemahan peserta didik digunakan untuk menentukan program yang sebaiknya ditempuh.

Dalam menulis instrumen perlu diperhatikan kisi-kisi instrumen (*grid* atau *blue-print*) Kisi-kisi instrumen merupakan matriks yang berisi spesifikasi instrumen yang akan ditulis. Kisi-kisi instrumen merupakan acuan bagi penulisan instrumen.

Contoh kisi-kisi instrumen penilaian afektif terlihat dalam tabel berikut.

Tabel 11. Contoh Kisi-Kisi Instrumen Penilaian Afektif

No	Indikator	Jumlah butir	Nomor Pertanyaan/ Pernyataan	Skala
1.	Empati	5	1,2,3,4,5	<i>Likert</i>
2	Peduli	6	6,7,8,9,10,11	<i>Gutman</i>
3	Sayang	7	12,13,14,15,16,17,18	<i>Bogardus</i>

4	Toleransi	5	19,20,21,22,23	Semantic Diffrencial
5	Jujur	7	24,25,26,27,28,29,30	Likert

Catatan: Indikator dijabarkan dari definisi operasional. Indikator merupakan pedoman dalam penulisan instrument. Tiap indikator bisa dikembangkan menjadi dua atau lebih instrument (jumlah instrumen yang dikembangkan dari indikator dinyatakan dalam jumlah butir).

D. Skala Instrumen Penilaian Afektif dan Sistem Pemberian Skorya

Skala yang sering digunakan dalam instrurnen penilaian afektif adalah skala Thurstone, skala Likert, dan skala beda semantik (*semantic differential scaling*), dan skala Guttman.

Skala Thurstone merupakan teknik formal yang pertama bagi pengukuran suatu sikap kala ini, dikembangkan oleh Louis L. Thurstone pada tahun 1928. Skala Thurstone terdiri dari sejumlah pernyataan dalam hal mana responden ditanya tentang setuju atau tidak setujunya. Dalam skala Thurstone penilai memberikan tanda centang (✓) di bawah kolom setuju (*agree*) atau tidak setuju (*disagree*) atau di bawah kolom angka yang menggambarkan kontinum, dari yang dianggap paling sesuai dengan pernyataan sampai yang dianggap tidak sesuai dengan pernyataan.

Contoh skala Thurstone: minat dalam pelajaran kimia dalam bentuk yang sederhana.

Pernyataan	setuju	tak setuju
1 Saya suka belajar kimia.		
2 Pelajaran kimia bermanfaat.		
3 Saya berusaha hadir setiap ada pelajaran kimia.		
4 Saya berusaha memiliki buku pelajaran kimia.		
5 Pelajaran kimia membosankan.		
6. Saya tidak senang kepada cara guru kimia mengajar.		
dan seterusnya....		

Contoh yang lebih kompleks untuk pernyataan yang sama.

Pernyataan	7	6	5	4	3	2	1
1. Saya suka belajar kimia.							
2. Pelajaran kimia bermanfaat.							
3. Saya berusaha hadir setiap ada pelajaran kimia.							
4. Saya berusaha memiliki buku pelajaran kimia							
5. Pelajaran kimia membosankan.							
6. Saya tidak senang kepada cara guru kimia mengajar.							
dan seterusnya....							

Skala Likert adalah suatu skala psikometrik yang umum digunakan dalam riset yang menggunakan kuesioner sebagai instrumen pengambilan data. Butir pernyataan Likert

merupakan pernyataan sederhana dalam hal mana responden diminta untuk mengevaluasi menurut setiap jenis kriteria subjektif atau objektif, kemudian tingkat kesetujuan atau ketidaksetujuannya diukur. Skala Likert disukai yang berbentuk simetris atau seimbang, dengan butir skala ganjil, terdapat jumlah yang sama dari posisi pilihan positif maupun posisi negatif. Di antara kedua posisi tersebut ada posisi tidak tahu (entah). Skala Likert adalah metode pemberian skala bipolar, mengukur baik respons positif maupun respons negatif terhadap suatu pernyataan. Format yang umum adalah format yang menggunakan lima pilihan yang meliputi:

1. Sangat tidak setuju (STS). 4. Setuju (S)
2. Tidak setuju (TS). 5. Sangat Setuju (SS)
3. Entah (E).

Contoh skala Likert yang umum dipergunakan adalah sebagai berikut.

Pernyataan	SS	S	E	TS	STS
1. Saya suka belajar kimia.					
2. Pelajaran kimia bermanfaat.					
3. Saya berusaha hadir setiap ada pelajaran kimia.					
4. Saya berusaha memiliki buku pelajaran kimia					
5. Pelajaran kimia membosankan					
6. Saya tidak senang kepada cara guru kimia mengajar.					
dan seterusnya.....					

Namun, kadang-kadang juga digunakan butir skala yang genap sehingga pilihan "entah" atau "tidak memilih", setuju "maupun" "tak setuju" dihilangkan. Hal ini memaksa responden untuk memilih (*forced choice method*) karena pilihan netral, sehingga menjadi:

- Sangat tidak setuju (STS). Setuju (S)
Tidak setuju (TS). Sangat Setuju (SS)

Contoh skala Likert yang memaksa responden untuk bersikap memilih.

Pernyataan	SS	S	TS	STS
Saya suka belajar kimia.				
Pelajaran kimia bermanfaat.				
Saya berusaha hadir setiap ada pelajaran kimia.				
Saya berusaha memiliki buku pelajaran kimia				
Pelajaran kimia membosankan				
Saya tidak senang kepada cara guru kimia mengajar.				
dan seterusnya.....				

Skala beda semantik adalah skala yang mengharuskan responden memilih di antara dua kutub yang berlawanan. pilihannya mendekati kutub mana dan dinyatakan dengan memberikan tanda cek di bawah kolom yang dipilih. Skala ini dikembangkan oleh Charles E. Osgood.

Contoh skala beda semantik (*semantic differential scaling*).

Pelajaran kimia									
	a	b	c	d	e	f	g	h	I
Menyenangkan									Membosankan
Sukar									Mudah
Bermanfaat									Sia-sia
Menantang									Menjemukan

Contoh skala beda semantik dapat ditampilkan secara lebih sederhana dengan semacam garis kontinum.

Pelajaran kimia							
						X	
Membosankan						Menyenangkan	

Pada contoh di atas, responden memilih pelajaran kimia itu menyenangkan (memberikan tanda X), tetapi belum sampai sangat menyenangkan yang ada pada garis kontinum paling kanan.

Skala Guttman adalah skala yang dikembangkan oleh Louis Guttman yang digunakan jika penilai menginginkan untuk merancang kuesioner singkat yang mampu membuat perbedaan kemampuan dari responden. Dalam skala Guttman, ada pengertian bahwa pertanyaan/pernyataan yang lebih atas membawahi pertanyaan/ pernyataan di bawahnya.

Contohnya adalah sebagai berikut:

(Paling khusus)

1. Apakah Anda ingin belajar unsur Fluor
2. Apakah Anda ingin belajar unsur-unsur golongan halogen?
3. Apakah Anda ingin belajar sistem periodik?
4. Apakah Anda ingin belajar ilmu kimia?

(Paling umum)

Keterangan:

jika Anda setuju dengan pertanyaan nomor 2, berarti Anda setuju dengan 3 dan nomor 4.

E. Telaah Instrumen dan Perakitannya

Tujuan telaah instrumen adalah untuk menjawab sejumlah pertanyaan.

- a. Apakah butir pertanyaan/pernyataan sesuai dengan indikator?
- b. Apakah bahasa yang digunakan komunikatif dan menggunakan bahasa Indonesia yang baku?
- c. Apakah butir pertanyaan/pernyataan tidak bias?
- d. Apakah format instrumen menarik untuk dibaca?
- e. Apakah petunjuk untuk menjawab atau mengisi instrumen jelas?
- f. Apakah jumlah butir dan/atau panjang kalimat pertanyaan/pernyataan sudah sesuai sehingga tidak menjemukan untuk dibaca/dijawab?

Sejumlah hal yang perlu diperhatikan dalam menggunakan kata-kata untuk kuesioner, antara lain:

- Gunakan kata-kata sederhana sesuai dengan tingkat pendidikan responden.
- Pertanyaan/ Pernyataan harus cukup jelas dan tegas, jangan samar-samar.
- Hindari pertanyaan/ pernyataan yang bias atau bermakna ganda.
- Hindari pertanyaan/ pernyataan hipotetik atau bersifat pengandaian.

Hasil telaah instrumen digunakan untuk memperbaiki instrumen. Hal yang perlu diperbaiki antara lain konstruksi instrumen, yaitu kalimat yang digunakan, waktu yang diperlukan untuk mengisi instrumen, cara pengisian atau cara menjawab instrumen, dan revisi terhadap hasil pengetikan.

F. Uji Coba Instrumen dan Analisis Hasil Uji Coba

Setelah dirakit, instrumen diujicobakan kepada responden sesuai dengan tujuan penilaian. Apakah instrumen ditujukan kepada guru, peserta didik, orang tua, atau staf non-pendidik di sekolah. Untuk itu dipilih sampel yang karakteristiknya mewakili populasi responden yang akan dinilai. Agar jumlah sampel yang diperlukan cukup representatif untuk dianalisis, minimal ada 30 responden, dapat berasal dari satu sekolah atau lebih. Misalnya ada 3 SMA yang dipilih, masing-masing diambil 10 responden.

Pada saat uji coba, yang perlu diperhatikan adalah saran-saran dari responden atas kejelasan pedoman pengisian instrumen, kejelasan kalimat yang digunakan, serta waktu yang diperlukan untuk mengisi instrumen. Waktu yang dipilih untuk pengambilan data disarankan bukan waktu saat responden sudah lelah. Perlu diingat bahwa pengisian instrumen penilaian afektif bukan merupakan suatu tes sehingga walau ada batasan waktu tetapi boleh agak longgar. Berdasarkan pengalaman, waktu yang diperlukan untuk mengisi kuesioner agar responden tidak jenuh dan lelah adalah 30 menit atau kurang. Dengan mengetahui patokan waktu ini maka dapat diperkirakan panjang instrumen.

Analisis hasil uji coba meliputi analisis terhadap variasi jawaban tiap butir pertanyaan/ pernyataan. Jika menggunakan skala instrumen 1 sampai 5, dan jawaban responden bervariasi dari 1 sampai 5, maka butir pertanyaan/ pernyataan pada instrumen dapat dikatakan baik. Namun apabila jawaban responden hanya pada satu pilihan jawaban saja, misalnya pada pilihan nomor 3, maka butir instrumen itu tergolong tidak baik.

Analisis hasil uji coba penilaian skala afektif juga bersifat kuantitatif. Indikator yang digunakan untuk mengetahui kebaikan suatu pertanyaan adalah daya beda (discrimination index). Jika daya beda butir instrument lebih dari 0,30 maka butir instrumen tergolong baik. Indikator lainnya yang digunakan untuk menilai kebaikan instrumen penilaian berupa skala sikap adalah indeks keandalan (reliability index). Batas indeks keandalan minimal 0,70. Jika indeks reliabilitas < 0,70, kesalahan pengukuran akan melebihi batas.

Indeks reliabilitas atau lebih sering disebut koefisien reliabilitas dapat diukur misalnya dengan menggunakan rumus Spearman-Brown, rumus Rulon, dan rumus alfa (Cronbach alfa). Anda dapat mendalaminya dalam buku-buku tentang pengukuran sikap.

**BAB
VI**

PENGEMBANGAN INSTRUMEN RANAH PSIKOMOTOR

A. Pengertian Ranah Penilaian Psikomotor

Ranah psikomotor merupakan ranah yang berkaitan dengan keterampilan (*skill*) atau kemampuan bertindak setelah seseorang menerima pengalaman belajar tertentu. Ranah psikomotor adalah ranah yang berhubungan dengan aktivitas fisik, misalnya lari, melompat, melukis, menari, memukul, dan sebagainya. Hasil belajar ranah psikomotor dikemukakan oleh Simpson (1956) yang menyatakan bahwa hasil belajar psikomotor ini tampak dalam bentuk keterampilan (*skill*) dan kemampuan bertindak individu. Hasil belajar psikomotor ini sebenarnya merupakan kelanjutan dari hasil belajar kognitif (memahami sesuatu) dan dan hasil belajar afektif (yang baru tampak dalam bentuk kecenderungan-kecenderungan berperilaku). Hasil belajar kognitif dan hasil belajar afektif akan menjadi hasil belajar psikomotor apabila peserta didik telah menunjukkan perilaku atau perbuatan tertentu sesuai dengan makna yang terkandung dalam ranah kognitif dan ranah afektif dengan materi kedisiplinan menurut agama Islam sebagaimana telah dikemukakan pada pembicaraan terdahulu, maka wujud nyata dari hasil psikomotor yang merupakan kelanjutan dari hasil belajar kognitif dan afektif itu adalah; (1) peserta didik bertanya kepada guru pendidikan agama Islam tentang contoh-contoh kedisiplinan yang telah ditunjukkan oleh Rasulullah SAW, para sahabat, para ulama dan lain-lain; (2) peserta didik mencari dan membaca buku-buku, majalah-majalah atau brosur-brosur, surat kabar dan lain-lain yang membahas tentang kedisiplinan; (3) peserta didik dapat memberikan penjelasan kepada teman-teman sekelasnya di sekolah, atau kepada adik-adiknya di rumah atau kepada anggota masyarakat lainnya, tentang kedisiplinan diterapkan, baik di sekolah, di rumah maupun di tengah-tengah kehidupan masyarakat; (4) peserta didik menganjurkan kepada teman-teman sekolah atau adik-adiknya, agar berlaku disiplin baik di sekolah, di rumah maupun di tengah-tengah kehidupan masyarakat; (5) peserta didik dapat memberikan contoh-contoh kedisiplinan di sekolah, seperti datang ke sekolah sebelum pelajaran di mulai, tertib dalam mengenakan seragam sekolah, tertib dan tenang dalam mengikuti pelajaran, disiplin dalam mengikuti tata tertib yang telah ditentukan oleh sekolah, dan lain-lain; (6) peserta didik dapat memberikan contoh kedisiplinan di rumah, seperti disiplin dalam belajar, disiplin dalam menjalankan ibadah shalat, ibadah puasa, disiplin dalam menjaga kebersihan rumah, pekarangan, saluran air, dan lain-lain; (7) peserta didik dapat memberikan contoh kedisiplinan di tengah-tengah kehidupan masyarakat, seperti menaati rambu-rambu lalu lintas, tidak kebut-kebutan, dengan suka rela mau antri waktu membeli karcis, dan lain-lain, dan (8) peserta didik mengamalkan dengan konsekuen kedisiplinan dalam belajar, kedisiplinan dalam beribadah, kedisiplinan dalam menaati peraturan lalu lintas, dan sebagainya.

B. Karakteristik Ranah Penilaian Psikomotor

Ranah psikomotor berhubungan dengan hasil belajar yang pencapaiannya melalui keterampilan manipulasi yang melibatkan otot dan kekuatan fisik. Ranah psikomotor adalah ranah yang berhubungan aktivitas fisik, misalnya; menulis, memukul, melompat dan lain sebagainya.

Tabel 12. Kaitan Antara Kegiatan Pembelajaran dan Domain Tingkatan Aspek Psikomotorik

Tingkat	Deskripsi
Gerakan Refleks	Arti: gerakan refleks adalah basis semua perilaku bergerak, respons terhadap stimulus tanpa sadar. Misalnya: melompat, menunduk, berjalan, menggerakkan leher dan kepala, menggenggam, memegang Contoh kegiatan belajar: • mengupas mangga dengan pisau • memotong dahan bunga • menampilkan ekspresi yang berbeda • meniru gerakan polisi lalu lintas, juru parkir • meniru gerakan daun berbagai tumbuhan yang diterpa angin
Gerakan dasar (<i>basic fundamental movements</i>)	Arti: gerakan ini muncul tanpa latihan tapi dapat Diperhalus melalui praktik gerakan ini terpola dan dapat ditebak Contoh kegiatan belajar: • contoh gerakan tak berpindah: bergoyang, membungkuk, merentang, mendorong, menarik, memeluk, berputar • contoh gerakan berpindah: merangkak, maju perlahan-lahan, muluncur, berjalan, berlari, melompat-loncat, berputar mengitari, memanjat. • Contoh gerakan manipulasi: menyusun balok/blok, menggantung, menggambar dengan krayon, memegang dan melepas objek, blok atau mainan. • Keterampilan gerak tangan dan jari-jari: memainkan bola, menggambar.
Gerakan Persepsi (<i>Perceptual abilities</i>)	Arti : Gerakan sudah lebih meningkat karena dibantu kemampuan perseptual Contoh kegiatan belajar: • menangkap bola, mendribble bola • melompat dari satu petak ke petak lain dengan 1 kali sambil menjaga keseimbangan • memilih satu objek kecil dari sekelompok objek yang ukurannya bervariasi • membaca melihat terbangnya bola pingpong • melihat gerakan pendulun menggambar simbol geometri • menulis alfabet • mengulangi pola gerak tarian • memukul bola tenis, pingpong • membedakan bunyi beragam alat musik • membedakan suara berbagai binatang • mengulangi ritme lagu yang pernah didengar

	• membedakan berbagai tekstur dengan meraba
Gerakan Kemampuan fisik (<i>Psychical abilities</i>)	Arti: gerak lebih efisien, berkembang melalui kematangan dan belajar Contoh kegiatan belajar: • menggerakkan otot/sekelompok otot selama waktu tertentu • berlari jauh • mengangkat beban • menarik-mendorong • melakukan push-up • kegiatan memperkuat lengan, kaki dan perut • menari • melakukan senam • melakukan gerakan pesenam, pemain biola, pemain bola
Gerakan terampil (<i>Skilled movements</i>)	Arti: dapat mengontrol berbagai tingkat gerak – terampil, tangkas, cekatan melakukan gerakan yang sulit dan rumit (kompleks) Contoh kegiatan belajar: • melakukan gerakan terampil berbagai cabang olahraga • menari, berdansa • membuat kerajinan tangan • menggergaji • mengetik • bermain piano • memanah • skating • melakukan gerak akrobatik • melakukan koprok yang sulit
Gerakan indah dan kreatif (<i>Non-discursive communicatio</i>)	Arti: mengkomunikasikan perasaan melalui gerakan • gerak estetik: gerakan-gerakan terampil yang efisien dan indah • gerakan kreatif: gerakan-gerakan pada tingkat tertinggi untuk mengkomunikasikan peran Contoh kegiatan belajar: ○ kerja seni yang bermutu (membuat patung, melukis, menari balet) ○ melakukan senam tingkat tinggi ○ bermain drama (acting) ○ keterampilan olahraga tingkat tinggi

Tabel. 13. Kompetensi dan Kata Operasional dalam Ranah Psikomotor

Domain	Kategori Jenis Perilaku	Kemampuan Internal	Kata Kerja Operasional
Psikomotor	Persepsi	Menafsirkan rangsangan Peka terhadap	Memilih Membedakan Mempersiapkan

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

		rangsangan Mendiskriminasikan	Menyisihkan Menunjukkan Mengidentifikasi Menghubungkan
	Kesiapan	Berkonsentrasi : Menyiapkan diri (fisik dan mental)	Memulai Mengawali Bereaksi Mempersiapkan Menanggapi Mempertunjukkan
	Gerakan Terbimbing	Meniru contoh	Mempraktekkan Memainkan Mengikuti Mengerjakan Membuat Mencoba Memperlihatkan Memasang Membongkar
	Gerakan Mekanism	Berketrampilan Berpegang pada pola	Mengoperasionalkan Membangun Memasang Membongkar Memperbaiki Melaksanakan Mengerjakan Menyusun Menggunakan Mengatur Mendemonstarikan Memainkan Menangani
	Respon yang kompleks	Berketerampilan secara: Misalnya : Lancar Luwes Supel Gesit Lincih	Mengoperasikan Membangun Memasang Membongkar Memperbaiki Melaksanakan Mengerjakan Menyusun Menggunakan Mengatur Mendemonstrasikan Memainkan Menangani
	Penyesesuaian dan keaslian	Menyesuaikan diri bervariasi	Mengubah Mengadaptasikan Mengatur kembali Membuat variasi

Tabel 14. Kata Operasional dalam Ranah Psikomotor

Meniru (P1)	Manipulasi (P2)	Presisi (P3)	Artikulasi (P4)	Naturalisasi (P5)
Menyalin Mengi kuti Mereplikasi Mengulangi Mematuhi	Kembali membuat Membangun Melakukan, Melaksanakan, Menerapkan	Menunjukkan Melengkapi Menunjukkan, Menyempurnakan Mengkalibrasi Mengendalikan	Membangun Mengatasi Menggabungkan Koordinat, Mengintegrasikan Beradaptasi Mengembangkan Merumuskan, Memodifikasi Master	Mendesain Menentukan Mengelola Menciptakan

C. Pengukuran Ranah Penilaian Psikomotor

Ada beberapa ahli yang menjelaskan cara menilai hasil belajar psikomotor. Ryan (1980) menjelaskan bahwa hasil belajar keterampilan dapat diukur melalui (1) pengamatan langsung dan penilaian tingkah laku peserta didik selama proses pembelajaran praktik berlangsung, (2) sesudah mengikuti pembelajaran, yaitu dengan jalan memberikan tes kepada peserta didik untuk mengukur pengetahuan, keterampilan, dan sikap, (3) beberapa waktu sesudah pembelajaran selesai dan kelak dalam lingkungan kerjanya. Sementara itu Leighbody (1968) berpendapat bahwa penilaian hasil belajar psikomotor mencakup: (1) kemampuan menggunakan alat dan sikap kerja, (2) kemampuan menganalisis suatu pekerjaan dan menyusun urutan pengerjaan, (3) kecepatan mengerjakan tugas, (4) kemampuan membaca gambar dan atau simbol, (5) keserasian bentuk dengan yang diharapkan dan atau ukuran yang telah ditentukan.

Dari penjelasan di atas dapat dirangkum bahwa dalam penilaian hasil belajar psikomotor atau keterampilan harus mencakup persiapan, proses, dan produk. Penilaian dapat dilakukan pada saat proses berlangsung yaitu pada waktu peserta didik melakukan praktik, atau sesudah proses berlangsung dengan cara mengetes peserta didik.

Penilaian psikomotorik dapat dilakukan dengan menggunakan observasi atau pengamatan. Observasi sebagai alat penilaian banyak digunakan untuk mengukur tingkah laku individu ataupun proses terjadinya suatu kegiatan yang dapat diamati, baik dalam situasi yang sebenarnya maupun dalam situasi buatan. Dengan kata lain, observasi dapat mengukur atau menilai hasil dan proses belajar atau psikomotorik. Misalnya tingkah laku peserta didik ketika praktik, kegiatan diskusi peserta didik, partisipasi peserta didik dalam simulasi, dan penggunaan alins ketika belajar.

Observasi dilakukan pada saat proses kegiatan itu berlangsung. Pengamat terlebih dahulu harus menetapkan kisi-kisi tingkah laku apa yang hendak diobservasinya, lalu dibuat pedoman agar memudahkan dalam pengisian observasi. Pengisian hasil observasi dalam pedoman yang dibuat sebenarnya bisa diisi secara bebas dalam bentuk uraian mengenai tingkah laku yang tampak untuk diobservasi, bisa pula dalam bentuk memberi tanda cek (✓) pada kolom jawaban hasil observasi.

Tes untuk mengukur ranah psikomotorik adalah tes untuk mengukur penampilan atau kinerja (*performance*) yang telah dikuasai oleh peserta didik. Tes tersebut dapat berupa tes *paper and pencil*, tes identifikasi, tes simulasi, dan tes unjuk kerja.

1) Tes simulasi

Kegiatan psikomotorik yang dilakukan melalui tes ini, jika tidak ada alat yang sesungguhnya yang dapat dipakai untuk memperagakan penampilan peserta didik, sehingga peserta didik dapat dinilai tentang penguasaan keterampilan dengan bantuan peralatan tiruan atau berperaga seolah-olah menggunakan suatu alat yang sebenarnya.

2) Tes unjuk kerja (*work sample*)

Kegiatan psikomotorik yang dilakukan melalui tes ini, dilakukan dengan sesungguhnya dan tujuannya untuk mengetahui apakah peserta didik sudah menguasai/terampil menggunakan alat tersebut. Misalnya dalam melakukan praktik pengaturan lalu lintas lalu lintas di lapangan yang sebenarnya

Tes simulasi dan tes unjuk kerja, semuanya dapat diperoleh dengan observasi langsung ketika peserta didik melakukan kegiatan pembelajaran. Lembar observasi dapat menggunakan daftar cek (*check-list*) ataupun skala penilaian (*rating scale*). Psikomotorik yang diukur dapat menggunakan alat ukur berupa skala penilaian terentang dari sangat baik, baik, kurang, kurang, dan tidak baik.

Dengan kata lain, kegiatan belajar yang banyak berhubungan dengan ranah psikomotor adalah praktik di aula/lapangan dan praktikum di laboratorium. Dalam kegiatan-kegiatan praktik itu juga ada ranah kognitif dan afektifnya, namun hanya sedikit bila dibandingkan dengan ranah psikomotor. Pengukuran hasil belajar ranah psikomotor menggunakan tes unjuk kerja atau lembar tugas.

Contohnya kemampuan psikomotor yang dibina dalam belajar matematika misalnya berkaitan dengan kemampuan mengukur (dengan satuan tertentu, baik satuan baku maupun tidak baku), menggambar bentuk-bentuk geometri (bangun datar, bangun ruang, garis, sudut,dll) atau tanpa alat. Contoh lainnya, siswa dibina kompetensinya menyangkut kemampuan melukis jaring-jaring kubus. Kemampuan dalam melukis jaring-jaring kubus secara psikomotor dapat dilihat dari gerak tangan siswa dalam menggunakan peralatan (jangka dan penggaris) saat melukis. Secara teknis penilaian ranah psikomotor dapat dilakukan dengan pengamatan (perlu lembar pengamatan) dan tes perbuatan.

Dalam ranah psikomotorik yang diukur meliputi (1) gerak refleks, (2) gerak dasar fundamen, (3) keterampilan perseptual; diskriminasi kinestetik, diskriminasi visual, diskriminasi auditoris, diskriminasi taktis, keterampilan perseptual yang terkoordinasi, (4) keterampilan fisik, (5) gerakan terampil, (6) komunikasi non diskusi (tanpa bahasa-melalui gerakan) meliputi: gerakan ekspresif, gerakan interprestatif.

Lembar observasi

Beri Tanda (√)

Nama Siswa	Mengerjakan Tugas (<i>On-Task</i>)	Tidak Mengerjakan Tugas (<i>Off-Task</i>)	Catatan Guru
Damar			
Ayu			
Dst.....			

Tabel Instrumen (alat) Asesmen Kinerja (unjuk kerja) Berpidato dengan *numerical Rating Scale*

Nama : Kelas :					
Petunjuk: Berilah skor untuk setiap aspek kinerja yang sesuai dengan ketentuan berikut: (4) bila aspek tersebut dilakukan dengan benar dan cepat (3) bila aspek tersebut dilakukan dengan benar tapi lama (2) bila aspek tersebut dilakukan selesai tetapi salah (1) bila dilakukan tapi tidak selesai (0 = tidak ada usaha)					
No	Aspek yang dinilai	Skor			
		4	3	2	1
1.	Berdiri tegak menghadap penonton				
2.	Mengubah ekspresi wajah sesuai dengan pernyataan				
3.	Berbicara dengan kata-kata yang jelas				
4.	Tidak mengulang-ulang pernyataan				
5.	Berbicara cukup keras untuk didengar penonton				

D. Penerapan Pengukuran dalam Domain

1. Kawasan Psikomotor

Untuk mengevaluasi hasil belajar domain psikomotor harus dilakukan melalui pengamatan (observasi) terhadap si pebalajar (siswa). Dalam melakukan observasi ini pengamat (observer) selain melakukan evaluasi terhadap perilaku yang ditampilkan siswa, tetapi sekaligus pengamat dapat belajar dari hasil pengamatan terhadap perilaku yang ditampilkan oleh siswa yang dalam hal ini sebagai model. Belajar melalui pengamatan (observastion learning) didasarkan pada teori dari Bandura (1986:262) yang mengatakan bahwa pengamat dapat mempelajari ketrampilan kognitif, afektif, atau psikomotor yang baru, dengan cara memperhatikan (attention) bagaimana orang melakukan suatu ketrampilan atau hal-hal baru. Dalam melakukan pengamatan terhadap perilaku yang ditampilkan si pebelajar (siswa), dapat digunakan lembar pengamatan (observasi) untuk tujuan belajar maupun sebagai evaluasi. Di bawah ini disajikan contoh lembar observasi untuk ketrampilan anak berdasarkan tahap-tahap perkembangan Piaget, dan lembar observasi untuk mengevaluasi ketrampilan pada tahap operasi formal (dalam hal ini evaluasi praktek bengkel untuk siswa SMK).

Contoh Lembar Observasi perkembangan Kognitif pada anak berdasarkan perkembangan Piaget. Menurut teori Piaget ada empat tahap perkembangan kognitif seperti berikut :

Lembar Observasi Tahap Perkembangan Anak Menurut Piaget

Rentang Usia	Tahap	Ciri Pokok
Lahir- (1,5-2) th	Sensori Motor	- Prasimbolik dan Praverbal - Perkembangan pada gerak

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

(2-3) – (7-8) th	Pra Operasional	- Mulai berfikir logis parsial - Proses berfikir atas dasar isyarat persepsual - Perkembangan bahasa
(7-8) – (12-14) th	Operasi konkrit	- Muncul refleksi dasar - Melihat pandangan lain - Kesepakatan aturan - Kerjasama dengan aturan - Berfikir logis terkait obyek - Berfikir bebas dari isyarat perseptual
Lebih dari 14 th	Operasi formal	- Berfikir masa depan - Mulai peran dewasa - Mulai berfikir logis tentang banyak factor - Berfikir hipkretis kekonkrit

**BAB
VII**

**PENGUJIAN INSTRUMEN
PENELITIAN DAN PENILAIAN**

A. Pengertian Validitas dan Reliabilitas

Dalam sains, pengukuran oleh suatu instrumen merupakan sesuatu yang esensial dan terstandardisasi. Contohnya, untuk mengukur suhu tubuh digunakan thermometer yang skalanya sudah ditetapkan dan dikalibrasi berulang kali dan hasil bacaannya memberikan informasi yang bisa menyimpulkan sesuatu. Contohnya, suhu tubuh normal adalah 37°C. Apabila hasil bacaan thermometer pada suhu tubuh seseorang adalah 40°C dan dokter percaya dengan hasil bacaan tersebut, dokter akan melakukan tindakan medis karena hal itu menunjukkan gejala yang tidak normal. Siswa telah mengikuti tes pendidikan agama Islam dengan alat tes yang terstandar dengan hasil 50. Padahal siswa dianggap memiliki nilai ketuntasan belajar jika nilainya sebesar 75 dianggap telah memenuhi kriteria ketuntasan belajar (*mastery learning*) atau yang sekarang dikenal KKM sekolah. Dimana penentuan nilai KKM sangatlah beragam ada yang skor 70, 75, 80 dan 85. Apabila ada siswa pada mata pelajaran pendidikan agama Islam bernilai 50, maka bagi guru pendidikan agama Islam siswa tersebut dianggap bermasalah dalam proses kegiatan belajar mengajar di sekolah.

Alat berupa Thermometer yang sama juga akan memberikan hasil bacaan yang dipercaya oleh dokter untuk mengukur suhu pasien lain (baik pasien itu besa-kecil, tua-muda, lelaki-perempuan, dll.) yang berada ditempat manapun di muka bumi. Hal ini menunjukkan thermometer sebagai instrumen yang valid (tepat) untuk mengukur suhu tubuh, dan reliabel memberikan hasil pengukuran yang konsisten). Dalam penelitian kuantitatif, isu realibilitas dan validitas instrumen memegang peranan yang sangat penting karena dapat menunjukkan kualitas riset secara mendasar.

Secara metodologis, penelitian dan penilaian kuantitatif memiliki dua bagian yaitu (a) pengembangan dan pengujian instrumen pengukuran dan (b) pengujian hipotesis. Bagian pertama mengkaji kualitas instrumen yang memuat penjelasan mengenai reliabilitas, validitas, tingkat kesukaran soal, daya beda, dan keberfungsian pengecoh. Dalam bidang ilmu sosial, bagian ini dinamakan dengan kajian psikometri. Bagian kedua berisi deskripsi statistik data dan pengujian hipotesis penelitian. Bagian ini dinamakan dengan kajian statistik yang membuat berbagai prosedur untuk mengidentifikasi keterkaitan antar variabel, baik melalui uji perbedaan (misalnya uji-t) maupun uji korelasi (seperti korelasi Pearson). Paparan ini menunjukkan bahwa reliabilitas dan validitas adalah salah satu pilar utama dalam riset kuantitatif. Instrumen yang reliabel dan valid akan memberikan informasi yang dapat dipercaya. Sebaliknya, instrumen yang gagal memenuhi kedua syarat tersebut akan memberikan hasil yang bisa atau menyesatkan sehingga dapat menurunkan kualitas penelitian.

Untuk lebih jelasnya tentang uji reliabilitas, validitas, tingkat kesukaran soal, daya beda dan keberfungsian pengecoh, maka dapat dilihat pada sub bab penjelasan dibawah ini.

1. Validitas

Validitas tes berhubungan dengan ketepatan terhadap apa yang mesti diukur oleh tes dan beberapa seberapa cermat tes melakukan pengukurannya. Atau dengan kata lain validitas tes berhubungan dengan ketepatan tes tersebut terhadap konsep yang diukur, sehingga betul-betul bisa mengukur apa yang seharusnya diukur.

Konsep utama kedua dalam menentukan kualitas instrumen adalah validitas atau ketepatan. Singkatnya, validitas instrumen adalah seberapa jauh pengukuran oleh instrumen dapat atribut apa yang seharusnya diukur. Hal ini bermakna bahwa instrumen yang digunakan mengukur sesuai dengan tujuan yang telah ditetapkan. Berhubung terdapat sifat alami yang tersembunyi dari atribut atau konstruk yang akan diukur maka akan selalu terdapat kesimpulan yang terjadi antara indikator yang diamati (yaitu respons partisipan terhadap butir-butir instrumen) dan konsepsi yang kita ukur. Dengan demikian, validitas menjelaskan seberapa tepat atau sah simulasi yang dilakukan oleh riset tersebut. Oleh karena itu, validitas memuat pertanyaan mengenai seberapa jauh kesimpulan yang dibuat mengenai hubungan antara indikator dan atribut ukur sesuai dengan konsep atau teori yang dipakai. Sama seperti reliabilitas, instrumen pengukuran atau prosedurnya sendiri tidak bisa disebut valid atau tidak valid sehingga pertanyaan validitas selalu ditunjukkan kepada kesimpulan yang dibuat dari apa yang kita amati.

Terdapat tiga jenis validitas yang harus jadi perhatian yaitu validitas isi (*content validity*), validitas konstruksi (*construct validity*), dan validitas kriteria terkait (*criterion related validity*).

a) Validitas Isi (*content validity*)

Validitas isi berhubungan dengan tes untuk mengukur isi yang seharusnya diukur. Dengan kata lain validitas isi menyatakan apakah tes sudah mencakup sampel yang representatif dari domain pelaku yang diukur.

Validitas isi berfokus pada apakah keseluruhan definisi konseptual telah sepenuhnya direpresentasikan dalam pengukuran. Definisi konseptual yang dimaksud adalah pada yang mengikat semua ide dan konsep, dan indikator yang digunakan untuk mengukur haruslah berisi semua ide yang ada dalam definisi tersebut. Terdapat dua tahapan yang harus dilakukan yaitu (a) menemukan isi definisi yang digunakan, dan (b) mengembangkan indikator yang mencakup semua hal yang terdapat dalam definisi tersebut. Contohnya, Anda hendak mengukur kecerdasan maka agar menyimpulkan yang nanti dilakukan dapat dikatakan valid, peneliti perlu menggunakan definisi yang bersifat komprehensif tentang kecerdasan. Definisi tersebut menjadi dasar pengembangan instrumen yang mengukur semua aspek kecerdasan yang meliputi logika, berpikir abstrak, pemahaman, kesadaran diri, komunikasi, belajar perencanaan, pemecahan masalah, dll. Dengan landasan teoritis yang baik maka instrumen yang dibuat dapat dikatakan mempunyai validitas isi yang bagus.

Misalnya tes hasil belajar Pendidikan Agama Islam (PAI) di SMU harus mampu mengungkapkan isi pelajaran PAI di SMU.

b) Validitas Konstruk (*construct validity*)

Validitas konstruk menunjukkan pada sejauh mana suatu instrumen mampu mengukur pengertian-pengertian yang terkandung dalam materi yang akan diukur. Misalnya instrumen minat harus mampu mengukur pengertian-pengertian yang terkandung pada variabel minat. Agar lebih jelas, biasanya variabel tersebut diuraikan dulu menjadi indikator-indikator.

Pada validitas konstruksi ini penilai dan peneliti menjelaskan seberapa baik pengukuran telah sesuai dengan espektasi teoritis. Maksudnya adalah berbagai ukuran yang terdapat dalam berbagai konteks teoritis, kesemuanya harus menunjukkan hubungan dengan konsep lain yang bisa diprediksi dan diinterpretasi dalam konteks tersebut. Sebuah contoh, hasil pengukuran terhadap perilaku agresi melalui instrumen A dapat dikatakan valid ketika meneliti korelasi yang tinggi dengan instrumen B yang juga mengukur agresi (validitas konvergen). Disisi lain, korelasi antara instrumen A dan instrumen B tersebut nilainya lebih tinggi dibandingkan korelasi instrumen A dan instrumen C yang mengukur kecerdasan (validitas divergen), yang secara teoritis tidak berkaitan dengan agresi.

c) Validitas Kriteria Terkait (*criterion related validity*).

Untuk validitas jenis terakhir adalah suatu indikator yang dibandingkan pengukuran lain dengan konsep yang sama. Terdapat dua jenis validitas ini yaitu validitas bandingan (*concurrent validity*) dan validitas ramalan (*predictive validity*).

1. Validitas bandingan (*concurrent validity*) menekankan pada variabel kriteria yang ada dalam saat ini atau dapat juga dibandingkan kondisi sebelumnya. Validitas ini diberi makna juga sebagai validitas kesamaan menunjuk kepada sejauh mana tes memiliki kesamaan dengan tes yang sudah ada atau yang sudah dibakukan . Kesamaan yang dimaksud meliputi kemampuan yang diukur, obyek yang diukur, dan waktu yang diperlukan .
2. Validitas ramalan (*predictive validity*) menekankan variabel kriteria yang baru akan ada pada masa depan. Validitas prediksi menunjuk pada sejauh mana tes dapat menentukan atau meramalkan kriteria tertentu yang diinginkan. Misalnya: a). apakah NEM di SLTP dapat digunakan untuk meramalkan prestasi belajar di SMU; b). apabila kita melakukan studi tentang prestasi seorang mahasiswa. Karena konteks validitas bandingan adalah saat ini, pertanyaannya adalah “Berapakah IPK yang Anda dapatkan tahun lalu?” disisi lain untuk validitas ramalan karena konteksnya belum terjadi, pertanyaannya adalah, “Berapakah kira-kira IPK yang akan Anda peroleh tahun depan?”

2. Reliabilitas

Reliabilitas tes berhubungan dengan konsistensi hasil pengukuran , yaitu seberapa konsistensi skor tes dari satu pengukuran ke pengukuran berikutnya. Reliabilitas dinyatakan dengan koefisien reabilitas ,yaitu koefisien korelasi yang menunjukkan derajat hubungan antara dua hasil pengukuran yang diperoleh dari instrument atau prosedur yang sama. Reliabilitas merujuk pada ketetapan/ keajegan alat tersebut dalam menilai apa yang diinginkan , artinya kapanpun alat tersebut digunakan akan memberikan hasil yang relative sama. Tes hasil belajar dikatakan ajeg apabila hasil pengukuran saat ini menunjukkan kesamaan hasil pada saat yang berlainan, jika dikenakan pada siswa yang sama. Meskipun masih memungkinkan terjadi perbedaan hasil untuk hal-hal tertentu akibat dari factor kebetulan ,selang waktu, perubahan pandangan siswa terhadap soal.

Reliabilitas menjelaskan seberapa jauh pengukuran yang dilakukan berkali-kali akan menghasilkan informasi yang sama. Artinya, tidak menghasilkan banyak perbedaan informasi yang berarti. Oleh karena perbedaan informasi itu akan selalu ada, pengukuran yang meyakinkan (*reliable*) tidak selalu menghasilkan informasi yang benar-benar sama persis. Informasi yang dihasilkan dapat memiliki perbedaan, akan tetapi nilainya kecil dan masih dalam batas toleransi.

Pertanyaan selanjutnya adalah apa dasar untuk menyimpulkan bahwa informasi yang dihasilkan adalah sama?. Ada empat metode dasar yang dapat dipakai, yaitu kesamaan berdasarkan antar-waktu (*stabilitas*), instrument paralelnya (*ekivalensi*), elemen-elemen didalam instrumen (*internal consistency*), dan kesepakatan penilai (*agreement*).

a) Reliabilitas stabilitas (*test-retest method*)

Dalam menentukan reabilitas tes melalui metode test-retest, tes yang sama dikenakan dua kali terhadap kelompok siswa yang sama dengan interval tertentu. Kedua hasil tes dikorelasikan, dan koefisien korelasi yang diperoleh menyatakan reabilitas tes, yaitu seberapa kestabilan tes tersebut dalam selang waktu tertentu. Semakin tinggi koefisien reliabilitas tes maka tes semakin stabil. Bila tes stabil maka skor siswa dalam kedua tes cenderung tetap. Satu hal yang penting dalam metode test-retest adalah menentukan interval waktu pelaksanaan tes, jika interval terlalu pendek maka siswa masih ingat hasil terdahulu, sebaliknya semakin besar interval waktu maka semakin banyak variable yang mempengaruhi hasil tes. Koefisien reliabilitas 0,80 sampai 0,90 dianggap standar untuk tes bakat dan tes *achievement* dalam tahun yang sama.

Reliabilitas stabilitas menunjukkan bahwa sebuah instrument yang dikenakan kepada orang yang sama dan dalam kondisi yang sama, namun dalam waktu yang berbeda, akan memberikan hasil yang kurang lebih setara. Kesetaraan ini menunjukkan bahwa instrument yang digunakan adalah reliable. Namun, apabila hasilnya jauh berbeda, instrument tersebut kurang reliable dan tidak layak digunakan dalam mengukur. Stabilitas hasil pengukuran dari waktu ke waktu ini bisa di estimasi dengan melakukan pengukuran instrument yang sama dalam kondisi tertentu dalam waktu yang berbeda. Hasil pengukuran dari kedua waktu tersebut di korelasikan untuk melihat kestabilan hasilnya (*test-retest reliability*).

Tabel 15. Perbandingan Prosedur Mengestimasi Reliabilitas

Jenis Reliabilitas	Tes Ulang	Tes Paralel	Konsistensi Internal	Antar-Penilai
Fokus	Butir	Butir	Butir	Penilai
Jumlah pengukuran	2 kali	1 kali	1 kali	1 kali
Jumlah Instrumen	1 buah	2 buah	1 buah	1 buah
Penunjuk Reliabilitas	Korelasi	Korelasi	Koefisien Reliabilitas	Koefisien Reliabilitas

(Sumber : diadaptasi dari Urbina, 2004)

Rumus yang digunakan untuk mencari harga koefisien korelasi kedua tes tersebut dapat digunakan korelasi pearson sebagai berikut :

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

Keterangan :

N= Banyaknya peserta tes

$\sum X$ =Jumlah skor tes pertama

$\sum Y$ =Jumlah skor tes kedua

r_{xy} = Harga Koefisien korelasi

Kriteria Pengujian reliabilitas :

- Harga r_{xy} = 1 korelasi sempurna
- Harga r_{xy} = 0.99-0.81 korelasi sangat tinggi
- Harga r_{xy} = 0.80-0.61 korelasi tinggi
- Harga r_{xy} = 0.60-0.41 korelasi sedang
- Harga r_{xy} = 0.40-0.21 korelasi rendah
- Harga r_{xy} = 0.20-0.01 korelasi sangat rendah

Dengan kriteria pengujian tersebut, maka kita dapat memberikan interpretasi reliabilitas dari suatu instrument penilaian dan penelitian.

b) Reliabilitas ekivalensi

Reliabilitas ekivalensi menunjukkan seberapa besar hasil yang dihasilkan oleh dua instrumen paralel menghasilkan informasi yang setara. Dua buah instrumen dikatakan paralel ketika memiliki kesamaan dalam hal : (a) target ukur, (b) spesifikasi (*blue print*) (c) jumlah butir, (d) tingkat kesukaran (untuk tes), serta (e) administrasi pengukuran. Dalam pendekatan ini besarnya reliabilitas ditunjukkan dengan nilai korelasi antarskor individu melalui kedua instrumen yang paralel. Menentukan reliabilitas tes dengan metode tes ekivalen dilakukan dengan menggunakan dua tes yang berbeda tetapi ekivalen (*parallel*). kedua tes dikenakan kepada kelompok siswa yang sama dalam waktu yang berurutan, kemudian hasilnya dikorelasikan dengan formula korelasi *product moment*. Koefisien korelasi yang diperoleh menyatakan koefisien reliabilitas, yaitu ukuran ekivalensi tes . Koefisien reliabilitas yang tinggi menunjukkan bahwa kedua tes menghasilkan hasil yang cenderung sama.

c) Konsistensi internal

Konsistensi internal menekankan pada konsistensi pada butir-butir yang terdapat didalam instrumen. Instrumen yang bertujuan mengungkap ciri-sifat laten (*latent traits*) biasanya aitem-aitem yang berupa pernyataan yang harus di respons oleh partisipan. Konsistensi internal menjelaskan bagaimana berbagai butir tersebut menghasilkan informasi yang konsisten satu sama lainnya dan berada dalam satu arah yang sama. Konsistensi internal pengukuran dapat diestimasi dengan menggunakan berbagai koefisien sesuai dengan karakteristik item-item di dalam instrumen. Jenis koefisien tersebut dapat dilihat pada Tabel 16.

Tabel 16. Pebandingan Koefisien Reliabilitas

Jenis Data dan Koefisien		
	Data Pilah (misal: 0.1)	Data Berurutan (misal: 0,1,2,...)
I. Teori Klasik		
Pengukuran Unidimensi	KR-20	Spearman Brown
	KR-21	Alpha Cronbach
		Koefisien Feldt
		Theta Armor
		Omega McDonald, dsb.
Pengukuran Multidimensi		Koefisien Mosier
		Koefisien Alpha Berstrata
		Reliabilitas Komposit
II. Teori Modern (Rasch)		
Pengukuran Unidimensi	Fungsi Informasi & Indeks Separasi	Fungsi Informasi
Pengukuran Multidimensi		Indeks Separasi

(Sumber: Widhiarso, 2008)

Adapun metode pengujian reliabilitas konsistensi internal yang biasa digunakan yaitu:

1) Metode Belah-Dua (*Split-Half Method*)

Realibilitas juga bias ditentukan dengan pemberian satu tes dalam sekali waktu. Tes dikenakan kepada siswa seperti biasanya, kemudian tes dibagi dua dalam pemberian skor. Pembagian biasanya dilakukan dengan patokan nomor ganjil dan genap. Kedua skor bagian tes kemudian dikorelasikan dengan teknik korelasikan dengan teknik korelasi product moment. Koefisien korelasi yang diperoleh menyatakan ukuran konsistensi internal ,yaitu seberapa konsisten hasil yang diperoleh dari kedua bagian tes. Apabila diperlukan untuk memperpanjang atau memperpendek tes maka dapat digunakan Formula Spearmen-Brown :

$$r_{11} = \frac{2r^{1/2^{1/2}}}{(1+r^{1/2^{1/2}})}$$

Dimana :

$r^{1/2^{1/2}}$ = korelasi antara skor-skor belahan tes

r_{11} = koefisien korelasi yang sudah disesuaikan.

2) Metode Kuder Richadson, Alpha Cronbach's, dan Anova Hoyt.

Metode Kuder-Richadson hampir sama dengan metode *split-half*, hanya saja tes tidak dibagi dua, melainkan berdasarkan pada proporsi siswa yang menjawab tes dengan benar dan standar deviasi skor total. Metode kuder-Richardson mendasarkan pada kovarian butir. Ada tiga prosedur yang dapat digunakan untuk memperoleh dan sampel individu pada satu kali pengetesan. Ketiga prosedur dengan menggunakan formula: Kuder Ricardson 20 (KR-20); Cronbac's Alpa; dan

analisis varians Hoyt. KR-20 hanya bias digunakan untuk butir-butir yang di skor secara ekonomi. Sebenarnya rumus Kuder Ricardson ini ada dua, yaitu KR-20 dan KR-21. KR-20 digunakan untuk tes yang seluruh butir soal memiliki saraf kesukaran sama, sedangkan KR-21 digunakan untuk tes yang taraf kesukaran butir soalnya tidak sama. Cronbac's Alpha atau koefisien alpha digunakan untuk menaksir kosistensi interanal butir yang diskor secara dikotomi maupun butir-butir yang memiliki rentang bobot penskoran yang lebar seperti halnya tes sikap dan tes uraian. Analisis varians Hoyt memberikan hasil yang identik dengan koefisien alpha.

Faktor-faktor yang Mempengaruhi Reliabilitas

- a. Panjang tes, secara umum jika tes semakin panjang maka semakin tinggi reliabilitasnya.
- b. Penyebaran skor, semakin besar penyebaran skor, maka akan semakin tinggi perkiraan reliabilitasnya.
- c. Kesulitan tes, umumnya tes terlalu mudah atau terlalu sulit akan menyebabkan reliabilitas tes semakin rendah, hal ini karena disebabkan karena terbatasnya penyebaran skor.
- d. Objektif tes, yakni tes yang objektifnya tinggi memiliki reliabilitas yang lebih tinggi, karena hasil tesnya tidak dipengaruhi oleh prosedur penskorannya.
- e. Interval waktu tes, tes dengan interval waktu yang pendek menyebabkan koefisien reliabilitas tes yang besar.

Berikut ini beberapa catatan mengenai reliabilitas.

- *Reliabilitas dan Karakteristik Sampel.* Besarnya nilai reliabilitas bergantung pada karakteristik sampel. Penggunaan instrumen yang sama dalam pengukuran harga diri pada siswa di Sekolah A dapat berbeda hasilnya dengan Sekolah B. oleh karena itu, ketika penilai menggunakan pendekatan analisis yang masih berbasis pada sampel, sebaiknya menggunakan istilah "Reliabilitas skor atau pengukuran" dan bukan "Reliabilitas instrumen". Instrumen tidak dapat dikatakan *reliable* atau tidak *reliable*, tetapi pengukuran atau skorlah yang dapat dikatakan demikian. Beberapa pendekatan telah diperkenalkan untuk mengestimasi reliabilitas yang telah terbebas dari karakteristik sampel, salah satunya adalah reliabilitas yang berbasis teori modern (IRT/Rasch).
- *Reliabilitas Dan Pengukuran Unidimensi.* Reliabilitas adalah property untuk mengukur satu atribut ciri-sifat oleh karena itu, ketika instrumen yang dipakai oleh peneliti mengukur atribut yang majemuk (multidimensional), reliabilitas diestimasi pada tiap atributnya. Sebagai contoh, BFI-44 adalah instrumen multidimensi yang mengukur lima ciri sifat kepribadian (ekstraversi, keramahan, keterbukaan, kestabilan dan ketabahan). Pada kasus ini, reliabilitas perlu diestimasi pada tiap dimensi secara terpisah. Jika peneliti menggunakan koefisien Alpha, akan didapatkan lima koefisien reliabilitas untuk tiap-tiap dimensi. Beberapa peneliti telah menggunakan pengembangan koefisien reliabilitas untuk pengukuran multidimensi. Dengan menggunakan koefisien ini meski pengukuran dilakukan pada ciri-sifat yang majemuk, banyaknya koefisien reliabilitas yang dihasilkan adalah satu.

d) Kesepakatan penilai (*agreement*)

Pengukuran yang dilakukan beberapa kali dengan alat yang sama terkadang akan menghasilkan hasil yang beragam. Dalam kasus seperti ini diperlukan kesepakatan penilai untuk

menentukan reliabilitas dari alat ukur yang dipakai. Hal ini didasarkan pada asumsi bahwa reliabilitas alat ukur itu adalah satu reliabilitas.

B. Prosedur Pengujian Instrumen Penelitian

Pengujian instrumen di mulai dengan proses uji coba instrumen terhadap sejumlah responden. Hasil yang diperoleh kemudian di analisis untuk menguji instrumen tersebut. Ukuran umum digunakan adalah validitas dan realibilitas. Khusus untuk instrumen berbentuk tes hasil belajar, selain validitas dan reabilitas masih ada lagi ukuran lain yang digunakan, antara lain tingkat kesukaran, tingkat pembeda, dan analisis pengecoh. (Untuk lebih rinci lihat Anne Anastasi & Susana Urbina (1997 : 82-200), Norman E.Gronlund & Robert L.Linn (1990: 44 – 7), dan L.Linn & Norman E.Gronlund, (1995: 47-112). Pengujian ini sering dikenal analisis item (*Item Analysis*).

Item analysis adalah analisis yang dilakukan terhadap item tes untuk menentukan tingkat kesukaran, daya beda, dan efektifitas distrakter. Item analysis hanya diperlukan pada instrumen berupa *achievement test*, sedangkan **instrument non-tes tidak memerlukan item analysis**. Item analysis terdiri dari (a) analisis tingkat kesukaran, (b) analisis daya pembeda, (c) analisis efektifitas pengecoh (khusus untuk tes berbentuk pilihan ganda).

a. Analisis Tingkat kesukaran atau disebut dengan istilah (*item difficulties*)

Analisis tingkat kesukaran dimaksudkan untuk mengkaji soal yang mudah, sedang dan sukar, sehingga bias menyeimbangkan proporsi soal yang mudah, sedang dan sukar dalam tes. Tingkat kesukaran tes dipandang dari kemampuan siswa untuk menjawabnya. Tingkat kesukaran tes menunjukkan presentase siswa yang menjawab item dengan benar. Pada analisis tingkat kesukaran akan dicari indeks tingkat kesukaran biasanya dirumuskan dengan $P = B/T$:

- P = indeks tingkat kesukaran
B = banyaknya siswa yang menjawab soal dengan betul
T = total seluruh peserta tes

Soal dengan indeks kesukaran 0 – 0,30 tergolong sukar, soal dengan indeks kesukaran 0,31 – 0,70 tergolong sedang, dan soal dengan indeks kesukaran 0,71 – 1 tergolong mudah.

b. Analisis daya pembeda tes (*item discriminating power*)

Analisis daya pembeda suatu soal tes dimaksud untuk mengkaji kemampuan soal untuk membedakan antara siswa yang memiliki prestasi tinggi dan memiliki prestasi rendah. Dengan kata lain jika soal dikenakan pada kelompok siswa yang kurang hasilnya jelek. Soal dikatakan tidak memiliki daya pembeda jika soal tersebut diberikan kepada anak yang pintar maupun anak yang kurang hasilnya sama. Atau bias terjadi bisa jadi berbalik, yaitu jika diberikan kepada anak yang pintar hasilnya jelek, sedangkan jika dikasi kepada anak yang kurang hasilnya baik. Analisis daya pembeda bisa dilakukan dengan menggunakan kriteria Rose dan Stanley. Selain itu analisis itu analisis daya pembeda juga bisa dilakukan dengan menghitung indeks diskriminasi, dengan rumus:

$$D = BA/TA - BB/TB$$

TA=banyak peserta kelompok atas

TB=banyak peserta kelompok bawah

BA= banyak peserta kelompok atas yang menjawab soal dengan benar

BB= banyak peserta kelompok bawah yang menjawab soal dengan benar.

c. Analisis efektifitas pengecoh (*distractor*)

Khusus untuk tes pilihan ganda, selain daya pembeda, juga diperlukan analisis efektifitas pengecoh. Analisis efektifitas pengecoh digunakan untuk menentukan pengecoh (*distractor*) sudah berfungsi sebagai pengecoh dengan baik atau tidak. Sebuah pengecoh dapat dikatakan berfungsi dengan baik bila pengecoh tersebut mempunyai daya tarik yang besar bagi peserta tes yang kurang memahami materi. Sebelum menentukan efektifitas pengecoh maka terlebih dahulu harus ditentukan pola jawaban, yaitu dengan menghitung banyaknya peserta tes yang memilih masing-masing pilihan jawaban atau tidak memilih sama sekali.

Kegunaan item analysis antara lain adalah:

- Menyeimbangkan proporsi item mudah, sedang, dan sulit dalam tes
- Item analysis bisa digunakan untuk memperpanjang dan memperpendek tes dalam rangka meningkatkan validitas dan reliabilitas tes
- Kesederhanaan konsep dasar dan teknik item analysis dapat membantu pemakai tes untuk mengevaluasi tes yang dipublikasikan

1. Pengujian Instrumen Tes

Berbeda dengan pengujian instrumen kuisioner (instrumen non tes) yang menggunakan rumus korelasi Pearson Product Moment, untuk menghitung validitas soal tes, dapat menggunakan korelasi *Point Biserial* atau korelasi Pearson Product Moment. Dalam hal ini disarankan dengan menggunakan formula *Point Biserial*. Hal ini disebabkan tes tersebut hanya menghendaki jawaban benar diberi skor 1 (satu) dan jawaban yang salah diberi skor 0

Hasil uji coba *achievement test* untuk mengukur kemampuan menyelesaikan persoalan program linier, yang dikenakan kepada 20 orang siswa memberikan hasil yang mengujiannya dapat dilihat sebagai berikut:

Tabel 17. Hasil Uji Coba *Achievement Test* dari 20 Orang

Testee	Skor untuk butir pada item nomor									
	1	2	3	4	5	6	7	8	9	10
Kunci Jawaban	A	B	B	D	B	C	D	B	B	C
A	B	B	D	C	C	C	C	C	C	C
B	A	D	B	C	B	D	D	B	B	C
C	D	B	D	D	B	D	C	B	B	C
D	A	B	B	D	B	C	D	B	B	C
E	A	C	B	D	D	C	C	B	B	C
F	B	B	D	C	D	C	C	B	C	D
G	A	C	C	D	B	C	D	B	B	C

H	A	C	B	D	B	C	D	B	B	C
I	C	B	D	D	D	C	B	B	D	C
J	A	B	B	D	B	C	D	B	B	C
K	B	B	B	D	B	C	D	D	D	D
L	B	B	B	B	B	B	D	C	B	B
M	B	B	C	C	B	D	D	B	D	B
N	A	D	B	D	D	C	C	B	B	C
O	A	D	C	D	B	C	D	B	B	C
P	B	B	D	D	D	C	B	B	D	C
Q	A	D	B	D	B	C	D	B	B	C
R	B	B	D	D	B	C	B	B	A	C
S	A	D	A	D	B	C	D	B	B	C
T	D	B	B	A	C	D	D	D	C	C

Tabel 18. Skoring Hasil Uji Coba Achievement Test dari 20 Orang

Testee	Skor untuk butir pada item nomor									
	1	2	3	4	5	6	7	8	9	10
A	0	1	0	0	0	1	0	0	0	1
B	1	0	1	0	1	0	1	1	1	1
C	0	1	0	1	1	0	0	1	1	1
D	1	1	1	1	1	1	1	1	1	1
E	1	0	1	1	0	1	0	1	1	1
F	0	1	0	0	0	1	0	1	0	0
G	1	0	0	1	1	1	1	1	1	1
H	1	0	1	1	1	1	1	1	1	1
I	0	1	0	1	0	1	0	1	0	1
J	1	1	1	1	1	1	1	1	1	1
K	0	1	1	1	1	1	1	0	0	0
L	0	1	1	0	1	0	1	0	1	0
M	0	1	0	0	1	0	1	1	0	0
N	1	0	1	1	0	1	0	1	1	1
O	1	0	0	1	1	1	1	1	1	1
P	0	1	0	1	0	1	0	1	0	1
Q	1	0	1	1	1	1	1	1	1	1
R	0	1	0	1	1	1	0	1	0	1
S	1	0	0	1	1	1	1	1	1	1
T	0	1	1	0	0	0	1	0	0	1

1.1. Pengujian Validitas Instrumen Tes

Ada beberapa langkah yang harus dilakukan dalam melaksanakan uji validitas soal. Hasil tabulasi jawaban dan skoringnya dari *achievement test*, yang memiliki nilai benar (1) – salah (0), maka pengujian validitas item instrumen dilakukan dengan menggunakan rumus koefisien korelasi point biserial. Rumus yang digunakan adalah sebagai berikut:

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{p}{q}}$$

M_p = Rerata skor total dari subyek yang menjawab betul item yang dicari validitasnya.

M_t = Rerata skor total

S_t = standar Deviasi skor total

p = proporsi siswa yang menjawab benar item yang dicari validitasnya

q = proporsi siswa yang menjawab salah item yang dicari validitasnya

$q = (1 - p)$

Adapun langkah-langkah tersebut adalah :

1. Langkah pertama yang harus dilakukan adalah membuat tabel kerja atau tabel penolong.

Tabel 19. Skoring Hasil Uji Coba Achievement Test dari 20 Orang

Tes-tee	Skor untuk butir item nomor:										Skor Total (X _i)	X _i ²
	1	2	3	4	5	6	7	8	9	10		
A	0	1	0	0	0	1	0	0	0	1	3	9
B	1	0	1	0	1	0	1	1	1	1	7	49
C	0	1	0	1	1	0	0	1	1	1	6	36
D	1	1	1	1	1	1	1	1	1	1	10	100
E	1	0	1	1	0	1	0	1	1	1	7	49
F	0	1	0	0	0	1	0	1	0	0	3	9
G	1	0	0	1	1	1	1	1	1	1	8	64
H	1	0	1	1	1	1	1	1	1	1	9	81
I	0	1	0	1	0	1	0	1	0	1	5	25
J	1	1	1	1	1	1	1	1	1	1	10	100
K	0	1	1	1	1	1	1	0	0	0	6	36
L	0	1	1	0	1	0	1	0	1	0	5	25
M	0	1	0	0	1	0	1	1	0	0	4	16
N	1	0	1	1	0	1	0	1	1	1	7	49
O	1	0	0	1	1	1	1	1	1	1	8	64
P	0	1	0	1	0	1	0	1	0	1	5	25
Q	1	0	1	1	1	1	1	1	1	1	9	81

R	0	1	0	1	1	1	0	1	0	1	6	36
S	1	0	0	1	1	1	1	1	1	1	8	64
T	0	1	1	0	0	0	1	0	0	1	4	16
N=20	10	12	10	14	13	15	12	16	12	16	130	934
P	0,5	0,6	0,5	0,7	0,65	0,75	0,6	0,8	0,6	0,8		
Q	0,5	0,4	0,5	0,3	0,35	0,25	0,4	0,2	0,4	0,2		

2. Menilai M_p dan M_t

$$M_t = \frac{\sum X_t}{N}$$

$$M_t = \frac{130}{20}$$

$$M_t = 6,5$$

$$M_p = \frac{\sum X_i}{n}$$

Keterangan :

$\sum X_i$ = Jumlah skor total testee yang menjawab benar pada soal yang diuji validitasnya

N = Jumlah testee yang menjawab benar pada soal yang diuji validitasnya

$$M_p = \frac{7 + 10 + 7 + 8 + 9 + 10 + 7 + 8 + 9 + 8}{10}$$

$$M_p = 8,3$$

3. Memasukkan nilai-nilai yang telah didapat ke dalam formula *point biserial*.

$$r_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{P}{q}}$$

Selengkapnya, pengujian validitas tiap item dalam instrumen *achievement tes* yang digunakan adalah sebagai berikut.

a. Validitas Item Nomor 1

Berdasarkan tabel kerja diatas diketemukan nilai-nilai sebagai berikut : $p = 0,5$ dan $q = 0,5$

$$P = \frac{\sum X_i}{N}$$

$$P = \frac{10}{20}$$

$$P = 0,5$$

$$q = 1 - 0,5 = 0,5$$

$$S_t = \sqrt{\left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2}$$

$$S_t = \sqrt{\left[\frac{934}{20}\right] - \left[\frac{130}{20}\right]^2}$$

$$S_t = \sqrt{[46,7] - [6,5]^2}$$

$$S_t = \sqrt{4,45}$$

$$S_t = 2,1095023109729$$

$$\gamma_{pbi} = \frac{Mp - Mt}{S_t} \sqrt{\frac{P}{q}}$$

$$\gamma_{pbi} = \frac{8,3 - 6,5}{2,1095023109729} \sqrt{\frac{0,5}{0,5}}$$

$$\gamma_{pbi} = 0,853281833652\sqrt{1}$$

$$\gamma_{pbi} = 0,853281833652$$

Kriteria Pengujian :

- Harga $r_{xy} = 1$ korelasi sempurna
- Harga $r_{xy} = 0.99-0.81$ korelasi sangat tinggi
- Harga $r_{xy} = 0.80-0.61$ korelasi tinggi
- Harga $r_{xy} = 0.60-0.41$ korelasi sedang
- Harga $r_{xy} = 0.40-0.21$ korelasi rendah
- Harga $r_{xy} = 0.20-0.01$ korelasi sangat rendah

Hasil yang didapat dari perhitungan korelasi point biserial ditemukan angka sebesar 0,853 termasuk dalam kategori tinggi (harga $r_{xy} = 0.99-0.81$ korelasi tinggi). Dengan soal nomor 1 dapat disimpulkan valid.

Cara lain

Tabel kerja untuk menghitung koefisien korelasi point biserial untuk item nomor 1 adalah sebagai berikut.

Tabel 20. Kerja Persiapan analisis Validitas soal Nomor :1

No. Testee	Skor item (Xi)	Total Y	Y jika X = 1	Y ²
1	0	9	0	81
2	1	49	49	2401
3	0	36	0	1296
4	1	100	100	10000
5	1	49	49	2401
6	0	9	0	81
7	1	64	64	4096
8	1	81	81	6561
9	0	25	0	625
10	1	100	100	10000
11	0	36	0	1296
12	0	25	0	625
13	0	16	0	256
14	1	49	49	2401

15	1	64	64	4096
16	0	25	0	625
17	1	81	81	6561
18	0	36	0	1296
19	1	64	64	4096
20	0	16	0	256
	10	934	701	59050

$$y_{pbi} = \frac{M_p - M_t}{S_t} \sqrt{\frac{P}{q}}$$

$$M_p = \frac{\sum XiY}{n}$$

Dimana:

M_p = Mean peserta menjawab benar yang dicari validatasnya

$\sum XiY$ = Jumlah perkalian antara peserta yang menjawab benar (X_i) pada item yang diuji dengan Total Y

n = Jumlah peserta yang menjawab benar.

$$M_p = \frac{\sum XiY}{n}$$

$$M_p = \frac{701}{10}$$

$$M_p = 70,1$$

$$M_t = \frac{\sum Y}{N}$$

Keterangan :

M_t = Mean total

$\sum Y$ = Jumlah Total Skor peserta yang menjawab benar

N = Jumlah peserta tes (*testee*)

$$M_t = \frac{934}{20}$$

$$M_t = 46,7$$

$$S_t = \sqrt{\frac{N\sum Y^2 - (\sum Y)^2}{N(N-1)}}$$

Keterangan :

S_t = Standar deviasi total

$\sum Y^2$ = Jumlah Kuadrat Total Skor peserta yang menjawab benar

$(\sum Y)^2$ = Kuadrat Total Skor peserta yang menjawab benar

N = Jumlah peserta tes (*testee*)

$$S_t = \sqrt{\frac{20 \times 59050 - (934)^2}{20(20-1)}}$$

$$S_t = \sqrt{\frac{1181000 - 872356}{20(19)}}$$

$$S_t = \sqrt{\frac{308644}{380}}$$

$$S_t = \sqrt{\frac{308644}{380}}$$

$$S_t = \sqrt{812,2210526}$$

$$S_t = 28,49949$$

$$P = \frac{\sum Xi}{N}$$

Keterangan :

p atau q = proporsi

$\sum Xi$ = Total Skor peserta yang menjawab benar

N = jumlah peserta tes (*testee*)

$$P = \frac{10}{20}$$

$$P = 0,5$$

$$q = 1 - 0,5 = 0,5$$

$$\gamma_{pbi} = \frac{Mp - Mt}{S_t} \sqrt{\frac{P}{q}}$$

$$\gamma_{pbi} = \frac{70,1 - 46,7}{28,49949} \sqrt{\frac{0,5}{0,5}}$$

$$\gamma_{pbi} = \frac{23,4}{28,49949} \sqrt{1}$$

$$\gamma_{pbi} = 0,821067 \times 1$$

$$\gamma_{pbi} = \mathbf{0,821067}$$

Cara lain uji validitas item dengan menggunakan korelasi Product Moment

Ada beberapa langkah yang harus dilakukan dalam melakukan analisis korelasi product moment sebagaimana kasus yang sama, yaitu :

- Membuat tabel kerja atau tabel penolong

Tabel Kerja. 21. Korelasi Product Moment

No. Resp	X	Y	XY	X ²	Y ²
1	0	3	0	0	9
2	1	7	7	1	49
3	0	6	0	0	36
4	1	10	10	1	100
5	1	7	7	1	49
6	0	3	0	0	9
7	1	8	8	1	64
8	1	9	9	1	81

9	0	5	0	0	25
10	1	10	10	1	100
11	0	6	0	0	36
12	0	5	0	0	25
13	0	4	0	0	16
14	1	7	7	1	49
15	1	8	8	1	64
16	0	5	0	0	25
17	1	9	9	1	81
18	0	6	0	0	36
19	1	8	8	1	64
20	0	4	0	0	16
N=20	10	130	83	10	934

- b. Memasukkan jumlah nilai-nilai $N=20$, $\sum X=10$, $\sum Y=130$, $\sum XY=83$, $\sum X^2=10$, dan $\sum Y^2=934$ yang telah didapat ke dalam formula korelasi Product Moment.

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

$$r_{xy} = \frac{20 \times 83 - 10 \times 130}{\sqrt{[20 \times 10 - (10)^2][20 \times 934 - (130)^2]}}$$

$$r_{xy} = \frac{1660 - 1300}{\sqrt{[200 - 100][18680 - 16900]}}$$

$$r_{xy} = \frac{360}{\sqrt{[100][1780]}}$$

$$r_{xy} = \frac{360}{\sqrt{178000}}$$

$$r_{xy} = \frac{360}{421.90046219458}$$

$$r_{xy} = 0,853281833652$$

$$r_{xy} = \mathbf{0,853}$$

Correlations			
		q1	Skor Total
q1	Pearson Correlation	1	.853**
	Sig. (2-tailed)		.000
	N	20	20
Skor Total	Pearson Correlation	.853**	1
	Sig. (2-tailed)	.000	
	N	20	20

** . Correlation is significant at the 0.01 level (2-tailed).

Kesimpulannya : Soal nomor 1 termasuk dalam kategori valid.

Dari jumlah soal sebanyak 10 soal, cara menyelesaikan atau mencari validitas butir secara manual dilakukan satu persatu sebagaimana contoh pada validitas soal nomor satu ini sampai dengan soal terakhir yakni soal nomor 10. Hasil perhitungan keseluruhan dengan program SPSS dapat dilihat pada tabel berikut

Correlations											
	Q1	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9	Q10	TOTALSKOR
Q1	1	-.816**	.400	.436	.314	.346	.408	.500*	.816**	.500*	.853**
Pearson Correlation		.000	.081	.054	.177	.135	.074	.025	.000	.025	.000
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q2	-.816**	1	-.204	-.312	-.171	-.236	-.250	-.408	-.667**	-.408	-.532*
Pearson Correlation			.388	.181	.471	.317	.288	.074	.001	.074	.016
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q3	.400	-.204	1	.000	.105	-.115	.408	-.250	.408	.000	.427
Pearson Correlation		.081	.388	1.000	.660	.628	.074	.288	.074	1.000	.061
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q4	.436	-.312	.000	1	.206	.630**	-.089	.491*	.356	.491*	.672**
Pearson Correlation		.054	.181	1.000	.384	.003	.709	.028	.123	.028	.001
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q5	.314	-.171	.105	.206	1	-.182	.685**	.157	.471*	-.105	.571**
Pearson Correlation		.177	.471	.660	.384	.444	.001	.508	.036	.660	.008
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q6	.346	-.236	-.115	.630**	-.182	1	-.236	.289	.000	.289	.356
Pearson Correlation		.135	.317	.628	.003	.444	.317	.217	1.000	.217	.124
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q7	.408	-.250	.408	-.089	.685**	-.236	1	-.153	.375	-.153	.484*
Pearson Correlation		.074	.288	.074	.709	.001	.317	.519	.103	.519	.031
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q8	.500*	-.408	-.250	.491*	.157	.289	-.153	1	.357	.375	.474*
Pearson Correlation		.025	.074	.288	.028	.508	.217	.519	.122	.103	.035
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q9	.816**	-.667**	.408	.356	.471*	.000	.375	.357	1	.357	.774**
Pearson Correlation		.000	.001	.074	.123	.036	1.000	.103	.122	.122	.000
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
Q10	.500*	-.408	.000	.491*	-.105	.289	-.153	.375	.357	1	.474*
Pearson Correlation		.025	.074	1.000	.028	.660	.217	.519	.103	.122	.035
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20
TOTALSKOR	.853**	-.532*	.427	.672**	.571**	.356	.484*	.474*	.774**	.474*	1
Pearson Correlation		.000	.016	.061	.001	.008	.124	.031	.035	.000	.035
Sig. (2-tailed)											
N	20	20	20	20	20	20	20	20	20	20	20

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Tabel 22. Hasil Pengujian 10 Soal Program SPSS dengan korelasi Point Biserial

Item	r_{xy}	$r_{tabel} \Rightarrow N=20$ dg 5%	Tanda Bintang (**/*)	Angka Signifikan si < 0,05	Keterangan
1	0,853**	0,444	**	0,000	Valid
2	-0,532*	0,444	*	0,016	Tidak Valid
3	0,427	0,444		0,061	Tidak Valid
4	0,672**	0,444	**	0,001	Valid
5	0,571**	0,444	**	0,008	Valid
6	0,356	0,444		0,124	Tidak Valid
7	0,484*	0,444	*	0,031	Valid
8	0,474*	0,444	*	0,035	Valid
9	0,774**	0,444	**	0,000	Valid
10	0,474*	0,444	*	0,035	Valid

Selanjutnya koefisien korelasi point biserial (γ_{pbi}) yang ada pada kolom 2 dari semua item, kemudian dibandingkan dengan harga r_{tabel} (kolom 3) untuk mengetahui validitas masing-masing item. Jika $\gamma_{pbi} > r_{tabel}$ maka item bersangkutan dinyatakan valid, sebaliknya jika $\gamma_{pbi} \leq r_{tabel}$ maka item bersangkutan dinyatakan tidak valid. Harga r_{xy} dari masing-masing item beserta

perbandingannya dengan r_{tabel} dan keputusan validitas item adalah seperti tercantum pada tabel tersebut.

1.2. Menentukan Uji reliabilitas instrument Tes

Ada banyak cara untuk menentukan reliabilitas instrument dengan melihat karakteristik dari instrumen yang akan diuji. Apakah berbentuk tes obyektif, seperti pilihan benar salah, pilihan ganda, sebab akibat, dan menjodohkan atau berbentuk tes subyektif, seperti tes uraian.

1.2.1. Pengujian Reliabilitas Instrumen untuk Item Berbentuk Tes Obyektif

Ada beberapa formula yang dapat digunakan untuk menguji reliabilitas instrumen yang dihasilkan dari tes berbentuk obyektif yang biasanya dikotomik, menjawab benar diberi skor 1 dan salah diberi skor (0). Dalam buku ini kita bicarakan hanya Spearman Brown dan KR-20.

a. Formula Spearman Brown

Formula Spearman Brown merupakan sebuah formula komputasi yang sangat populer untuk estimasi reliabilitas tes yang dibelah menjadi dua bagian yang saling paralel satu dengan yang lainnya. Artinya formula ini digunakan dengan belahan tes yang sama, dengan demikian soalnya berjumlah genap. Formula ini dapat digunakan untuk tes yang item-itemnya diberi skor dikotomik atau non dikotomik.

Contoh kasus untuk uji item dari 10 soal dari 10 testee

Tabel 23. Distribusi Hasil Skor tes dari 20 Testee pada 10 soal

Testee	Skor untuk butir item nomor:											
	1	2	3	4	5	Belahan 1	6	7	8	9	10	Belahan 2
A	0	1	0	0	0	1	1	0	0	0	1	2
B	1	0	1	0	1	3	0	1	1	1	1	4
C	0	1	0	1	1	3	0	0	1	1	1	3
D	1	1	1	1	1	5	1	1	1	1	1	5
E	1	0	1	1	0	3	1	0	1	1	1	4
F	0	1	0	0	0	1	1	0	1	0	0	2
G	1	0	0	1	1	3	1	1	1	1	1	5
H	1	0	1	1	1	4	1	1	1	1	1	5
I	0	1	0	1	0	2	1	0	1	0	1	3
J	1	1	1	1	1	5	1	1	1	1	1	5
K	0	1	1	1	1	4	1	1	0	0	0	2
L	0	1	1	0	1	3	0	1	0	1	0	2
M	0	1	0	0	1	2	0	1	1	0	0	2
N	1	0	1	1	0	3	1	0	1	1	1	4
O	1	0	0	1	1	3	1	1	1	1	1	5
P	0	1	0	1	0	2	1	0	1	0	1	3
Q	1	0	1	1	1	4	1	1	1	1	1	5
R	0	1	0	1	1	3	1	0	1	0	1	3

S	1	0	0	1	1	3	1	1	1	1	1	5
T	0	1	1	0	0	2	0	1	0	0	1	2

Ujilah Reliabilitas soal dari tabel tersebut diatas!

Ada beberapa langkah yang harus dilakukan dalam proses uji reliabilitas instrument dengan menggunakan Spearman Brow, yaitu:

1. Membuat tabel penolong

Tabel 24. Penolong Proses Perhitungan Korelasi Product Moment

No. Resp	X	Y	XY	X ²	Y ²
1	1	2	2	1	4
2	3	4	12	9	16
3	3	3	9	9	9
4	5	5	25	25	25
5	3	4	12	9	16
6	1	2	2	1	4
7	3	5	15	9	25
8	4	5	20	16	25
9	2	3	6	4	9
10	5	5	25	25	25
11	4	2	8	16	4
12	3	2	6	9	4
13	2	2	4	4	4
14	3	4	12	9	16
15	3	5	15	9	25
16	2	3	6	4	9
17	4	5	20	16	25
18	3	3	9	9	9
19	3	5	15	9	25
20	2	2	4	4	4
N=20	59	71	227	197	283

2. Memasukkan nilai yang didapat dari tabel diatas ditemukan N = 20, $\sum X = 59$, $\sum Y = 71$, $\sum XY = 227$, $\sum X^2 = 197$, dan $\sum Y^2 = 283$

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

$$r_{xy} = \frac{20 \times 227 - 59 \times 71}{\sqrt{[20 \times 197 - (59)^2][20 \times 283 - (71)^2]}}$$

$$r_{xy} = \frac{4540 - 4189}{\sqrt{[20 \times 197 - (59)^2][20 \times 283 - (71)^2]}}$$

$$r_{xy} = \frac{351}{\sqrt{[3940 - 3481] [5660 - 5041]}}$$

$$r_{xy} = \frac{351}{\sqrt{[459] [619]}}$$

$$r_{xy} = \frac{351}{\sqrt{284121}}$$

$$r_{xy} = \frac{351}{533,03001791644}$$

$$r_{xy} = \mathbf{0,6584994994691}$$

Hasil olah data Program SPSS

Correlations		X	Y
X	Pearson Correlation	1	.658**
	Sig. (2-tailed)		.002
	N	20	20
Y	Pearson Correlation	.658**	1
	Sig. (2-tailed)	.002	
	N	20	20

**. Correlation is significant at the 0.01 level (2-tailed).

Kemudian nilai koefisien korelasi yang ditemukan dimasukkan ke dalam formula Spearman Brown:

$$r_{1.2} + r = \frac{2 r_{1.2}}{1 + r_{1.2}}$$

$$r_{1.2} = \frac{2 \times 0,6584994994691}{1 + 0,6584994994691}$$

$$r_{1.2} = \frac{1,3169989989382}{1,6584994994691}$$

$$r_{1.2} = \mathbf{0,7940906821858}$$

Reliability Statistics			
Cronbach's Alpha	Part 1	Value	-.033 ^a
		N of Items	5 ^b
	Part 2	Value	.452
		N of Items	5 ^c
	Total N of Items		10
	Correlation Between Forms		.658
Spearman-Brown Coefficient	Equal Length		.794
	Unequal Length		.794
Guttman Split-Half Coefficient			.789

a. The value is negative due to a negative average covariance among items. This violates reliability model assumptions. You may want to check item codings.

b. The items are: q1, q2, q3, q4, q5.

c. The items are: q6, q7, q8, q9, q10.

Kriteria interpretasi hasil pengujiannya :

< 0,5	= Buruk
0,5 – 0,6	= Jelek
0,6 – 0,7	= Cukup
0,7 – 0,8	= Bagus
>0,8	= Bagus Sekali

Kesimpulannya : berdasarkan kriteria interpretasi tersebut diatas, maka dapat disimpulkan instrument tes tersebut dalam kategori bagus (0,7 – 0,8).

Kesimpulannya: instrument tes reliabilitasnya termasuk dalam kategori tinggi.

b. Formula KR-20

Formula Kuder-Richardson digunakan dengan karakteristik, apabila suatu tes berisi item-item yang diberi skor dikotomi sedangkan jumlah itemnya sendiri tidak begitu banyak, kadang-kadang membagi tes menjadi dua bagian tidak dapat menghasilkan bagian yang setara sedangkan membagi tes menjadi lebih dari dua belahan akan mengakibatkan jumlah item-item dalam setiap belahan terlalu sedikit. Bila belahan item sedikit, maka reliabilitasnya tidak menghasilkan estimasi yang cermat. Solusi yang dapat dilakukan adalah membelah belahan tes sesuai dengan jumlah item. Koefisien reliabilitas instrument tes tersebut dihitung dengan menggunakan rumus KR-20, yaitu:

$$r_{11} = \left(\frac{n}{n-1} \right) \left(\frac{s^2 - \sum pq}{s^2} \right)$$

n = jumlah aitem dalam tes

S = standar deviasi skor total

p = proporsi siswa yang menjawab benar untuk tiap-tiap item

q = proporsi siswa yang menjawab salah untuk tiap-tiap item

Sebelum perhitungan terlebih dahulu dibuat tabel kerja dengan langkah-langkah sebagai berikut.

- Item-item yang dinyatakan tidak valid dikeluarkan dari instrument. Jadi reliabilitas instrument dihitung hanya untuk item-item yang dinyatakan valid. *Dalam buku referensi yang lain, tidak mensyaratkan harus yang tidak valid dikeluarkan.*
- Menghitung p, yaitu proporsi siswa yang menjawab benar untuk tiap-tiap item
- Menghitung q, yaitu proporsi siswa yang menjawab salah untuk tiap-tiap item
- (q = 1-p)
- Menghitung hasil kali p dengan q (pq) untuk tiap-tiap item

Contoh :

Formula KR-20 dengan mempersyaratkan item yang tidak valid dikeluarkan dari tabel persiapan, yaitu pada item 3 dan 6.

Tabel 25. Persiapan Distribusi Hasil Skor tes dari 20 Testee pada 8 soal

Testee	Skor untuk butir item nomor:								Skor Total (X _i)	X _i ²
	1	2	4	5	7	8	9	10		
A	0	1	0	0	0	0	0	1	2	4

B	1	0	0	1	1	1	1	1	6	36
C	0	1	1	1	0	1	1	1	6	36
D	1	1	1	1	1	1	1	1	8	64
E	1	0	1	0	0	1	1	1	5	25
F	0	1	0	0	0	1	0	0	2	4
G	1	0	1	1	1	1	1	1	7	49
H	1	0	1	1	1	1	1	1	7	49
I	0	1	1	0	0	1	0	1	4	16
J	1	1	1	1	1	1	1	1	8	64
K	0	1	1	1	1	0	0	0	4	16
L	0	1	0	1	1	0	1	0	4	16
M	0	1	0	1	1	1	0	0	4	16
N	1	0	1	0	0	1	1	1	5	25
O	1	0	1	1	1	1	1	1	7	49
P	0	1	1	0	0	1	0	1	4	16
Q	1	0	1	1	1	1	1	1	7	49
R	0	1	1	1	0	1	0	1	5	25
S	1	0	1	1	1	1	1	1	7	49
T	0	1	0	0	1	0	0	1	3	9
N=20	10	12	14	13	12	16	12	16	105	617
P	0.5	0.6	0.7	0.65	0.6	0.8	0.6	0.8		
Q	0.5	0.4	0.3	0.35	0.4	0.2	0.4	0.2		
Pq	0.25	0.24	0.21	0.2275	0.24	0.16	0.24	0.16	1.7275	

Berdasarkan tabel diatas diketahui $\sum pq = 1,7275$; $\sum Xt = 105$; $\sum Xt^2 = 617$; $N = 20$. Langkah selanjutnya yaitu mencari nilai koefisien Varians) untuk skor total, dengan menggunakan rumus :

$$S_t^2 = \frac{N\sum Y^2 - (\sum Y)^2}{N(N-1)}$$

$$S_t^2 = \frac{20 \times 617 - (105)^2}{20(20-1)}$$

$$S_t^2 = \frac{12340 - 11025}{20(19)}$$

$$S_t^2 = \frac{1315}{380}$$

$$S_t^2 = 3,4605263157895$$

Setelah semua langkah diatas dilakukan, kemudian dihitung harga koefisien reliabilitas instrument. Dengan menggunakan rumus KR-20 sebagai berikut :

$$r_{11} = \left[\frac{n}{n-1} \right] - \left[\frac{S^2 - \sum pq}{S^2} \right]$$

$$r_{11} = \left[\frac{8}{8-1} \right] - \left[\frac{3,4605263157895 - 1,7275}{3,4605263157895} \right]$$

$$r_{11} = \left[\frac{8}{7} \right] - \left[\frac{1,7330263157895}{3,4605263157895} \right]$$

$$r_{11} = 1,1428571428571 - [0,5007984790875]$$

$$r_{11} = \mathbf{0,6420586637696}$$

Kesimpulan : Instrumen Tes Reliabilitasnya Cukup.

Contoh :

Formula KR-20 dengan melibatkan semua item.

Tabel 26. Persiapan Distribusi Hasil Skor tes dari 20 Testee pada 10 soal

Tes Tee	Skor untuk butir item nomor:										Skor Total (X _i)	Xt ²
	1	2	3	4	5	6	7	8	9	10		
A	0	1	0	0	0	1	0	0	0	1	3	9
B	1	0	1	0	1	0	1	1	1	1	7	49
C	0	1	0	1	1	0	0	1	1	1	6	36
D	1	1	1	1	1	1	1	1	1	1	10	100
E	1	0	1	1	0	1	0	1	1	1	7	49
F	0	1	0	0	0	1	0	1	0	0	3	9
G	1	0	0	1	1	1	1	1	1	1	8	64
H	1	0	1	1	1	1	1	1	1	1	9	81
I	0	1	0	1	0	1	0	1	0	1	5	25
J	1	1	1	1	1	1	1	1	1	1	10	100
K	0	1	1	1	1	1	1	0	0	0	6	36
L	0	1	1	0	1	0	1	0	1	0	5	25
M	0	1	0	0	1	0	1	1	0	0	4	16
N	1	0	1	1	0	1	0	1	1	1	7	49
O	1	0	0	1	1	1	1	1	1	1	8	64
P	0	1	0	1	0	1	0	1	0	1	5	25
Q	1	0	1	1	1	1	1	1	1	1	9	81
R	0	1	0	1	1	1	0	1	0	1	6	36
S	1	0	0	1	1	1	1	1	1	1	8	64
T	0	1	1	0	0	0	1	0	0	1	4	16
N= 20	10	12	10	14	13	15	12	16	12	16	130	934
P	0.5	0.6	0.5	0.7	0.65	0.75	0.6	0.8	0.6	0.8	6.5	
Q	0.5	0.4	0.5	0.3	0.35	0.25	0.4	0.2	0.4	0.2		
p(1-p)	0.25	0.24	0.25	0.21	0.2275	0.187	0.24	0.16	0.24	0.16	2.165	

Berdasarkan tabel diatas diketahui $\sum p(1-p) = 2,165$; $\sum X_t = 130$; $\sum X_t^2 = 934$; $N = 20$. Langkah selanjutnya yaitu mencari nilai koefisien Varians untuk skor total, dengan menggunakan rumus :

$$S_t^2 = \frac{\sum X^2 - \frac{(\sum X)^2}{N}}{N}$$

$$S_t^2 = \frac{934 - \frac{(130)^2}{20}}{20}$$

$$S_t^2 = \frac{934 - \frac{(16900)}{20}}{20}$$

$$S_t^2 = \frac{934 - 845}{20}$$

$$S_t^2 = \frac{89}{20}$$

$$S_t^2 = 4,45$$

Atau

$$S_t^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_t^2 = \left[\frac{934}{20} \right] - \left[\frac{130}{20} \right]^2$$

$$S_t^2 = [46,7] - [6,5]^2$$

$$S_t^2 = [46,7] - [42,25]$$

$$S_t^2 = 4,45$$

Setelah semua langkah diatas dilakukan, kemudian dihitung harga koefisien reliabilitas instrument. Dengan menggunakan rumus KR-20 sebagai berikut :

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum p(1-p)}{S^2} \right]$$

Dimana :

k = banyaknya aitem

S^2 = Varians skor total

p dan q (1-p) = proporsi subyek yang menjawab aitem

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum p(1-p)}{S^2} \right]$$

$$r_{11} = \left[\frac{10}{10-1} \right] \left[1 - \frac{2,165}{4,45} \right]$$

$$r_{11} = \left[\frac{10}{9} \right] [1 - 0,4865168539326]$$

$$r_{11} = [1,111111111111] [0,513483146067]$$

$$r_{11} = 0,5705368289633$$

Atau

$$r_{11} = \left[\frac{n}{n-1} \right] \left[\frac{S^2 - \sum pq}{S^2} \right]$$

$$r_{11} = \left[\frac{10}{10-1} \right] \left[\frac{4,45 - 2,165}{4,45} \right]$$

$$r_{11} = \left[\frac{10}{9} \right] \left[\frac{2,285}{4,45} \right]$$

$$r_{11} = 1,111111111 \times [0,5134831460674]$$

$$r_{11} = 0,5705368289632$$

Kesimpulan : Instrumen Tes Reliabilitasnya rendah.

1.2.2. Pengujian Reliabilitas Instrumen untuk Item Berbentuk Tes Subyektif

Formula untuk menguji tes yang berbentuk uraian dapat menggunakan formula Alpha Cronbach. Dimana hasil pengukuran itu bisa dinilai mulai 0-5, 0-8 atau bentuk 0-10.

Contoh kasus.

Tes kemampuan dengan menggunakan 6 soal dan diikuti oleh peserta tes (*testee*) sebanyak

Tabel 27. Persiapan Data Hasil Skoring dari 6 soal dari 10 peserta tes

Testee	Skor untuk butir item nomor:						Skor Total (X _i)	X _i ²
	1	2	3	4	5	6		
A	8	6	8	8	10	10	50	2500
B	7	4	4	6	6	5	32	1024
C	8	2	6	8	7	8	39	1521
D	7	5	7	7	6	10	42	1764
E	5	5	5	7	5	4	31	961
F	2	4	2	8	6	8	30	900
G	4	5	6	6	6	6	33	1089
H	5	5	5	7	8	7	37	1369
I	5	5	4	6	8	5	33	1089
J	3	6	5	5	8	6	33	1089
N=15	54	47	52	68	70	69	360	13306
Jumlah Kuadrat	330	233	296	472	510	515		

Mencari nilai varians total (S_t²) dengan formula dibawah ini.

$$S_t^2 = \frac{\sum X_t^2 - \left[\frac{\sum X_t}{N} \right]^2}{N}$$

var tot	ΣX ²	ΣX	(ΣX) ²	N	(ΣX) ² /N	ΣX ² - (ΣX) ² /N	ΣX ² - (ΣX) ² /N : N
	13306	360	129600	10	12960	346	34.6

Mencari nilai varians total

$$\begin{aligned}
 S_t^2 &= \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2 \\
 S_t^2 &= \left[\frac{13306}{10} \right] - \left[\frac{360}{10} \right]^2 \\
 S_t^2 &= [1330,6] - [36]^2 \\
 S_t^2 &= [1330,6] - [1296] \\
 S_t^2 &= 34,6
 \end{aligned}$$

Mencari nilai varians masing-masing item dan dijumlahkan masing-masing varians item.

Tabel 28. Jumlah Kuadrat dari Masing-Masing Skor Item

1		2		3		4		5		6	
8	64	6	36	8	64	8	64	10	100	10	100
7	49	4	16	4	16	6	36	6	36	5	25
8	64	2	4	6	36	8	64	7	49	8	64
7	49	5	25	7	49	7	49	6	36	10	100
5	25	5	25	5	25	7	49	5	25	4	16
2	4	4	16	2	4	8	64	6	36	8	64
4	16	5	25	6	36	6	36	6	36	6	36
5	25	5	25	5	25	7	49	8	64	7	49
5	25	5	25	4	16	6	36	8	64	5	25
3	9	6	36	5	25	5	25	8	64	6	36
54	330	47	233	52	296	68	472	70	510	69	515

Berdasarkan jumlah kuadrat dan skor total item, maka dapat dicari Varians item (S_i^2) masing-masing pada setiap item.

$$S_i^2 = \frac{\sum X_i^2 - \left[\frac{\sum X_i}{N} \right]^2}{N}$$

No. Item	$\sum X^2$	$\sum X$	$(\sum X)^2$	N	$(\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N : N$
item 1	330	54	2916	10	291.6	38.4	3.84
item 2	233	47	2209	10	220.9	12.1	1.21
item 3	296	52	2704	10	270.4	25.6	2.56
item 4	472	68	4624	10	462.4	9.6	0.96
item 5	510	70	4900	10	490	20	2
item 6	515	69	4761	10	476.1	38.9	3.89
							14.46

$$S_1^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_1^2 = \left[\frac{330}{10} \right] - \left[\frac{54}{10} \right]^2$$

$$S_1^2 = [33] - [5,4]^2$$

$$S_1^2 = [33] - [29,16]$$

$$S_1^2 = \mathbf{3,84}$$

$$S_2^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_2^2 = \left[\frac{233}{10} \right] - \left[\frac{47}{10} \right]^2$$

$$S_2^2 = [23,3] - [4,7]^2$$

$$S_2^2 = [23,3] - [22,09]$$

$$S_2^2 = \mathbf{1,21}$$

$$S_3^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_3^2 = \left[\frac{296}{10} \right] - \left[\frac{52}{10} \right]^2$$

$$S_3^2 = [29,6] - [5,2]^2$$

$$S_3^2 = [29,6] - [27,04]$$

$$S_3^2 = \mathbf{2,56}$$

$$S_4^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_4^2 = \left[\frac{472}{10} \right] - \left[\frac{68}{10} \right]^2$$

$$S_4^2 = [47,2] - [6,8]^2$$

$$S_4^2 = [47,2] - [46,24]$$

$$S_4^2 = \mathbf{0,96}$$

$$S_5^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_5^2 = \left[\frac{510}{10} \right] - \left[\frac{70}{10} \right]^2$$

$$S_5^2 = [51] - [7]^2$$

$$S_5^2 = [51] - [49]$$

$$S_5^2 = \mathbf{2}$$

$$S_6^2 = \left[\frac{\sum X t^2}{N} \right] - \left[\frac{\sum X t}{N} \right]^2$$

$$S_6^2 = \left[\frac{515}{10} \right] - \left[\frac{69}{10} \right]^2$$

$$S_6^2 = [51,5] - [6,9]^2$$

$$S_6^2 = [51,5] - [47,61]$$

$$S_6^2 = 3,89$$

Menjumlahkan masing-masing varians pada setiap item

$$\Sigma S_i^2 = 3,84 + 1,21 + 2,56 + 0,96 + 2 + 3,89 = \underline{14,46}$$

Memasukkan kedalam formulan Alpha Cronbach;

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\Sigma S_i^2}{St^2} \right]$$

$$r_{11} = \left[\frac{6}{6-1} \right] \left[1 - \frac{14,46}{34,6} \right]$$

$$r_{11} = \left[\frac{6}{5} \right] [1 - 0,4179190751445]$$

$$r_{11} = 1,2 [0,582080924856]$$

$$r_{11} = \underline{0,6984971098272}$$

Reliability Statistics

Cronbach's Alpha	N of Items
.698	6

Kriteria interpretasi hasil pengujiannya :

< 0,5	= Buruk
0,5 – 0,6	= Jelek
0,6 – 0,7	= Cukup
0,7 – 0,8	= Bagus
>0,8	= Bagus Sekali

Kesimpulannya : berdasarkan kriteria interpretasi tersebut diatas, maka dapat disimpulkan instrument non tes tersebut dalam kategori cukup (0,60 – 0,7).

1.3. Uji taraf Kesukaran Item

Mengingat instrumen yang diuji coba merupakan *achievement test* maka setelah validitas dan reliabilitas tes diketahui maka dilanjutkan dengan uji taraf (indeks) kesukaran item. Rumus yang digunakan adalah :

$$I = \frac{B}{N}$$

Keterangan =

I = Indeks kesukaran item

B = Banyaknya siswa yang menjawab item tersebut dengan benar

N = Jumlah siswa yang mengikuti tes

Tabel 29. Skoring Hasil Uji Coba *Achievement Test* dari 20 Orang

Tes-tee	Skor untuk butir item nomor:									
	1	2	3	4	5	6	7	8	9	10

		1	0	0	A	0	0	0	0	1
B	1	0	1	0	1	0	1	1	1	1
C	0	1	0	1	1	0	0	1	1	1
D	1	1	1	1	1	1	1	1	1	1
E	1	0	1	1	0	1	0	1	1	1
F	0	1	0	0	0	1	0	1	0	0
G	1	0	0	1	1	1	1	1	1	1
H	1	0	1	1	1	1	1	1	1	1
I	0	1	0	1	0	1	0	1	0	1
J	1	1	1	1	1	1	1	1	1	1
K	0	1	1	1	1	1	1	0	0	0
L	0	1	1	0	1	0	1	0	1	0
M	0	1	0	0	1	0	1	1	0	0
N	1	0	1	1	0	1	0	1	1	1
O	1	0	0	1	1	1	1	1	1	1
P	0	1	0	1	0	1	0	1	0	1
Q	1	0	1	1	1	1	1	1	1	1
R	0	1	0	1	1	1	0	1	0	1
S	1	0	0	1	1	1	1	1	1	1
T	0	1	1	0	0	0	1	0	0	1
N=20	10	12	10	14	13	15	12	16	12	16

Soal nomor 1. B = 10 dan N = 20 maka :

$$I = \frac{B}{N}$$

$$I = \frac{10}{20}$$

$$I = 0,5$$

Berdasarkan cara yang sama diperoleh indeks kesukaran item nomor 3, 4, 5, dan seterusnya, sampai dengan item nomor 10. Selanjutnya disesuaikan dengan kriteria sebagai berikut :

Item dengan P 0,00 sampai 0,30 tergolong sukar

Item dengan P 0,31 sampai 0,70 tergolong sedang

Item dengan P 0,71 sampai 1,00 tergolong mudah

Berdasarkan patokan tersebut maka indeks kesukaran beserta kriteria kesukaran tiap-tiap item yang diujicobakan dapat dilihat pada table dihalaman berikut:

Tabel 30. Hasil Analisis Tingkat Kesukaran Soal Tes

No. Item	Banyaknya siswa yang menjawab item dengan benar (B)	N	Indeks Kesukaran Item (I) = B/N	Keterangan
1	10	20	0.5	Sedang

2	12	20	0.6	Sedang
3	10	20	0.5	Sedang
4	14	20	0.7	Sedang
5	13	20	0.65	Sedang
6	15	20	0.75	Mudah
7	12	20	0.6	Sedang
8	16	20	0.8	Mudah
9	12	20	0.6	Sedang
10	16	20	0.8	Mudah

1.4. Analisis Daya Pembeda

Analisis daya pembeda adalah pengkajian butir-butir soal yang dimaksudkan untuk mengetahui kesanggupan siswa untuk membedakan siswa yang tergolong mampu dengan siswa yang tergolong tidak mampu. Cara yang digunakan untuk analisis daya pembeda adalah dengan menggunakan kriteria Rose dan Stanley. Rumusnya :

SR-ST

SR = Jumlah siswa yang menjawab salah kelompok rendah

ST = Jumlah siswa yang menjawab salah kelompok tinggi

Tabel Persiapan Analisis Daya Pembeda Tes dengan tabel Ranging Skoring dari Hasil Uji Coba *Achievement Test* dari 20 Orang

Testee	Skor untuk butir item nomor:										Skor Total (X_i)
	1	2	3	4	5	6	7	8	9	10	
D	1	1	1	1	1	1	1	1	1	1	10
J	1	1	1	1	1	1	1	1	1	1	10
H	1	0	1	1	1	1	1	1	1	1	9
Q	1	0	1	1	1	1	1	1	1	1	9
O	1	0	0	1	1	1	1	1	1	1	8
S	1	0	0	1	1	1	1	1	1	1	8
G	1	0	0	1	1	1	1	1	1	1	8
B	1	0	1	0	1	0	1	1	1	1	7
E	1	0	1	1	0	1	0	1	1	1	7
N	1	0	1	1	0	1	0	1	1	1	7
R	0	1	0	1	1	1	0	1	0	1	6
C	0	1	0	1	1	0	0	1	1	1	6
K	0	1	1	1	1	1	1	0	0	0	6
L	0	1	1	0	1	0	1	0	1	0	5
P	0	1	0	1	0	1	0	1	0	1	5
I	0	1	0	1	0	1	0	1	0	1	5

M	0	1	0	0	1	0	1	1	0	0	4
T	0	1	1	0	0	0	1	0	0	1	4
A	0	1	0	0	0	1	0	0	0	1	3
F	0	1	0	0	0	1	0	1	0	0	3

Berdasarkan table diatas, terlihat $N = 20$ orang kelompok atas/bawah = 27% dari 20 = 5,4 atau 5. Kelompok tinggi dan kelompok rendah diperoleh dengan terlebih dahulu meranking peserta berdasarkan skor yang diperoleh. Kemudian 27% dari siswa yang memperoleh skor tertinggi diambil sebagai kelompok tinggi, dan 27% dari siswa yang memperoleh skor terendah diambil sebagai kelompok rendah.

Tabel 31. Persiapan Analisis Daya Pembeda Tes

No. Item	SR	ST	SR-ST
1	5	0	5
2	0	3	-3
3	4	1	3
4	4	0	4
5	4	0	4
6	2	0	2
7	3	0	3
8	2	0	2
9	5	0	5
10	2	0	2

Mengingat peserta tes sebanyak 20 orang maka kelompok tinggi dan kelompok rendah masing-masing diambil sebanyak $27\% \times 20 = 5,4$ atau 5 orang. Berdasarkan hasil uji coba diperoleh table kerja untuk analisis daya pembeda seperti table di atas. Kriteria yang digunakan dari table Ros dan Stanley adalah sebagai berikut:

Jumlah Testi	N (27%N)	Option			
		2	3	4	5
28-31	8	4	5	5	5
32-35	9	5	5	5	5
36-38	10	5	5	5	5

Dan seterusnya

Mengingat banyak testee dalam uji coba tes sebanyak 20 orang lebih kecil dari interval testee sebesar 28-31, maka banyak option pada instrument ada 2 maka kriteria yang digunakan adalah 4. Dengan demikian diperoleh hasil analisis daya pembeda sebagai berikut.

Tabel 32. Hasil Analisis Daya Pembeda 10 soal

Nomor Item	SR	ST	SR-ST	Kriteria	Keterangan
1	5	0	5	4	Diterima
2	0	3	-3	4	Ditolak
3	4	1	3	4	Ditolak
4	4	0	4	4	Diterima
5	4	0	4	4	Diterima
6	2	0	2	4	Ditolak
7	3	0	3	4	Ditolak
8	2	0	2	4	Ditolak
9	5	0	5	4	Diterima
10	2	0	2	4	Ditolak

1.5. Analisis Efektifitas Pengecoh (*Distractor*)

Analisis efektifitas pengecoh (*distractor*) atau analisis pola jawaban dilakukan dengan menghitung persentase yang memilih tiap alternative jawaban pada masing-masing item. Kriteria pengecoh yang baik adalah apabila pengecoh tersebut dipilih oleh paling sedikit 5% dari peserta tes. Mengingat banyak peserta tes 20 orang maka pengecoh dikatakan baik jika dipilih oleh minimal 5% dari 20 orang, atau minimal oleh 1 orang peserta tes. Data tentang banyaknya siswa yang memilih tiap-tiap pilihan pada masing-masing item, beserta keterangan baik-tidaknya pengecoh adalah seperti tercantum pada table berikut.

Tabel 33. Data Jawaban Hasil Uji Coba *Achievement Test* dari 20 Orang

Testee	Skor untuk butir pada item nomor									
	1	2	3	4	5	6	7	8	9	10
Kunci Jawaban	A	B	B	D	B	C	D	B	B	C
A	B	B	D	C	C	C	C	C	C	C
B	A	D	B	C	B	D	D	B	B	C
C	D	B	D	D	B	D	C	B	B	C
D	A	B	B	D	B	C	D	B	B	C
E	A	C	B	D	D	C	C	B	B	C
F	B	B	D	C	D	C	C	B	C	D
G	A	C	C	D	B	C	D	B	B	C
H	A	C	B	D	B	C	D	B	B	C
I	C	B	D	D	D	C	B	B	D	C
J	A	B	B	D	B	C	D	B	B	C
K	B	B	B	D	B	C	D	D	D	D
L	B	B	B	B	B	B	D	C	B	B
M	B	B	C	C	B	D	D	B	D	B
N	A	D	B	D	D	C	C	B	B	C

O	A	D	C	D	B	C	D	B	B	C
P	B	B	D	D	D	C	B	B	D	C
Q	A	D	B	D	B	C	D	B	B	C
R	B	B	D	D	B	C	B	B	A	C
S	A	D	A	D	B	C	D	B	B	C
T	D	B	B	A	C	D	D	D	C	C

Tabel 34. Data Pilihan Siswa dan Pengecoh

No. Item	Pilihan Jawaban				Jumlah Teste	Kunci Jawaban	Keterangan
	A	B	C	D			
1	10	7	1	2	20	A	Semua Pengecoh Baik
2	0	12	3	5	20	B	Pengecoh a kurang baik
3	1	10	3	6	20	B	Semua Pengecoh Baik
4	1	1	4	14	20	D	Semua Pengecoh Baik
5	0	13	2	5	20	B	Pengecoh a kurang baik
6	0	1	15	4	20	C	Pengecoh a kurang baik
7	0	3	5	12	20	D	Pengecoh a kurang baik
8	0	16	2	2	20	B	Pengecoh a kurang baik
9	1	12	3	4	20	B	Semua Pengecoh Baik
10		2	16	2	20	C	Pengecoh a kurang baik

2. Pengujian Instrumen Non Tes

2.1. Uji Validitas Instrumen Non Tes

Uji validitas item pada instrumen non tes dilakukan dengan menggunakan korelasi Product Moment. Ada beberapa langkah yang harus dilakukan dalam melakukan analisis korelasi product moment sebagaimana kasus yang sama, yaitu :

- Membuat tabel kerja atau tabel penolong.
- Memasukkan jumlah nilai-nilai $N=20$, $\sum X=10$, $\sum Y=130$, $\sum XY=83$, $\sum X^2=10$, dan $\sum Y^2=934$ yang telah didapat ke dalam formula korelasi Product Moment.
- Interpretasi.

Contoh Kasus:

Dilakukan penilaian tentang sikap siswa dalam belajar dari 5 item dari 21 siswa, dimana skala Likert A = 5, B = 4, C = 3, D = 2, E = 1 diketemukan data sebagai berikut:

Tabel 35. Item dan Jawaban Peserta Didik dari 21 Siswa

No. Peserta Didik	Butir Soal dan Skor Jawaban				
	1	2	3	4	5
1	C	B	C	B	A
2	D	D	C	B	C

3	C	A	B	D	D
4	B	A	A	C	D
5	C	B	D	B	A
6	B	B	A	A	B
7	A	B	A	A	A
8	C	A	B	C	D
9	A	A	B	A	C
10	C	B	A	C	D
11	B	A	B	C	D
12	C	A	C	A	C
13	B	A	B	A	D
14	C	C	C	B	A
15	B	A	B	B	C
16	A	A	A	C	B
17	B	A	B	C	B
18	B	A	B	B	A
19	B	A	B	A	B
20	B	B	A	B	C
21	C	B	B	E	B

Tabel 36. Skor dan Jawaban Peserta Didik dari 21 Siswa

No. Peserta Didik	Butir Soal dan Skor Jawaban					Skor total
	1	2	3	4	5	
1	3	4	3	4	5	19
2	1	2	3	4	3	13
3	3	5	4	2	2	16
4	4	5	5	3	2	19
5	3	4	2	4	5	18
6	4	4	5	5	4	22
7	5	4	5	5	5	24
8	3	5	4	3	2	17
9	5	5	4	5	3	22
10	3	4	5	3	2	17
11	4	5	4	3	2	18
12	3	5	3	5	3	19

13	4	5	4	5	2	20
14	3	3	3	4	5	18
15	4	5	4	4	3	20
16	5	5	5	3	4	23
17	4	5	4	3	4	20
18	4	5	4	4	5	22
19	4	5	4	5	4	22
20	4	4	5	4	3	20
21	3	4	4	1	4	19

Ujilah Validitas item dari data tersebut !

Jawab:

1. Membuat tabel kerja atau tabel penolong.

Tabel 37. Penolong Uji Validitas

No. Resp	X	Y	XY	X ²	Y ²
1	3	19	57	9	361
2	1	13	13	1	169
3	3	16	48	9	256
4	4	19	76	16	361
5	3	18	54	9	324
6	4	22	88	16	484
7	5	24	120	25	576
8	3	17	51	9	289
9	5	22	110	25	484
10	3	17	51	9	289
11	4	18	72	16	324
12	3	19	57	9	361
13	4	20	80	16	400
14	3	18	54	9	324
15	4	20	80	16	400
16	5	23	115	25	529
17	4	20	80	16	400
18	4	22	88	16	484
19	4	22	88	16	484
20	4	20	80	16	400

21	3	19	57	9	361
N = 21	76	408	1519	292	8060

2. Memasukkan jumlah nilai-nilai N= 21, $\sum X = 76$, $\sum Y = 408$, $\sum XY = 1519$, $\sum X^2 = 292$, dan $\sum Y^2 = 8060$ yang telah didapat ke dalam formula korelasi Product Moment.

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

$$r_{xy} = \frac{21 \times 1519 - 76 \times 408}{\sqrt{[21 \times 292 - (76)^2][21 \times 8060 - (408)^2]}}$$

$$r_{xy} = \frac{31899 - 31008}{\sqrt{[6132 - (5776)][169260 - (166464)]}}$$

$$r_{xy} = \frac{891}{\sqrt{[356][2796]}}$$

$$r_{xy} = \frac{891}{\sqrt{995376}}$$

$$r_{xy} = \frac{891}{997.68532113087}$$

$$r_{xy} = 0.8930671636925$$

$$r_{xy} = \mathbf{0,893}$$

Correlations			
		q1	Skor total X
q1	Pearson Correlation	1	.893**
	Sig. (2-tailed)		.000
	N	21	21
Skor total X	Pearson Correlation	.893**	1
	Sig. (2-tailed)	.000	
	N	21	21

** . Correlation is significant at the 0.01 level (2-tailed).

3. Interpretasinya = terdapat korelasi dengan kategori tinggi antara X (q1) dan Y (Skor Total X). Hal ini terlihat dari angka signifikan sebesar $0,000 < 0,05$ dan ada tanda bintang **, sehingga dapat dikatakan item atau soal nomor 1 termasuk *valid*.

Dari jumlah soal sebanyak 5 soal, cara menyelesaikan atau mencari validitas butir secara manual dilakukan satu persatu sebagaimana contoh pada validitas soal nomor satu ini sampai dengan soal terakhir yakni soal nomor 5. Hasil perhitungan keseluruhan dengan program SPSS dapat dilihat pada tabel berikut

Correlations							
		q1	q2	q3	q4	q5	skortotx
q1	Pearson Correlation	1	.632**	.584**	.254	.113	.893**
	Sig. (2-tailed)		.002	.005	.267	.625	.000
	N	21	21	21	21	21	21
q2	Pearson Correlation	.632**	1	.295	-.048	-.310	.458*
	Sig. (2-tailed)	.002		.194	.835	.171	.037
	N	21	21	21	21	21	21
q3	Pearson Correlation	.584**	.295	1	-.110	-.308	.463*
	Sig. (2-tailed)	.005	.194		.636	.175	.034
	N	21	21	21	21	21	21
q4	Pearson Correlation	.254	-.048	-.110	1	.242	.411
	Sig. (2-tailed)	.267	.835	.636		.291	.064
	N	21	21	21	21	21	21
q5	Pearson Correlation	.113	-.310	-.308	.242	1	.435*
	Sig. (2-tailed)	.625	.171	.175	.291		.049
	N	21	21	21	21	21	21
skortotx	Pearson Correlation	.893**	.458*	.463*	.411	.435*	1
	Sig. (2-tailed)	.000	.037	.034	.064	.049	
	N	21	21	21	21	21	21

** . Correlation is significant at the 0.01 level (2-tailed).

* . Correlation is significant at the 0.05 level (2-tailed).

Tabel 38. Hasil Pengujian 5 Soal Program SPSS dengan korelasi Product Moment dan Point Biserial

Item	r_{xy}	$r_{\text{tabel}} \Rightarrow N=21$ dg 5%	Tanda Bintang (**/*)	Angka Signifikan $si < 0,05$	Keterangan
1	0,893**	0,433	**	0,000	Valid
2	0,458*	0,433	*	0,037	Valid
3	0,463*	0,433		0,034	Valid
4	0,411	0,433		0,064	Tidak Valid
5	0,435*	0,433	*	0,049	Valid

2.2. Uji Reliabilitas Instrumen Non Tes

Formula untuk menguji reliabilitas instrument non tes dapat menggunakan formula *Alpha Cronbach*. Dimana hasil pengukuran itu bisa dinilai mulai 0-5, 0-8 atau bentuk 0-10.

Contoh kasus.

Dilakukan penilaian tentang sikap siswa dalam belajar dari 5 item dari 21 siswa, dimana skala Likert A = 5, B = 4, C = 3, D = 2, E = 1 diketemukan data sebagaimana data pada uji validitas instrumen non tes.

Tabel 39. Persiapan Data Hasil Skoring dari 5 item dari 21 peserta tes

Testee	Skor untuk butir item nomor:					Skor Total (X_t)	X_t^2
	1	2	3	4	5		
A	3	4	3	4	5	19	361
B	1	2	3	4	3	13	169

C	3	5	4	2	2	16	256
D	4	5	5	3	2	19	361
E	3	4	2	4	5	18	324
F	4	4	5	5	4	22	484
G	5	4	5	5	5	24	576
H	3	5	4	3	2	17	289
I	5	5	4	5	3	22	484
J	3	4	5	3	2	17	289
K	4	5	4	3	2	18	324
L	3	5	3	5	3	19	361
M	4	5	4	5	2	20	400
N	3	3	3	4	5	18	324
O	4	5	4	4	3	20	400
P	5	5	5	3	4	22	484
Q	4	5	4	3	4	20	400
R	4	5	4	4	5	22	484
S	4	5	4	5	4	22	484
T	4	4	5	4	3	20	400
U	3	4	4	1	4	16	256
N= 21	76	93	84	79	72	404	7910
Jumlah Kuadrat	292	425	350	321	274		

Langkah selanjutnya adalah mencari nilai varians total dalam varians setiap item.

Mencari varians total (S_t^2)

$$S_t^2 = \frac{\sum X_t^2 - \left[\frac{\sum X_t}{N} \right]^2}{N}$$

var	$\sum X^2$	$\sum X$	$(\sum X)^2$	N	$(\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N : N$
tot	7910	404	163216	21	7772.2	137.81	6.562358277

Atau mencari nilai varians total dengan formula di bawah ini.

$$S_t^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_t^2 = \left[\frac{7910}{21} \right] - \left[\frac{404}{21} \right]^2$$

$$S_t^2 = [376,666666666667] - [19,238095238095]^2$$

$$S_t^2 = [376,666666666667] - [370,10430839001]$$

$$S_t^2 = 6,56235827666$$

$$S_t^2 = \underline{\underline{6,562}}$$

Tabel 40. Jumlah Kuadrat dari Masing-Masing Skor Item

1		2		3		4		5	
3	9	4	16	3	9	4	16	5	25
1	1	2	4	3	9	4	16	3	9
3	9	5	25	4	16	2	4	2	4
4	16	5	25	5	25	3	9	2	4
3	9	4	16	2	4	4	16	5	25
4	16	4	16	5	25	5	25	4	16
5	25	4	16	5	25	5	25	5	25
3	9	5	25	4	16	3	9	2	4
5	25	5	25	4	16	5	25	3	9
3	9	4	16	5	25	3	9	2	4
4	16	5	25	4	16	3	9	2	4
3	9	5	25	3	9	5	25	3	9
4	16	5	25	4	16	5	25	2	4
3	9	3	9	3	9	4	16	5	25
4	16	5	25	4	16	4	16	3	9
5	25	5	25	5	25	3	9	4	16
4	16	5	25	4	16	3	9	4	16
4	16	5	25	4	16	4	16	5	25
4	16	5	25	4	16	5	25	4	16
4	16	4	16	5	25	4	16	3	9
3	9	4	16	4	16	1	1	4	16
76	292	93	425	84	350	79	321	72	274

Berdasarkan jumlah kuadrat dan skor total item, maka dapat dicari Varians item (S_i^2) masing-masing pada setiap item.

$$S_t^2 = \frac{\sum X_t^2 - \left[\frac{\sum X_t}{N} \right]^2}{N}$$

No. Item	$\sum X^2$	$\sum X$	$(\sum X)^2$	N	$(\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N$	$\sum X^2 - (\sum X)^2/N : N$
item 1	292	76	5776	21	275.0476	16.952381	0.807256
item 2	425	93	8649	21	411.8571	13.142857	0.62585
item 3	350	84	7056	21	336	14	0.666667
item 4	321	79	6241	21	297.1905	23.809524	1.133787
item 5	274	72	5184	21	246.8571	27.142857	1.292517
							4.526077

Atau menggunakan formula Varians sebagaimana dibawah ini.

$$S_1^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_1^2 = \left[\frac{292}{21} \right] - \left[\frac{76}{21} \right]^2$$

$$S_1^2 = [13,904761904762] - [3,6190476190476]^2$$

$$S_1^2 = [13,904761904762] - [13,097505668934]$$

$$S_1^2 = \mathbf{0,807256235828}$$

$$S_2^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_2^2 = \left[\frac{425}{21} \right] - \left[\frac{93}{21} \right]^2$$

$$S_2^2 = [20,2380955238095] - [4,4285714285714]^2$$

$$S_2^2 = [20,2380955238095] - [19,612244897959]$$

$$S_2^2 = \mathbf{0,625850340136}$$

$$S_3^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_3^2 = \left[\frac{350}{21} \right] - \left[\frac{84}{21} \right]^2$$

$$S_3^2 = [16,666666666667] - [4]^2$$

$$S_3^2 = [16,666666666667] - [16]$$

$$S_3^2 = \mathbf{0,6666666666667}$$

$$S_4^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_4^2 = \left[\frac{321}{21} \right] - \left[\frac{79}{21} \right]^2$$

$$S_4^2 = [15,285714285714] - [3,7619047619048]^2$$

$$S_4^2 = [15,285714285714] - [14,151927437642]$$

$$S_4^2 = \mathbf{1,133786848072}$$

$$S_5^2 = \left[\frac{\sum X_t^2}{N} \right] - \left[\frac{\sum X_t}{N} \right]^2$$

$$S_5^2 = \left[\frac{274}{21} \right] - \left[\frac{72}{21} \right]^2$$

$$S_5^2 = [13,047619047619] - [3,4285714285714]^2$$

$$S_5^2 = [13,047619047619] - [11,755102040816]$$

$$S_5^2 = \mathbf{1,292517006803}$$

Menjumlahkan masing-masing varians pada setiap item

$$\sum S_i^2 = 0,807256235828 + \mathbf{0,625850340136} + \mathbf{0,66666666666667}$$

$$+ \mathbf{1,133786848072} + \mathbf{1,292517006803} = \mathbf{4,526077097506}$$

Memasukkan kedalam formulasi Alpha Cronbach;

$$r_{11} = \left[\frac{k}{k-1} \right] \left[1 - \frac{\sum S_i^2}{S_t^2} \right]$$

$$r_{11} = \left[\frac{5}{5-1} \right] \left[1 - \frac{4,526077097506}{6,56235827666} \right]$$

$$r_{11} = \left[\frac{5}{4} \right] [1 - 0,6897028334469]$$

$$r_{11} = 1,25 \times [0,310297166553]$$

$$r_{11} = \mathbf{0,3878714581913}$$

$$r_{11} = \mathbf{0,388}$$

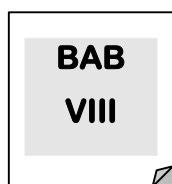
Hasil program SPSS

Reliability Statistics	
Cronbach's Alpha	N of Items
.388	5

Kriteria interpretasi hasil pengujiannya :

< 0,5	= Buruk
0,5 – 0,6	= Jelek
0,6 – 0,7	= Cukup
0,7 – 0,8	= Bagus
>0,8	= Bagus Sekali

Kesimpulannya: berdasarkan nilai uji Instrumen non tes yang ditemukan angka sebesar 0,388 berada pada interval interpretasi yaitu $< 0,5$ = Buruk. Dengan demikian dapat disimpulkan instrument non tes tersebut termasuk dalam kategori buruk.



APLIKASI PROGRAM DALAM PENGUJIAN INSTRUMEN PENILIAN DAN PENELITIAN

Ada beberapa alat yang dapat digunakan untuk melakukan pengujian instrument penilaian dan penelitian. Namun kesempatan ini penulis hanya menjelaskan tentang Program SPSS, Anates dan Rasch Model. Untuk mendapatkan penjelasan secara rinci aplikasi program dapat dilakukan dengan langkah –langkah sebagaimana dijelaskan dibawah ini.

A. Pengujian Instrumen Penelitian dengan Aplikasi Program SPSS

Pengujian dengan menggunakan program SPSS banyak dilakukan oleh para peneliti dan penilaian dalam dunia pendidikan.

Sebelum *questinare* benar-benar dibagikan kepada responden dengan sampel yang besar, hendaknya diuji coba kepada sampel yang lebih kecil. Hal bertujuan untuk memperbaiki questinare jika ternyata item pertanyaan yang disusun tidak dapat mengukur perilaku yang ingin diukur, atau tidak konsisten.

1. CONTOH DATA DARI TABULASI Uji Coba Lapangan

Misalnya kita memiliki data dari 11 responden berikut. Penelitian dilakukan untuk mengetahui **Pengaruh Kepuasan Upah dan Motivasi Terhadap Kinerja Guru di MAN 1 Kudus**. Untuk melaksanakan kegiatan penelitian tersebut peneliti membuat instrumen penelitian masing-masing variabel sebanyak 5 pertanyaan dengan nilai kuantitatif. Skala Likert digunakan dengan skala 5 (sangat tinggi/sangat setuju dengan pernyataan), 4, 3, 2, dan 1 (sangat rendah/sangat tidak setuju pernyataan). Skala tersebut menggambarkan aspek-aspek tertentu yang diobservasi dari kinerja, kepuasan upah, maupun memberi motivasi. Penelitian memodel untuk menjelaskan **Kinerja** dengan dua variabel penjelas, yaitu, **Kepuasan terhadap upah (UPAH)**, dan **Motivasi dari dalam (MOTIV)**.

Questionare yang dikirim ke responden misalnya diperoleh data berikut.

KINERJA

No.resp.	Q1	Q2	Q3	Q4	Q5	Skor KINERJA
1.	4	3	4	5	3	19
2.	3	3	3	4	5	18
3.	5	4	4	5	4	22
4.	3	2	2	3	2	12
5.	5	5	4	4	3	21
6.	4	3	3	3	4	17
7.	3	4	4	3	2	16

8.	4	3	5	5	4	21
9.	4	4	4	5	3	20
10.	3	4	5	5	3	20
11.	5	5	5	5	4	24

Kepuasan UPAH

No.resp.	Q1	Q2	Q3	Q4	Q5	Skor Upah
1.	5	4	5	5	4	23
2.	4	3	3	4	5	19
3.	4	4	4	4	3	19
4.	2	3	2	3	2	12
5.	4	5	3	4	3	19
6.	5	5	3	3	2	18
7.	4	4	3	3	2	16
8.	5	4	5	4	5	23
9.	5	5	5	5	4	24
10.	3	4	4	3	2	16
11.	5	5	5	4	4	23

Motivasi

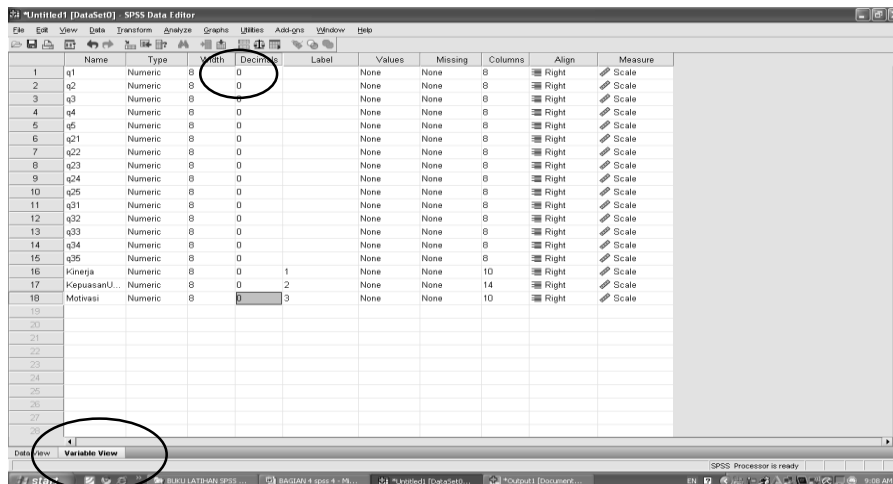
No.resp.	Q1	Q2	Q3	Q4	Q5	Skor MOTIVASI
1.	5	4	4	3	4	20
2.	4	3	3	4	5	19
3.	4	4	4	5	3	20
4.	2	3	3	4	2	14
5.	4	5	4	3	3	19
6.	5	5	4	4	2	20
7.	4	4	4	5	2	19
8.	5	4	5	5	5	24
9.	5	5	4	5	4	23
10.	3	4	4	4	2	17
11.	5	5	5	4	4	23

2. Pengolahan dengan Program SPSS

Sekarang bagaimana menggunakan data di atas dengan alat bantu SPSS.

- 1). Hidupkan program SPSS
- 2). Desain variabel dengan Klik variable (lihat tanda lingkaran pada gambar di bawah) pada data editor, beri nama item sesuai urutan q1, q2. q3. q4, q5, dst. Pada kolom decimal set nol desimal. (lihat chart).

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Kolom paling kiri isi dengan nama variabel (q1, q2, dst)

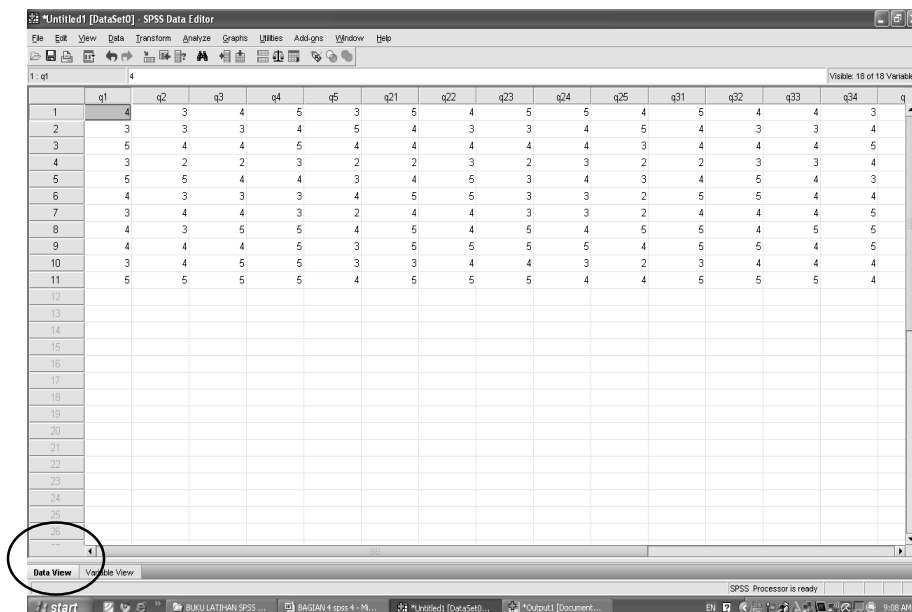
Type: Numeric – untuk variabel yang berisi angk-angka skala – klik string untuk variabel kategori (laki-wanita, besar-kecil, nama orang dst).

Width: set untuk banyaknya digit/lebar kolom yang diinginkan.

Decimal: set untuk berapa angka di belakang koma, SPSS akan langsung set 2 decimal,

3). Masukkan tabulasi data hasil penelitian ke dalam chart data SPSS.

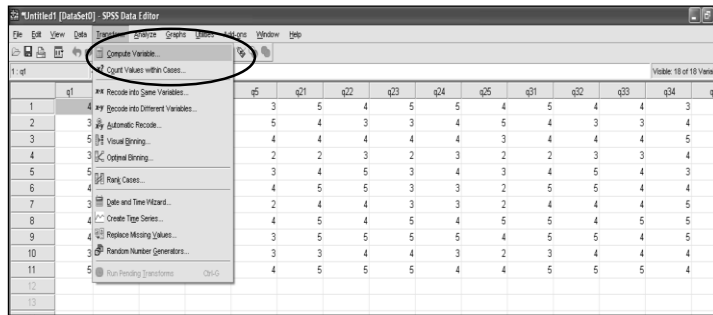
Di layar SPSS susunan data akan terlihat sebagai berikut.



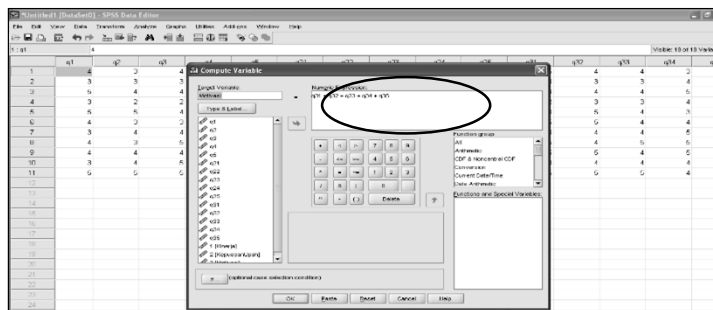
4). CARA MEMBUAT VARIABEL BARU dari variabel yang sudah ada.

Jika total skor untuk variabel KINERJA, KEPUASANUPAH, dan MOTIVASI belum dijumlah, SPSS dapat melakukan sistem penjumlahan dengan dengan prosedur sebagai berikut berikut.

1. Pada menu Transform di Klik
2. Kemudian Klik: Compute



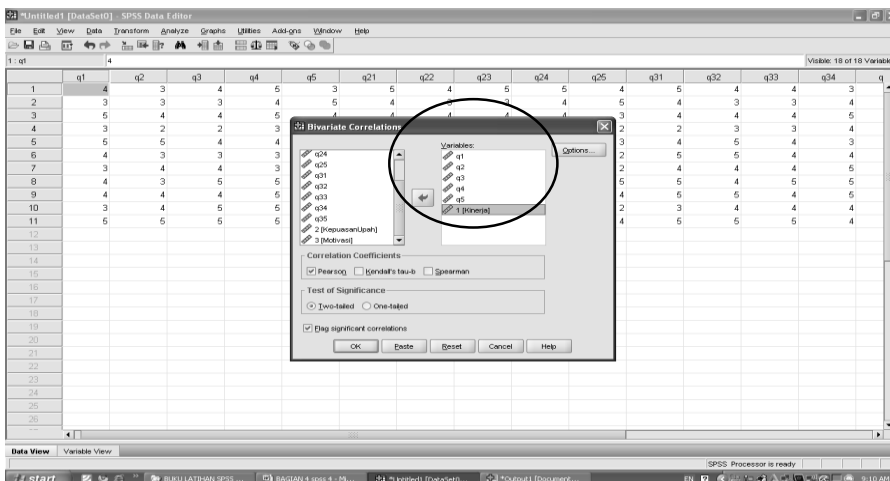
3. Masukkan nama variabel target: KINERJA
4. Kemudian jumlahkan nilai $q1 + q2 + q3 + q4 + q5$ dalam box **numeric expression** (lihat contoh dalam box berikut);
5. Kemudian klik OK (lihat panah).



Pada data editor akan muncul variabel baru yang merupakan skor variabel KINERJA yang berasal dari jumlah nilai-nilai dari pertanyaan 1 sampai 5 (jumlah $q1$ ---- s.d. $q5$).

2.1. Uji Validitas

- a) Klik Analyze, arahkan kursor ke Bivariate Correlation.
- b) Masukkan variabel $q1$, $q2$, $q3$, $q4$, $q5$ dan kinerja ke dalam kotak variabels dengan cara mengblok dan mengklik tanda panah.



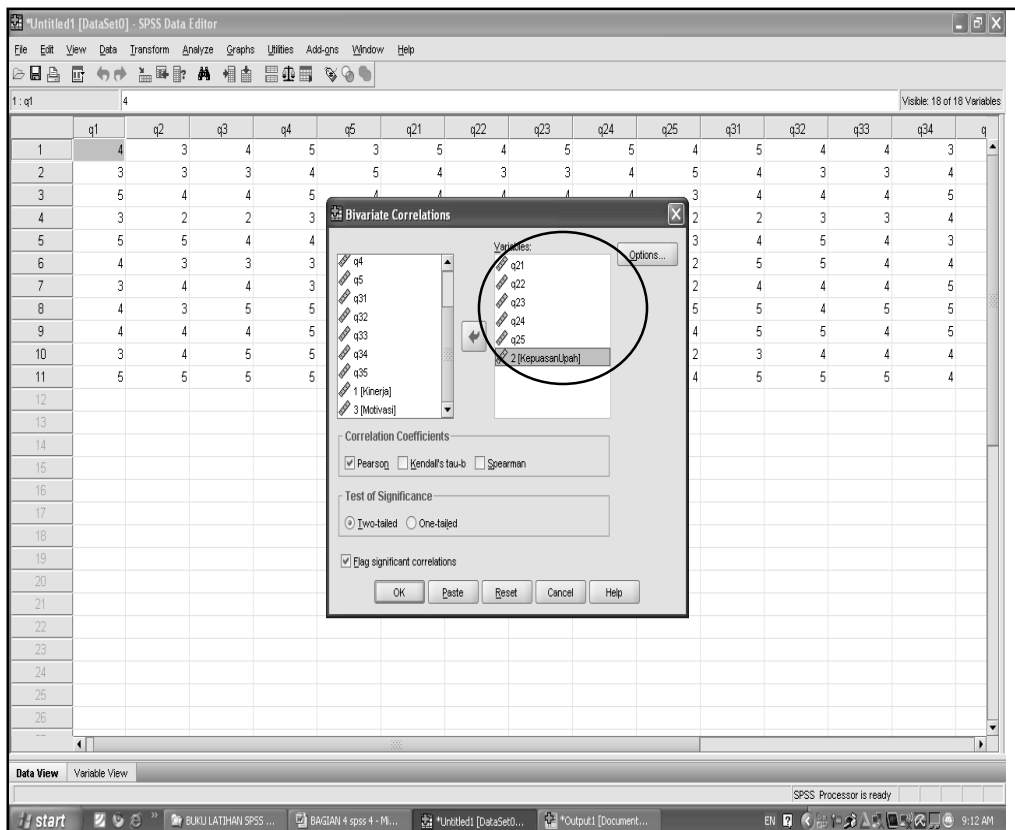
Correlations

		q1	q2	q3	Q4	q5	1
q1	Pearson Correlation	1	.603*	.371	.435	.308	.739**
	Sig. (2-tailed)		.049	.262	.181	.357	.009
	N	11	11	11	11	11	11
q2	Pearson Correlation	.603*	1	.646*	.370	.053	.740**
	Sig. (2-tailed)	.049		.032	.263	.877	.009
	N	11	11	11	11	11	11
q3	Pearson Correlation	.371	.646*	1	.735*	.156	.813**
	Sig. (2-tailed)	.262	.032		.010	.646	.002
	N	11	11	11	11	11	11
q4	Pearson Correlation	.435	.370	.735*	1	.348	.802**
	Sig. (2-tailed)	.181	.263	.010		.294	.003
	N	11	11	11	11	11	11
q5	Pearson Correlation	.308	.053	.156	.348	1	.517
	Sig. (2-tailed)	.357	.877	.646	.294		.103
	N	11	11	11	11	11	11
1	Pearson Correlation	.739**	.740**	.813**	.802**	.517	1
	Sig. (2-tailed)	.009	.009	.002	.003	.103	
	N	11	11	11	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

- c) Masukkan variabel q21, q22, q23, q24, q25 dan kepuasan upah ke dalam kotak variabels dengan cara mengblok dan mengklik tanda panah.



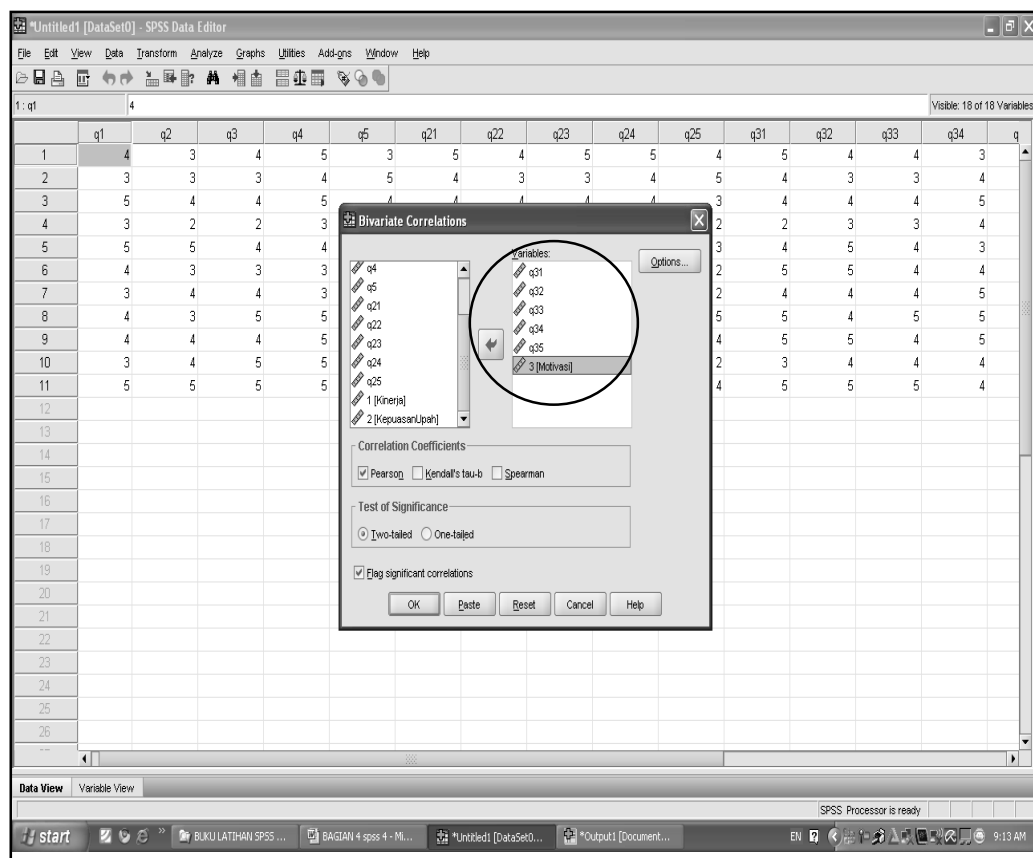
Correlations

		q21	q22	q23	q24	q25	2
q21	Pearson Correlation	1	.629*	.695*	.592	.552	.883**
	Sig. (2-tailed)		.038	.018	.055	.078	.000
	N	11	11	11	11	11	11
q22	Pearson Correlation	.629*	1	.415	.242	-.061	.514
	Sig. (2-tailed)	.038		.204	.474	.859	.106
	N	11	11	11	11	11	11
q23	Pearson Correlation	.695*	.415	1	.696*	.587	.880**
	Sig. (2-tailed)	.018	.204		.017	.057	.000
	N	11	11	11	11	11	11
q24	Pearson Correlation	.592	.242	.696*	1	.732*	.838**
	Sig. (2-tailed)	.055	.474	.017		.010	.001
	N	11	11	11	11	11	11
q25	Pearson Correlation	.552	-.061	.587	.732*	1	.767**
	Sig. (2-tailed)	.078	.859	.057	.010		.006
	N	11	11	11	11	11	11
2	Pearson Correlation	.883**	.514	.880**	.838**	.767**	1
	Sig. (2-tailed)	.000	.106	.000	.001	.006	
	N	11	11	11	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

**. Correlation is significant at the 0.01 level (2-tailed).

- d) Masukkan variabel q31, q32, q33, q34, q35 dan motivasi ke dalam kotak variabels dengan cara mengblok dan mengklik tanda panah.



Correlations

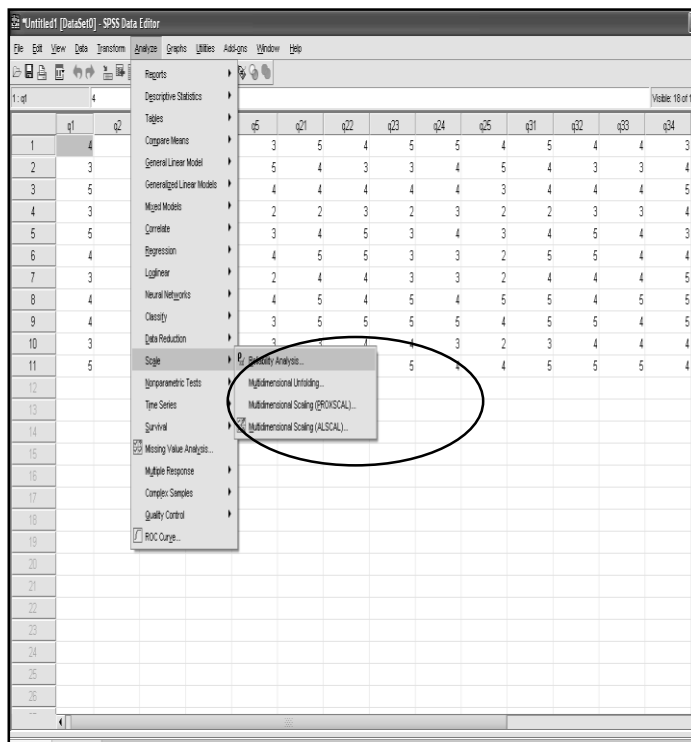
		q31	q32	q33	q34	q35	3
q31	Pearson Correlation	1	.629*	.644*	.086	.552	.904**
	Sig. (2-tailed)		.038	.032	.801	.078	.000
	N	11	11	11	11	11	11
q32	Pearson Correlation	.629*	1	.632*	-.065	-.061	.576
	Sig. (2-tailed)	.038		.037	.851	.859	.063
	N	11	11	11	11	11	11
q33	Pearson Correlation	.644*	.632*	1	.211	.266	.775**
	Sig. (2-tailed)	.032	.037		.534	.430	.005
	N	11	11	11	11	11	11
q34	Pearson Correlation	.086	-.065	.211	1	.051	.343
	Sig. (2-tailed)	.801	.851	.534		.882	.301
	N	11	11	11	11	11	11
q35	Pearson Correlation	.552	-.061	.266	.051	1	.663*
	Sig. (2-tailed)	.078	.859	.430	.882		.026
	N	11	11	11	11	11	11
3	Pearson Correlation	.904**	.576	.775**	.343	.663*	1
	Sig. (2-tailed)	.000	.063	.005	.301	.026	
	N	11	11	11	11	11	11

*. Correlation is significant at the 0.05 level (2-tailed).

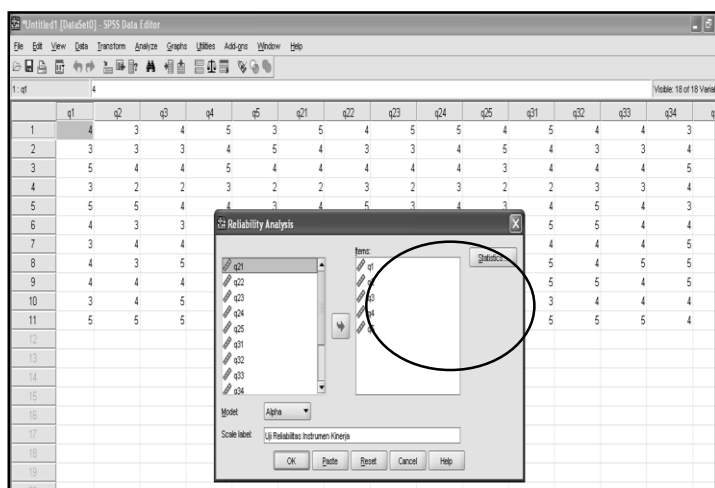
**. Correlation is significant at the 0.01 level (2-tailed).

2.2. Uji Reliabilitas

a) Klik Analyze, arahkan kursor ke scale dan reliability analyze.



b) Masukkan variable q1, q2, q3, q4, q5 ke dalam kotak variabels dengan cara mengblok dan mengklik tanda panah.



Reliability

Notes

Output Created		11-Jun-2011 09:15:25
Comments		
Input	Active Dataset	DataSet0
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	11
Missing Value Handling	Matrix Input	
	Definition of Missing	User-defined missing values are treated as missing.
Syntax	Cases Used	Statistics are based on all cases with valid data for all variables in the procedure.
		RELIABILITY
		/VARIABLES=q1 q2 q3 q4 q5
Resources		/SCALE('Uji Reliabilitas Instrumen Kinerja') ALL
		/MODEL=ALPHA.
	Processor Time	00:00:00.015
	Elapsed Time	00:00:00.014

[DataSet0]

Scale: Uji Reliabilitas Instrumen Kinerja

Case Processing Summary

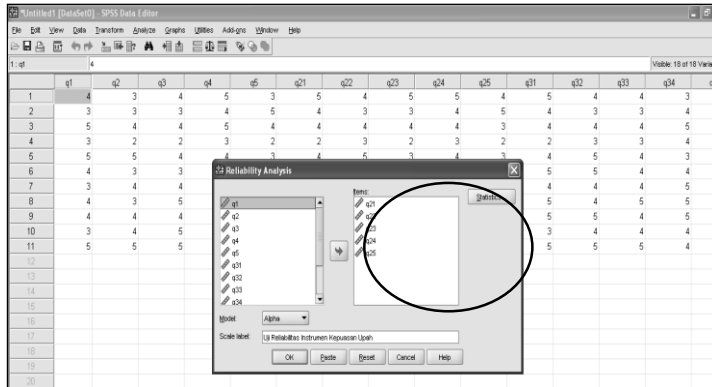
		N	%
Cases	Valid	11	100.0
	Excluded ^a	0	.0
	Total	11	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.770	5

- c) Masukkan variable q21, q22, q23, q24, q25 dan kepuasan upah ke dalam kotak variabels dengan cara mengblok dan mengklik tanda pandah.



Reliability

Notes

Output Created	11-Jun-2011 09:17:50
Comments	
Input	Active Dataset DataSet0 Filter <none> Weight <none> Split File <none> N of Rows in Working Data File 11 Matrix Input
Missing Value Handling	Definition of Missing User-defined missing values are treated as missing. Cases Used Statistics are based on all cases with valid data for all variables in the procedure.
Syntax	RELIABILITY /VARIABLES=q21 q22 q23 q24 q25 /SCALE('Uji Reliabilitas Instrumen Kepuasan Upah') ALL /MODEL=ALPHA.
Resources	Processor Time 00:00:00.016 Elapsed Time 00:00:00.017

[DataSet0]

Scale: Uji Reliabilitas Instrumen Kepuasan Upah
Case Processing Summary

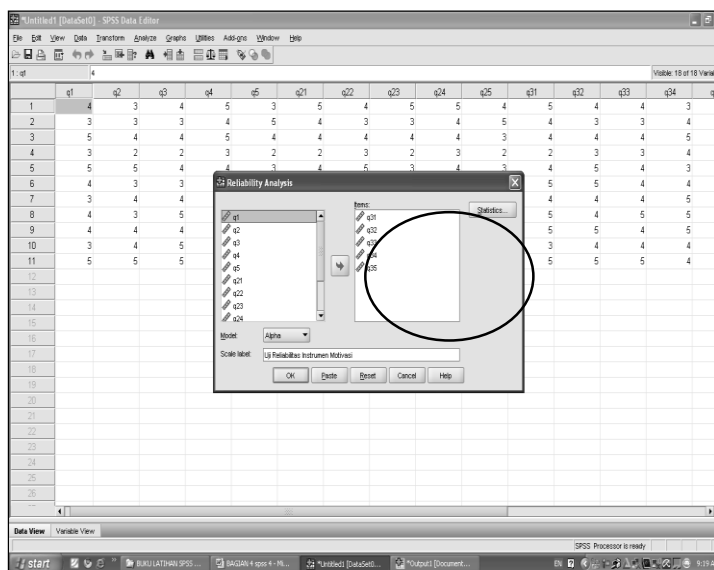
	N	%
Cases		
Valid	11	100.0
Excluded ^a	0	.0
Total	11	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.833	5

- d) Masukkan variable q21, q22, q23, q24, q25 dan motivasi ke dalam kotak variabels dengan cara mengblok dan mengklik tanda pandah.



Reliability

Notes

Output Created		11-Jun-2011 09:19:23
Comments		
Input	Active Dataset	DataSet0
	Filter	<none>
	Weight	<none>
	Split File	<none>
	N of Rows in Working Data File	11
	Matrix Input	
Missing Value Handling	Definition of Missing	User-defined missing values are treated as missing.
	Cases Used	Statistics are based on all cases with valid data for all variables in the procedure.
Syntax		RELIABILITY /VARIABLES=q31 q32 q33 q34 q35 /SCALE('Uji Reliabilitas Instrumen Motivasi') ALL /MODEL=ALPHA.
Resources	Processor Time	00:00:00.016
	Elapsed Time	00:00:00.014

[DataSet0]

Scale: Uji Reliabilitas Instrumen Motivasi

Case Processing Summary

	N	%
Cases Valid	11	100.0
Excluded ^a	0	.0
Total	11	100.0

a. Listwise deletion based on all variables in the procedure.

Reliability Statistics

Cronbach's Alpha	N of Items
.651	5

2.3. Interpretasi

2.3.1. Uji Validitas dari Var. Kinerja, Kepuasan Upah dan Motivasi

Tabel 41. Interpretasi Uji Validitas

No. Item	Koefisien Korelasi	Angka Signifikansi	Tanda Bintang	Keterangan
q1	0,739	0,009	**	Valid
q2	0,740	0,009	**	Valid
q3	0,813	0,002	**	Valid
q4	0,802	0,003	**	Valid
q5	0,517	0,103	-	Tidak valid
q21	0,883	0,000	**	Valid
q22	0,514	0,106	-	Tidak Valid
q23	0,880	0,000	**	Valid
q24	0,838	0,001	**	Valid
q25	0,767	0,006	**	Valid
q31	0,904	0,000	**	Valid
q32	0,576	0,063	-	Tidak Valid
q33	0,775	0,005	**	Valid
q34	0,343	0,301	-	Tidak Valid
q35	0,663	0,026	**	Valid

Kesimpulannya dari 15 soal terlihat yang tidak valid yaitu q5, q22, q32, dan q34

2.3.2. Uji Reliabilitas

Tabel 42. Interpretasi Reliabilitas

No. Variabel	Koefisien Cronbach Alpha	Standar Kriteria	Keterangan
Kinerja	.770	0,60>	Reliabel
Kepuasan Upah	.833	0,60>	Reliabel
Motivasi	.651	0,60>	Reliabel

B. Pengujian Instrumen Penelitian dengan Aplikasi Program Anates

1. Analisis Butir Soal dengan menggunakan

Software ANATES V4

1). Tentang Anates

Anates merupakan sebuah program aplikasi computer bertujuan untuk menganalisis butir butir soal. Program ini sangat bermanfaat khususnya bagi para guru, peneliti atau dosen pada umumnya dan pemerhati evaluasi pendidikan.

2). Fasilitas yang ada dalam program ANATES

- Penyebaran data meliputi : memasukan skor data hasil tes, dan Membobot skor data sesuai yang dibutuhkan.

- b. Pengolahan data meliputi: Reliabilitas, kelompok unggul dan asor, daya pembeda, tingkat kesukaran soal, korelasi skor butir soal dengan skor total, dan kualitas pengecoh.

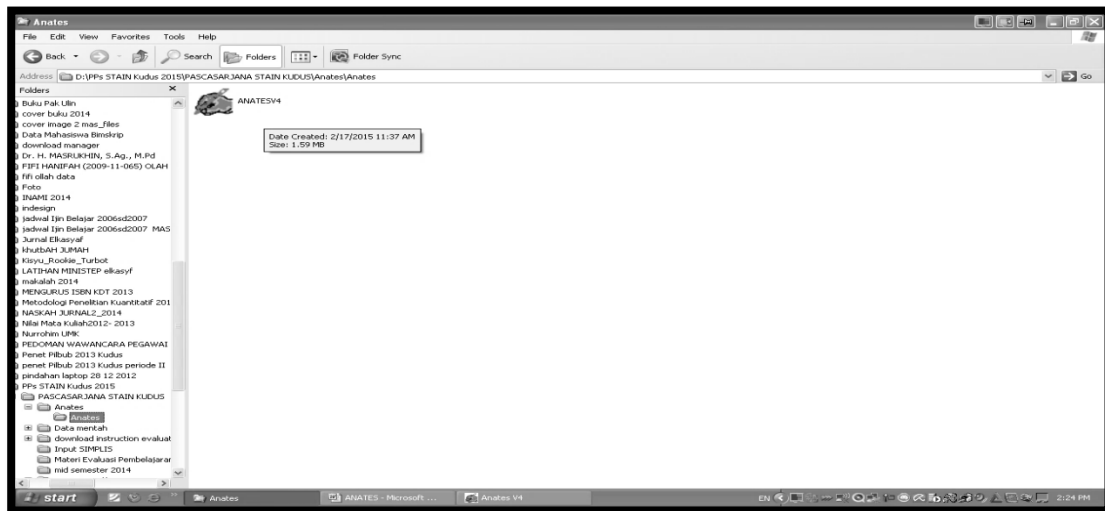
3). Menginstal Anates

Untuk menggunakan program ini, sebelumnya program perlu diinstallkan pada computer. Cara menginstallnya adalah :

- a) Buka dan kopikan file master ANATES (baik melalui CD maupun Flashdiks atau file dalam Hardisk), file terlihat pada gambar dibawah ini.

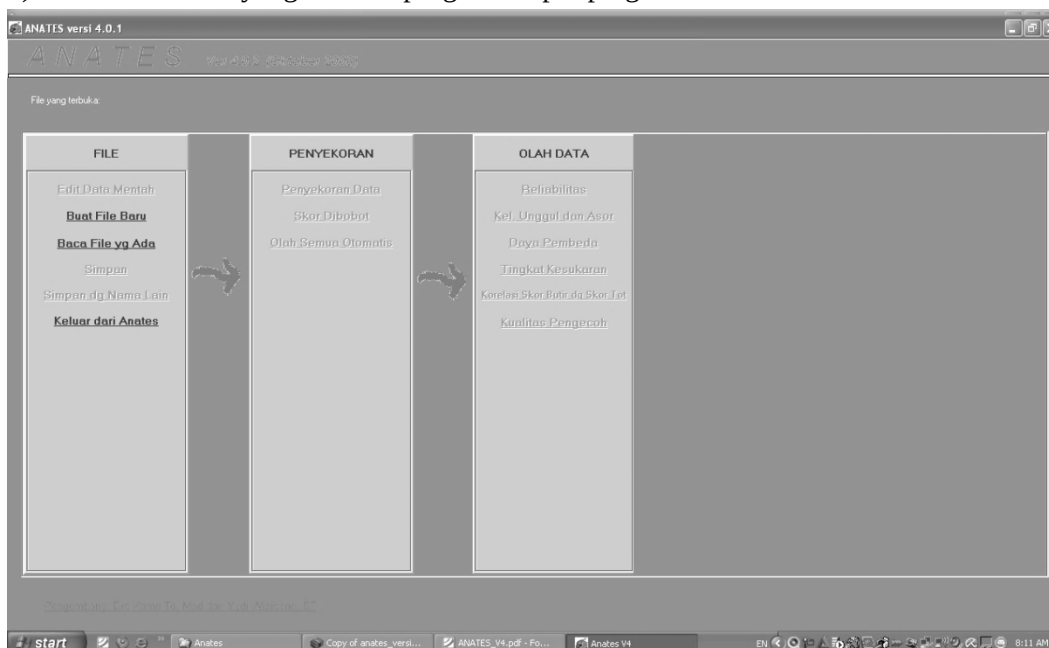


Anatesv4.exe



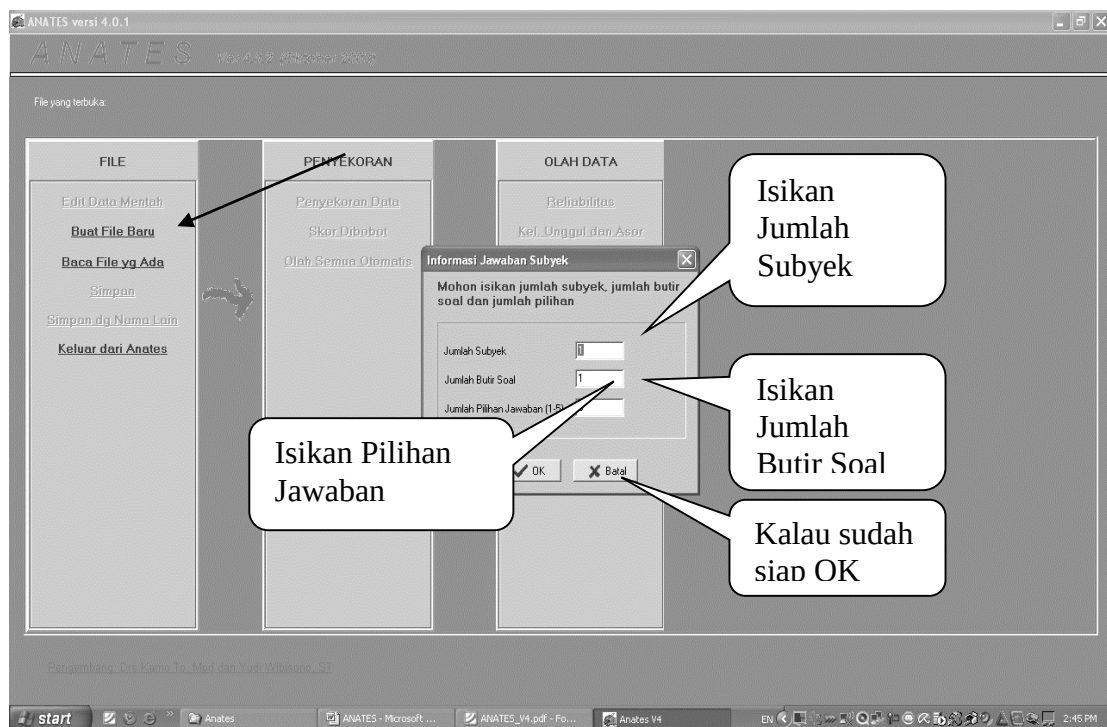
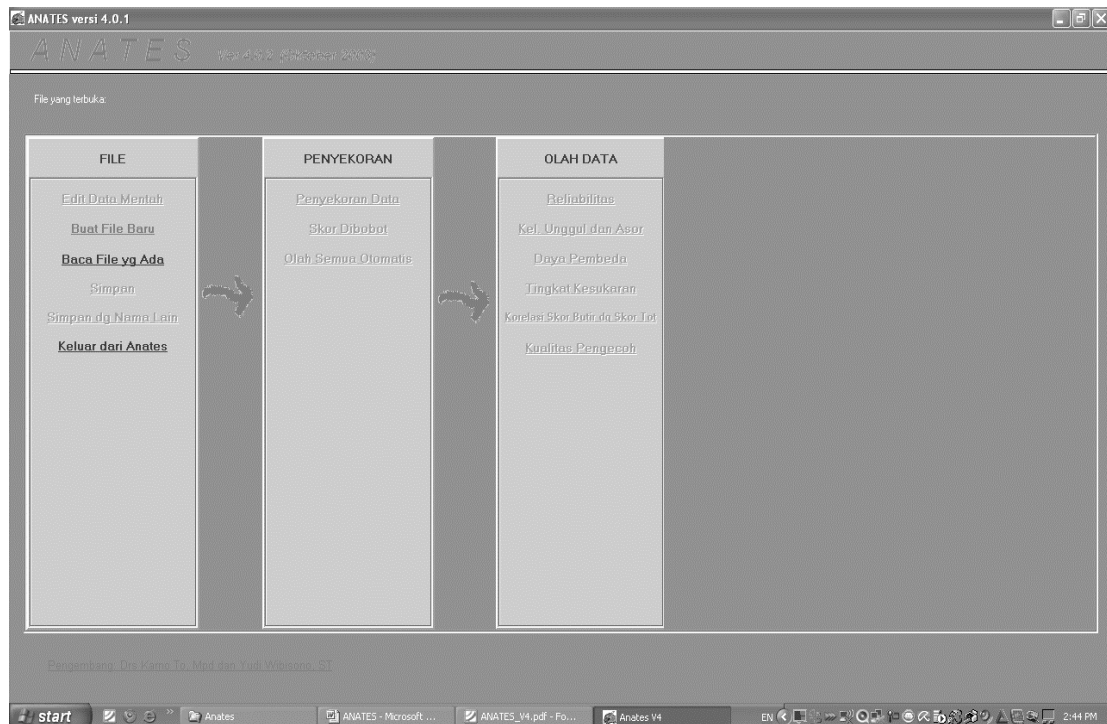
4). Memulai ANATES

- a) Klik dua kali pada gambar tersebut.
b) Jika muncul tanyang berikut, program siap dipergunakan.



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

- c) Memasukkan data, klik Buat File Baru, isikan data seperti berikut :
- d) Jika sudah terisi Jumlah Subjek, Jumlah butir soal, dan Jumlah Pilihan Jawaban, Klik OK, muncul seperti berikut:



- e) Kalau sudah terisi Jumlah Subjek, Jumlah butir soal, dan jumlah pilihan Jawaban, klik OK, muncul seperti berikut:

The screenshot shows a software window titled 'Edit Data Mentah: Belum_ada_nama.ana'. It has a menu bar with 'Kembali Ke Menu Utama', 'Buat File Baru', 'Simpan', 'Baca File', and 'Cetak'. Below the menu, it displays 'Jumlah Subjek: 10', 'Jumlah Butir Soal: 10', and 'Jumlah Pilihan Jawaban: 5'. A tip says 'Gunakan tombol F2 untuk pindah antar kolom'. The main area is a table with columns 'No.Urut', 'Kode/Nama Subjek', and 10 numbered columns (1-10). The first row is labeled 'KUNCI->' and contains the text 'KUNCI->' in the first column. Callouts point to specific cells: 'Sel untuk menuliskan jawaban sbyek' points to column 10; 'Jika subjek tidak menjawab berikan tanda bintang (*)' points to the cell at row 1, column 1; 'Tulis kode atau nama subiek' points to the 'Kode/Nama Subjek' column; and 'Tulis kunci jawaban pada baris ini' points to the 'KUNCI->' row.

No.Urut	Kode/Nama Subjek	1	2	3	4	5	6	7	8	9	10
KUNCI->	KUNCI->										
1											
2											
3											
4											
5											
6											
7											
8											
9											
10											

Tulis kode atau nama subiek

Tulis kunci jawaban pada baris ini

Sel untuk menuliskan jawaban sbyek

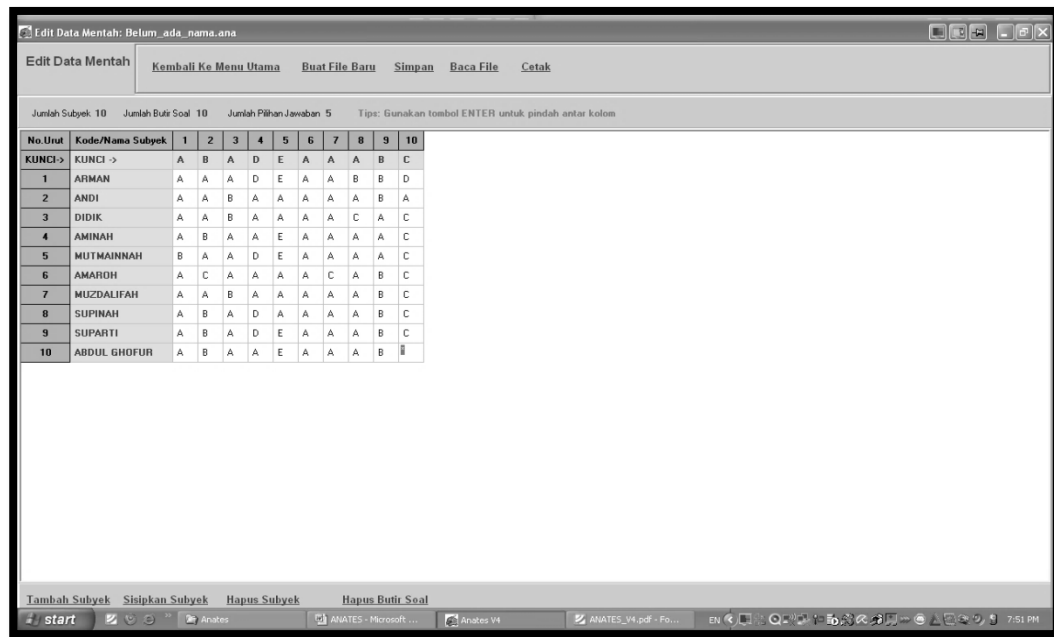
Jika subjek tidak menjawab berikan tanda bintang (*)

- f) Contoh pengisian data :

Masukkan KUNCI jawaban pada baris pertama

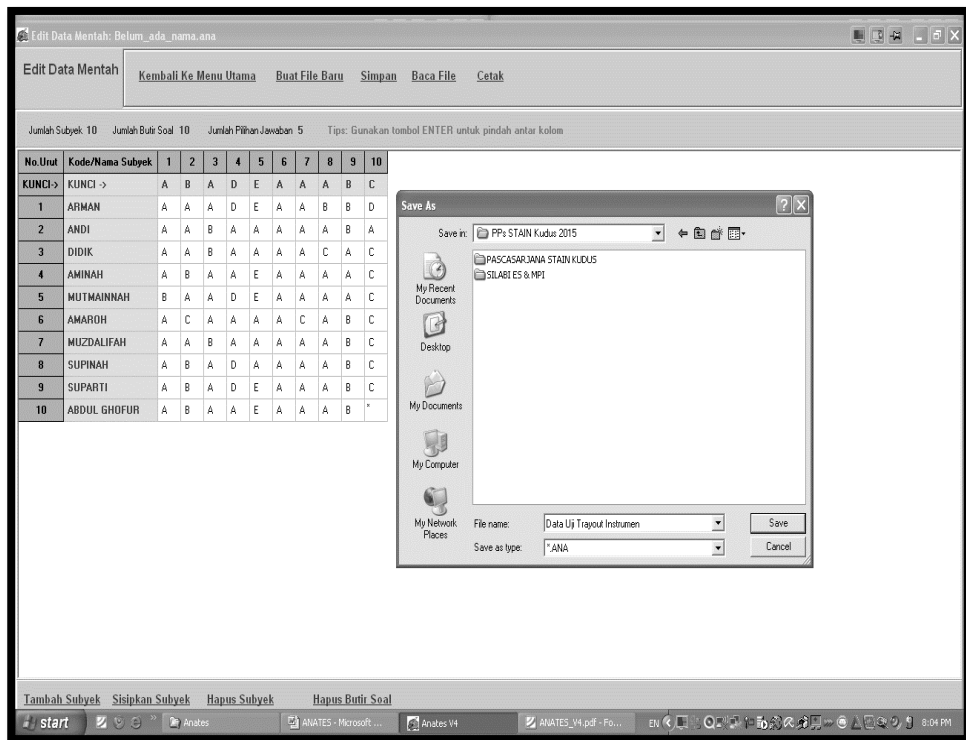
Masukkan nama siswa pada kolom nama/subyek

Masukkan jawaban pada sel/kolom yang berwarna putih, jika siswa tidak menjawab isikan tanda bintang (*)



g) Fungsi-fungsi pada Menu Bar

- **Kembali ke menu utama**
Jika diklik menu ini, tampilan akan ke tampilan pertama membuka program
- **Buat file Baru**
Untuk memasukkan data baru, klik menu Buat File Baru, kemudian isikan data-data seperti pada langkah di atas.
- **Simpan**
Data yang sudah dimasukkan simpan atau saving pada harddisk untuk pengamanan data
- **Baca File**
Menu ini untuk membuka file yang sudah ada pada data sebelumnya. Misalnya untuk mengeditnya.
- **Cetak**
Menu ini untuk mencetak lembar yang dibutuhkan, misalnya data keseluruhan atau hasil olahan data.
- **Untuk pengamanan**, data yang sudah diisikan disimpang (saving) dalam harddisk, dengan mengklik menu simpan yang ada pada menu bar di bagian atas.



- Jika data sudah diisikan pada kolom yang tersedia, maka data tersebut sudah siap untuk diolah sesuai keperluan dalam menganalisis butir soal.

5). Penyebaran

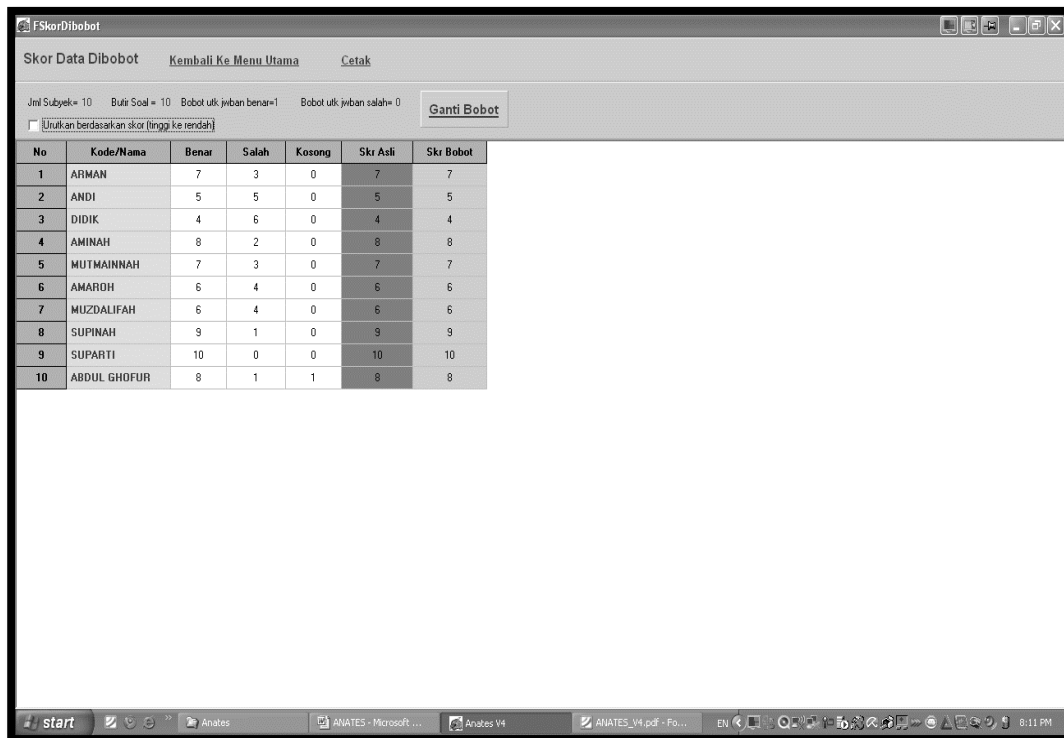
Menu penyebaran untuk melihat hasil jawaban siswa. Jawaban benar ditandai dengan angka 1, jawaban salah 0 (-), dan tidak menjawab dengan tanda bintang (*). Contoh :

The screenshot shows the 'FSkor' application window. The main window contains a table with columns for 'No.Urut', 'Kode>Nama Subyek', 'Skor', and 10 columns for answers (1-10). The table shows scores for 10 students, with 1 indicating a correct answer, 0 indicating an incorrect answer, and * indicating no answer.

No.Urut	Kode>Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10
1	ARMAN	7	1	-	1	1	1	1	1	-	1	-
2	ANDI	5	1	-	-	-	1	1	1	1	1	-
3	DIDIK	4	1	-	-	-	1	1	-	-	1	-
4	AMINAH	8	1	1	1	-	1	1	1	1	-	1
5	MUTMAINNAH	7	-	-	1	1	1	1	1	1	-	1
6	AMAROH	6	1	-	1	-	-	1	-	1	1	1
7	MUZDALIFAH	6	1	-	-	-	-	1	1	1	1	1
8	SUPINAH	9	1	1	1	1	-	1	1	1	1	1
9	SUPARTI	10	1	1	1	1	1	1	1	1	1	1
10	ABDUL GHOFUR	8	1	1	1	-	1	1	1	1	1	*

6). **Skor dibobot**

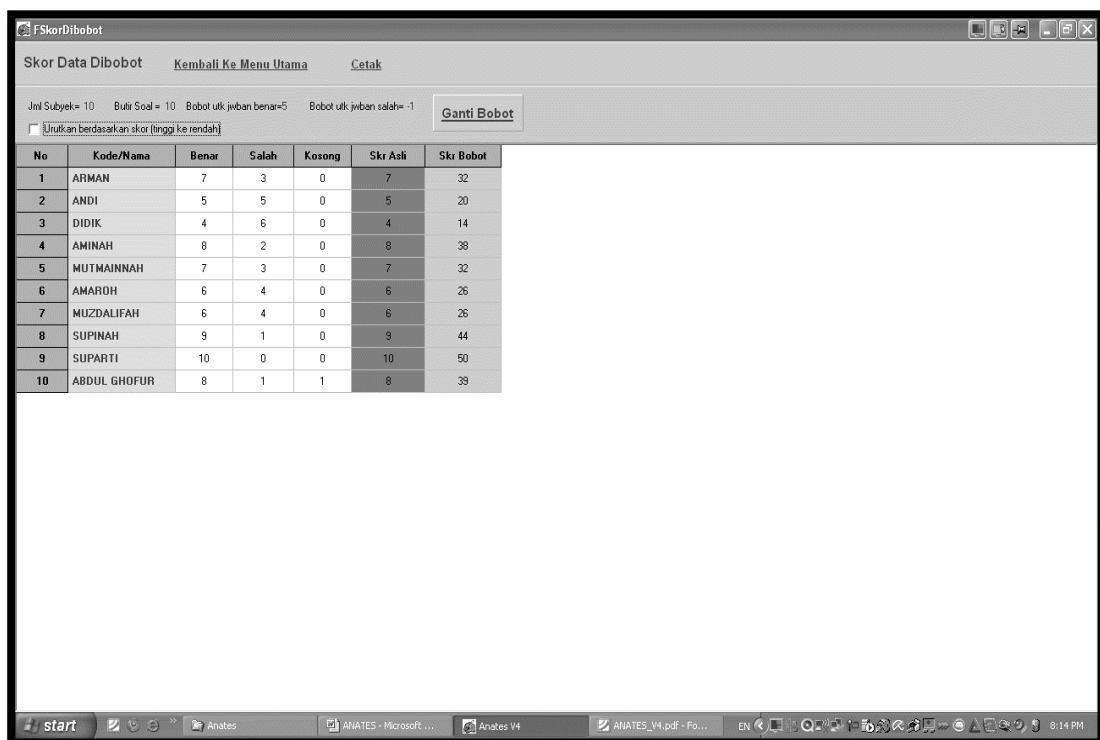
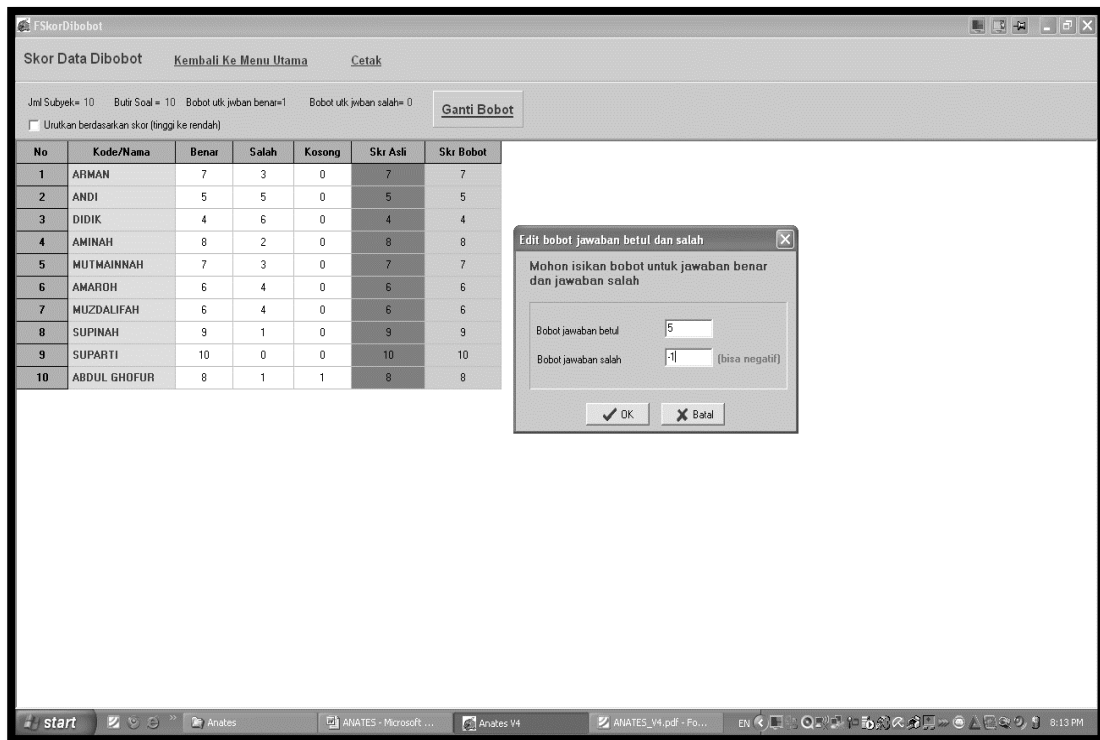
Menu ini untuk melihat jumlah hasil jawaban siswa. Jika akan mengganti bobot skor klik **Ganti Bobot** (bagian kanan atas). Bobot standar 1 jawaban benar = bernilai 1.



The screenshot shows a software window titled 'FSkorDibobot'. It has a menu bar with 'Skor Data Dibobot', 'Kembali Ke Menu Utama', and 'Cetak'. Below the menu bar, there are statistics: 'Jml Subjek= 10', 'Buat Soal = 10', 'Bobot utk jwban benar=1', and 'Bobot utk jwban salah= 0'. A 'Ganti Bobot' button is located to the right of these statistics. Below the statistics, there is a checkbox labeled 'Urutkan berdasarkan skor (tinggi ke rendah)'. The main area of the window contains a table with the following data:

No	Kode>Nama	Benar	Salah	Kosong	Skor Asli	Skor Bobot
1	AFMAN	7	3	0	7	7
2	ANDI	5	5	0	5	5
3	DIDIK	4	6	0	4	4
4	AMINAH	8	2	0	8	8
5	MUTMAINNAH	7	3	0	7	7
6	AMAROH	6	4	0	6	6
7	MUZDALIFAH	6	4	0	6	6
8	SUPINAH	9	1	0	9	9
9	SUPARTI	10	0	0	10	10
10	ABDUL GHOFUR	8	1	1	8	8

Misalnya data diatas bobotnya diubah menjadi 1 jawaban benar = bernilai 5, jawaban salah = bernilai -1, dan tidak menjawab = bernilai 0, maka data akan berubah seperti berikut :



Perhatikan kolom SKr bobot berbeda dengan tampilan di atasnya.

7). Reliabilitas

Untuk melihat hasil uji reliabilitas, klik Reliabilitas pada menu **Olah Data**.

Reliabilitas Tes

Kembali Ke Menu Utama Cetak

Rata2=7.00 Simpang Baku= 1.83 KorelasiXY= 0.63 Reliabilitas Tes = 0.77

No.Urut	Kode>Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	ARMAN	2	5	7
2	ANDI	2	3	5
3	DIDIK	1	2	3
4	AMINAH	3	4	7
5	MUTMAINNAH	3	3	6
6	AMAROH	2	3	5
7	MUZDALIFAH	2	3	5
8	SUPINAH	4	4	8
9	SUPARTI	4	5	9
10	ABDUL GHOFUR	3	5	8

start Anates ANATES - Microsoft ... Anates V4 ANATES_V4.pdf - Fo... EN 8:17 PM

Hasil olahan data yang diperoleh pada menu ini adalah :

Rata-rata nilai = 7,00
 Simpangan Baku = 1, 83
 Korelasi XY = 0,63
 Reliabilitas Tes = 0,77

8). Kelompok Unggul dan Asor

Untuk keperluan analisis butir soal yang dianalisis adalah 27 % dari kelompok unggul dan 27% dari kelompok asor.

Kelompok Unggul dan Asor

Kel. Unggul & Asor

Kembali Ke Menu Utama

Cetak

Kelompok Unggul

No. Urut	Kode/Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10
1	SUPARTI	10	1	1	1	1	1	1	1	1	1	1
2	SUPINAH	9	1	1	1	1	-	1	1	1	1	1
3	AMINAH	8	1	1	1	-	1	1	1	1	-	1
Jml Jwb Benar			3	3	3	2	2	3	3	3	2	3

Kelompok Asor

No. Urut	Kode/Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10
1	MUZDALIFAH	6	1	-	-	-	-	1	1	1	1	1
2	ANDI	5	1	-	-	-	-	1	1	1	1	-
3	DIDIK	4	1	-	-	-	-	1	1	-	-	1
Jml Jwb Benar			3	0	0	0	0	3	3	2	2	2

start

Anates

ANATES - Microsoft ...

Anates V4

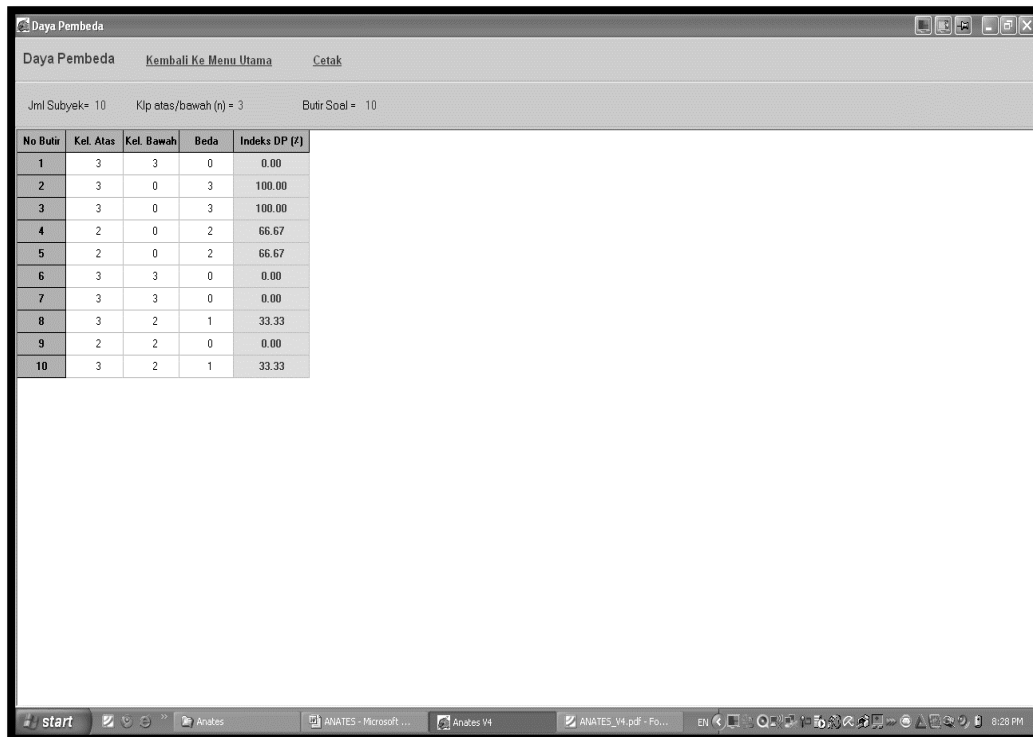
ANATES_V4.pdf - Fo...

EN

8:21 PM

9). Daya Pembeda

Daya pembeda merupakan perbandingan jawaban siswa kelompok unggul dan kelompok asor. Soal yang dianggap baik antara kelompok unggul dan kelompok asor memiliki perbedaan yang signifikan. Semakin besar perbedaannya dianggap semakin baik dan sebaliknya. Jika perbedaannya negative soal dianggap menyesatkan.



The screenshot shows the 'Daya Pembeda' (Discrimination Power) software interface. It displays a table with 5 columns: 'No Butir' (Item Number), 'Kel. Atas' (Upper Group), 'Kel. Bawah' (Lower Group), 'Beda' (Difference), and 'Indeks DP (%)' (Discrimination Index %). The table contains 10 rows of data for items 1 through 10. The software also shows 'Jml Subyek = 10', 'Kip atas/bawah (n) = 3', and 'Butir Soal = 10'.

No Butir	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	3	3	0	0.00
2	3	0	3	100.00
3	3	0	3	100.00
4	2	0	2	66.67
5	2	0	2	66.67
6	3	3	0	0.00
7	3	3	0	0.00
8	3	2	1	33.33
9	2	2	0	0.00
10	3	2	1	33.33

10).Tingkat Kesukaran

Tingkat kesukaran merupakan perbandingan antara jumlah jawaban benar dari setiap nomor soal dengan jumlah subyek

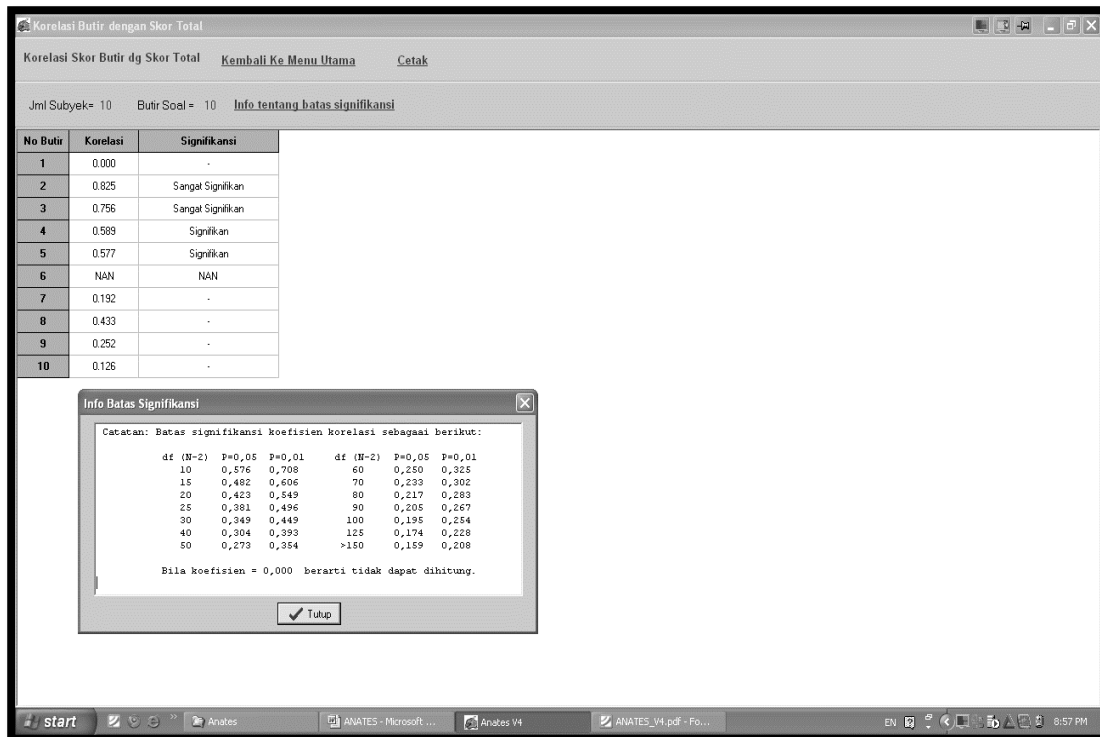


The screenshot shows the 'Tingkat Kesukaran' (Difficulty Level) software interface. It displays a table with 4 columns: 'No Butir' (Item Number), 'Jml Betul' (Number of Correct Answers), 'Tkt. Kesukaran (%)' (Difficulty Level %), and 'Tafsiran' (Interpretation). The table contains 10 rows of data for items 1 through 10. The software also shows 'Jml Subyek = 10' and 'Butir Soal = 10'.

No Butir	Jml Betul	Tkt. Kesukaran (%)	Tafsiran
1	9	90.00	Sangat Mudah
2	4	40.00	Sedang
3	7	70.00	Sedang
4	4	40.00	Sedang
5	5	50.00	Sedang
6	10	100.00	Sangat Mudah
7	9	90.00	Sangat Mudah
8	8	80.00	Mudah
9	7	70.00	Sedang
10	7	70.00	Sedang

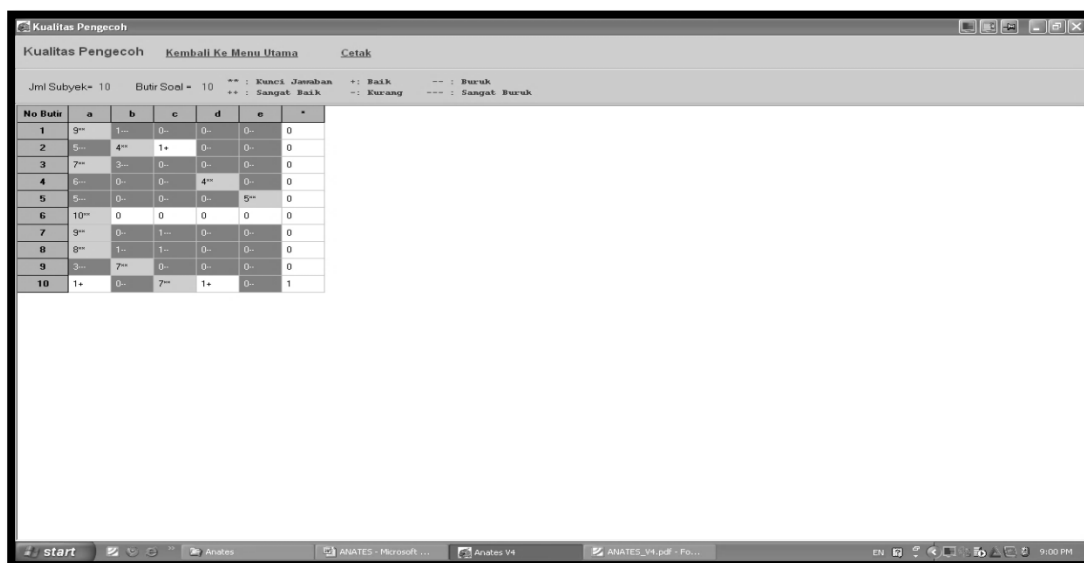
11). Korelasi Butir dengan Skor Total

Butir soal yang memiliki korelasi tinggi dianggap sebagai soal yang lebih baik dibandingkan dengan butir soal yang nilai korelasinya rendah. Dengan demikian soal yang memiliki korelasi tinggi dianggap sebagai signifikan untuk digunakan pada tes berikutnya, dan sebaliknya.



12). Kualitas Pengecoh

Kualitas pengecoh merupakan hasil analisis untuk melihat alternative (option atau pilihan) jawaban yang baik untuk digunakan kembali dan alternative jawaban yang buruk agar diubah atau diganti.



13).Olah Semua Otomatis

Langkah-langkah 5 hingga 12 ditampilkan setiap sesuai kebutuhan pengguna ANATES. Jika ingin melihat hasil pengolahan secara menyeluruh, klik Olah Semua Otomatis.

Tampilan Lainnya akan seperti berikut:

SKOR DATA DIBOBOT

=====

Jumlah Subyek = 10

Jumlah butir = 10

Bobot jwb benar = 5

Bobot jwb salah = -1

Keterangan: data terurut berdasarkan skor (tinggi ke rendah)

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No	Kode>Nama	Benar	Salah	Kosong	Skr Asli	Skr Bobot
1	SUPARTI	10	0	0	10	50
2	SUPINAH	9	1	0	9	44
3	AMINAH	8	2	0	8	38
4	ABDUL GHOFUR	8	1	1	8	39
5	ARMAN	7	3	0	7	32
6	MUTMAINNAH	7	3	0	7	32
7	AMAROH	6	4	0	6	26
8	MUZDALIFAH	6	4	0	6	26
9	ANDI	5	5	0	5	20
10	DIDIK	4	6	0	4	14

RELIABILITAS TES

=====

Rata2= 7.00

Simpang Baku= 1.83

KorelasiXY= 0.63

Reliabilitas Tes= 0.77

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No.Urut	Kode>Nama Subyek	Skor Ganjil	Skor Genap	Skor Total
1	ARMAN	2	5	7
2	ANDI	2	3	5
3	DIDIK	1	2	3
4	AMINAH	3	4	7
5	MUTMAINNAH	3	3	6
6	AMAROH	2	3	5
7	MUZDALIFAH	2	3	5
8	SUPINAH	4	4	8
9	SUPARTI	4	5	9
10	ABDUL GHOFUR	3	5	8

Kel Unggul & Asor

=====

Kelompok Unggul

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No.Urut	Kode/Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10
1	SUPARTI	10	1	1	1	1	1	1	1	1	1	1
2	SUPINAH	9	1	1	1	1	-	1	1	1	1	1
3	AMINAH	8	1	1	1	-	1	1	1	1	-	1
Jml Jwb Benar			3	3	3	2	2	3	3	3	2	3

Kelompok Asor

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No.Urut	Kode/Nama Subyek	Skor	1	2	3	4	5	6	7	8	9	10
1	MUZDALIFAH	6	1	-	-	-	-	1	1	1	1	1
2	ANDI	5	1	-	-	-	-	1	1	1	1	-
3	DIDIK	4	1	-	-	-	-	1	1	-	-	1
Jml Jwb Benar			3	0	0	0	0	3	3	2	2	2

DAYA PEMBEDA

=====

Jumlah Subyek= 10

Klp atas/bawah(n)= 3

Butir Soal= 10

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No Butir	Kel. Atas	Kel. Bawah	Beda	Indeks DP (%)
1	3	3	0	0.00
2	3	0	3	100.00
3	3	0	3	100.00
4	2	0	2	66.67
5	2	0	2	66.67
6	3	3	0	0.00
7	3	3	0	0.00
8	3	2	1	33.33
9	2	2	0	0.00
10	3	2	1	33.33

TINGKAT KESUKARAN

=====

Jumlah Subyek= 10

Butir Soal= 10

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No Butir	Jml Betul	Tkt. Kesukaran(%)	Tafsiran
1	9	90.00	Sangat Mudah

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

2	4	40.00	Sedang
3	7	70.00	Sedang
4	4	40.00	Sedang
5	5	50.00	Sedang
6	10	100.00	Sangat Mudah
7	9	90.00	Sangat Mudah
8	8	80.00	Mudah
9	7	70.00	Sedang
10	7	70.00	Sedang

KORELASI SKOR BUTIR DG SKOR TOTAL

=====

Jumlah Subyek= 10

Butir Soal= 10

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No Butir	Korelasi	Signifikansi
1	0.000	-
2	0.825	Sangat Signifikan
3	0.756	Sangat Signifikan
4	0.589	Signifikan
5	0.577	Signifikan
6	NAN	NAN
7	0.192	-
8	0.433	-
9	0.252	-
10	0.126	-

Catatan: Batas signifikansi koefisien korelasi sebagaai berikut:

df (N-2)	P=0,05	P=0,01	df (N-2)	P=0,05	P=0,01
10	0,576	0,708	60	0,250	0,325
15	0,482	0,606	70	0,233	0,302
20	0,423	0,549	80	0,217	0,283
25	0,381	0,496	90	0,205	0,267
30	0,349	0,449	100	0,195	0,254
40	0,304	0,393	125	0,174	0,228
50	0,273	0,354	>150	0,159	0,208

Bila koefisien = 0,000 berarti tidak dapat dihitung.

KUALITAS PENGECHOH

=====

Jumlah Subyek= 10

Butir Soal= 10

Nama berkas: D:\PPS STAIN KUDUS 2015\DATA UJI TRAYOUT INSTRUMEN.ANA

No Butir	a	b	c	d	e *
1	9**	1---	0--	0--	0-- 0
2	5---	4**	1+	0--	0-- 0
3	7**	3---	0--	0--	0-- 0

4	6---	0--	0--	4**	0--	0
5	5---	0--	0--	0--	5**	0
6	10**	0	0	0	0	0
7	9**	0--	1---	0--	0--	0
8	8**	1--	1--	0--	0--	0
9	3---	7**	0--	0--	0--	0
10	1+	0--	7**	1+	0--	1

Keterangan:

** : Kunci Jawaban

++ : Sangat Baik

+ : Baik

- : Kurang Baik

-- : Buruk

---: Sangat Buruk

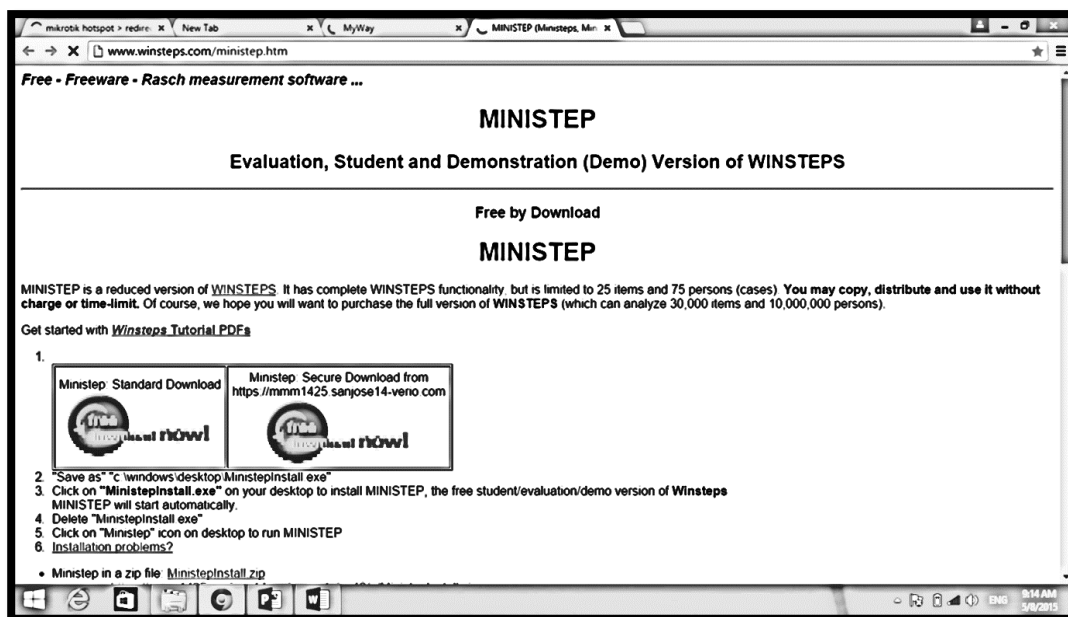
C. Pengujian Instrumen Penelitian dengan Aplikasi Program Rasch Model

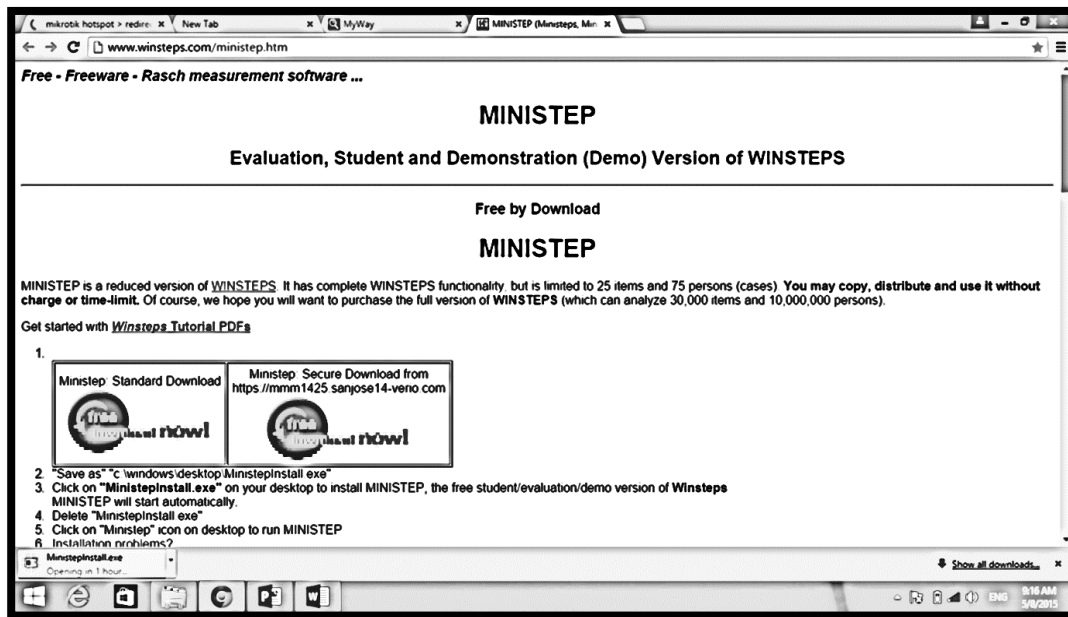
1. Petunjuk Penggunaan Perangkat Lunak Ministep Dalam Analisis Pemodelan Rasch

Ada beberapa langkah yang perlu dilakukan dalam menggunakan program Ministep yaitu :

a) Instalasi Ministep

Perangkat Lunak (*software*) Ministep adalah program komputer khusus untuk analisis pemodelan Rasch yang bisa bekerja di bawah system Microsoft Windows yang dibuat oleh Jhon Linacre. Paket program ini bersifat freeware (bisa digunakan tanpa perlu membeli atau mendapatkan izin pemakaian), yang merupakan versi terbatas dari Winsteps. Kemampuan pengolahan data untuk Ministep adalah 25 aitem dan 75 responden. Program Ministep ini dapat diunduh (*download*) secara gratis di laman situs (*website*):





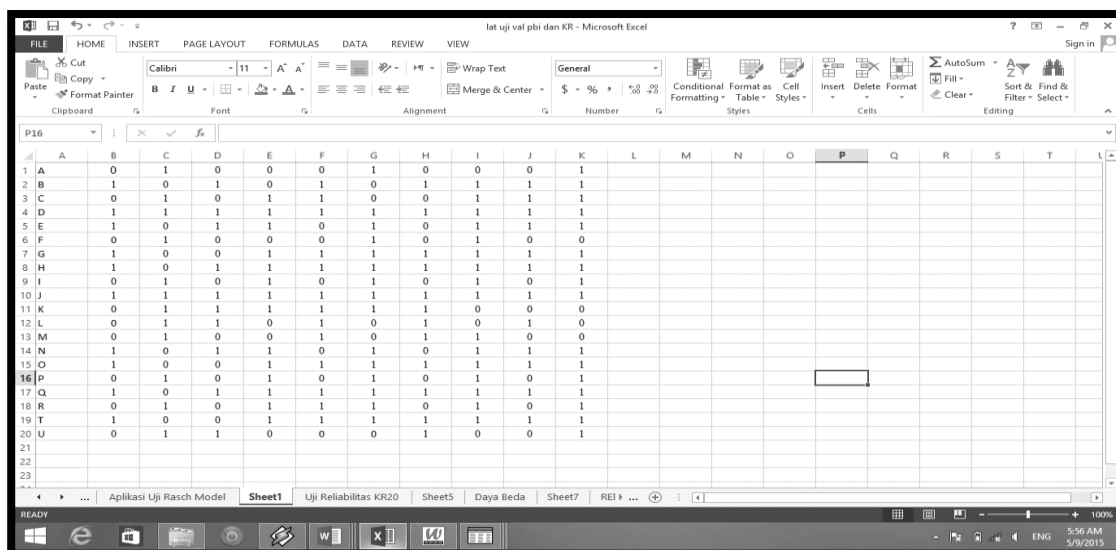
Klik tombol *download* untuk memulai mengunduh program, dan simpan dalam folder yang telah dipersiapkan.

Apabila sudah selesai mengunduh, klik tombol untuk menginstalasi program dalam komputer atau laptop Anda. Program pun siap untuk digunakan.

b) Penyiapan Berkas Data (File Data)

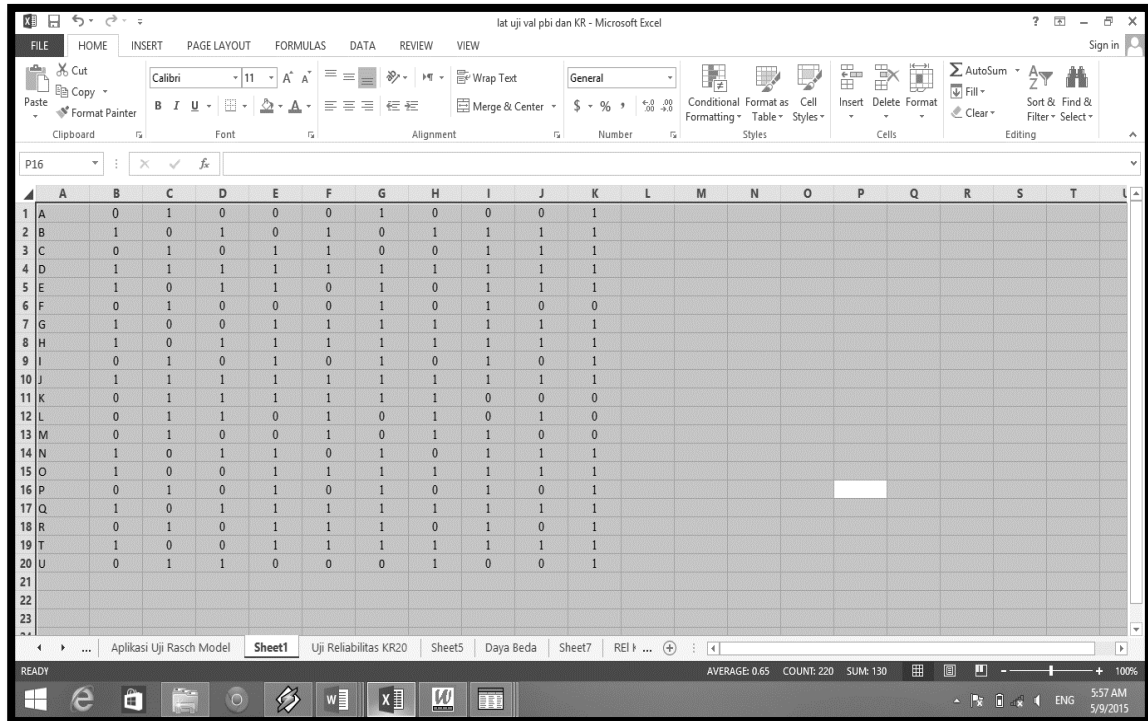
Pengolahan data dengan menggunakan Ministep memerlukan jenis fail tertentu, yang hanya berbentuk data. Oleh karena itu, data perlu dipersiapkan dengan perangkat lunak *spreadsheet* seperti Microsoft Excel. Untuk contoh penyiapan data dan penggunaan Ministep, digunakan data dari scalogram yang dibahas di bagian 2.4.1 pada buku ini.

Mengisi data (data entry) dengan menggunakan Microsoft Excel :

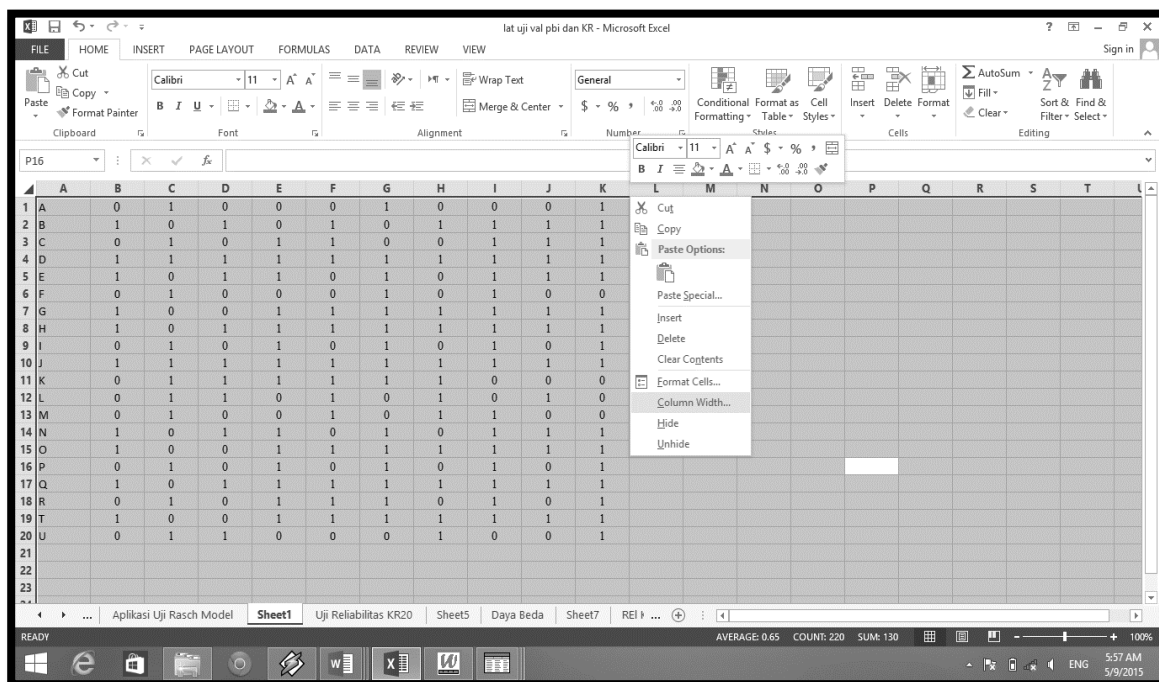


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

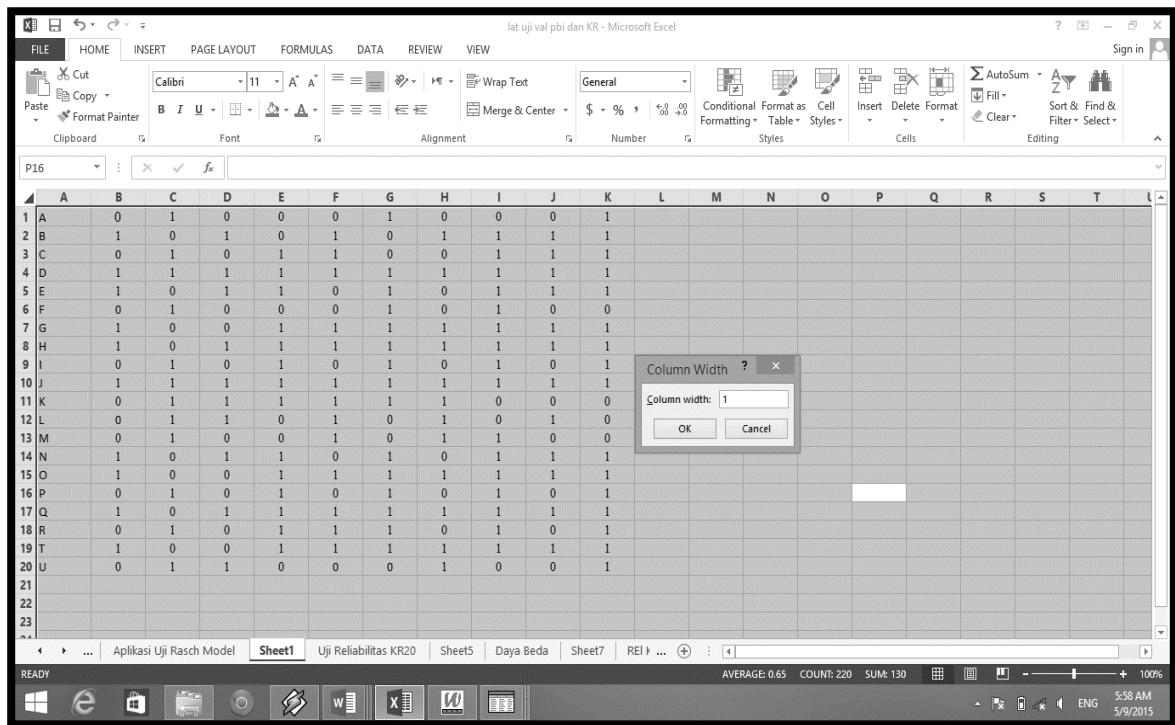
Kolom ke-1 berisi identitas responden (A sampai J), sedangkan untuk kolom berikutnya langsung diisi dengan data yang didapat dari scalogram. Apabila sudah selesai mengisi data, simpan (save) fail dengan nama scalogram, misalnya.



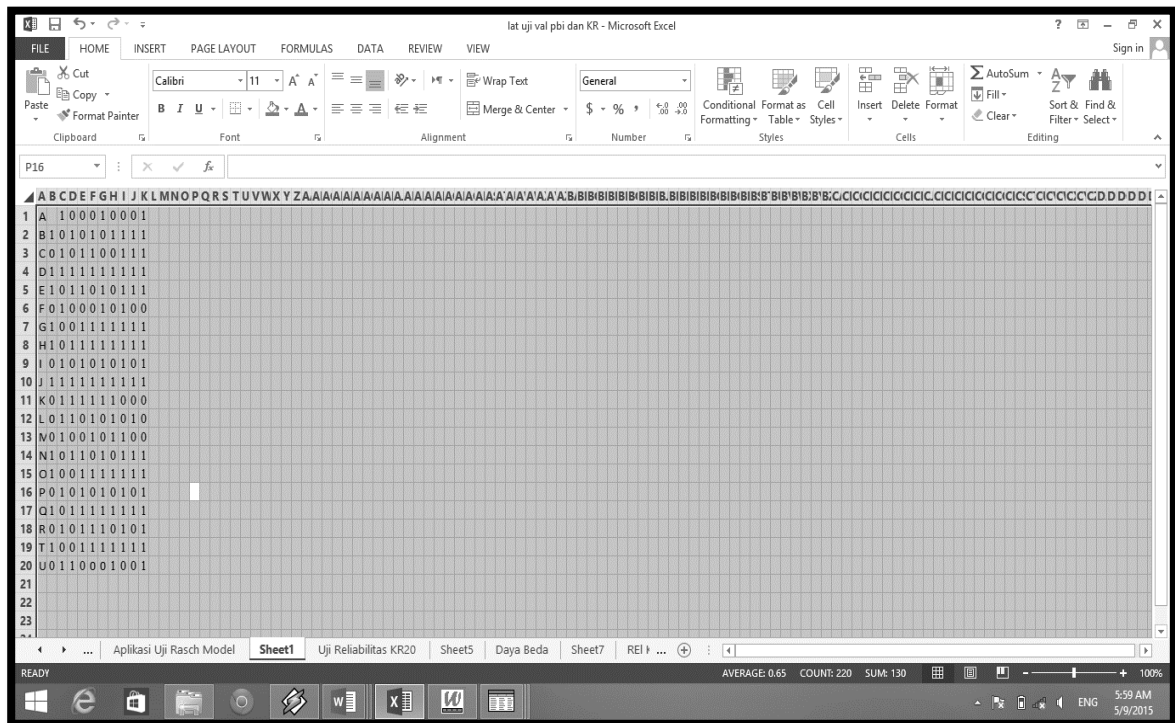
Kemudian, blok semua bagian dan ubah panjang kolom menjadi 1 dengan klik kanan dari tetikus anda, dan Klik OK.



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

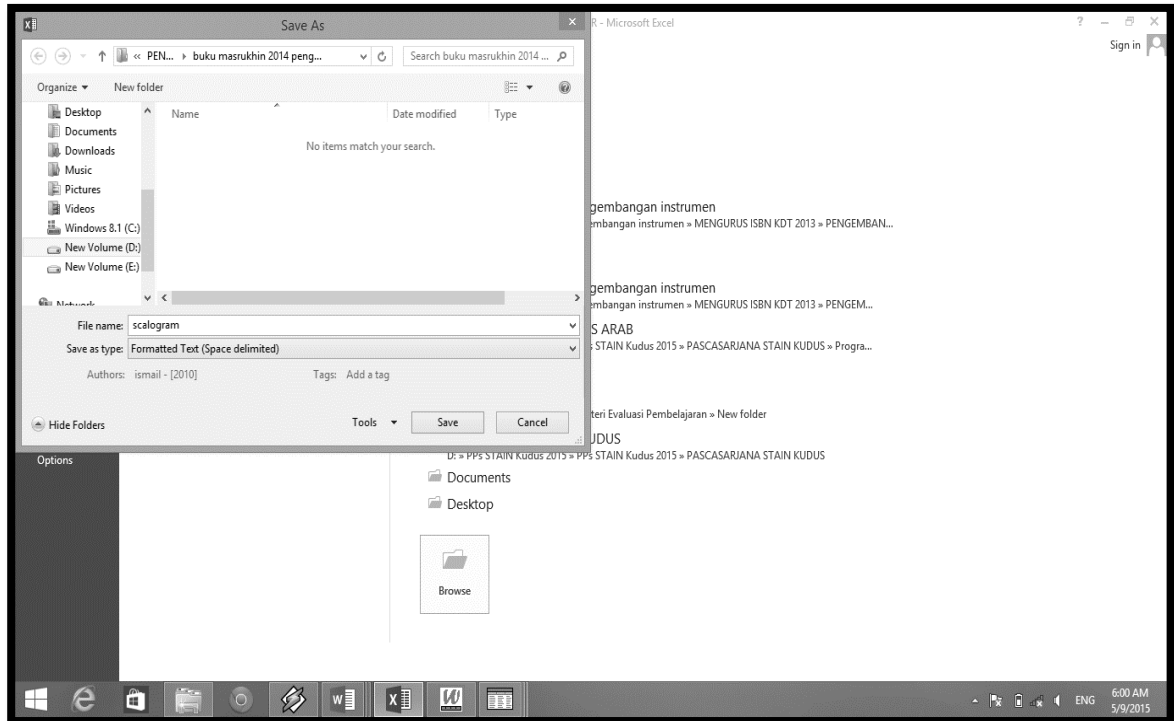


Klik OK

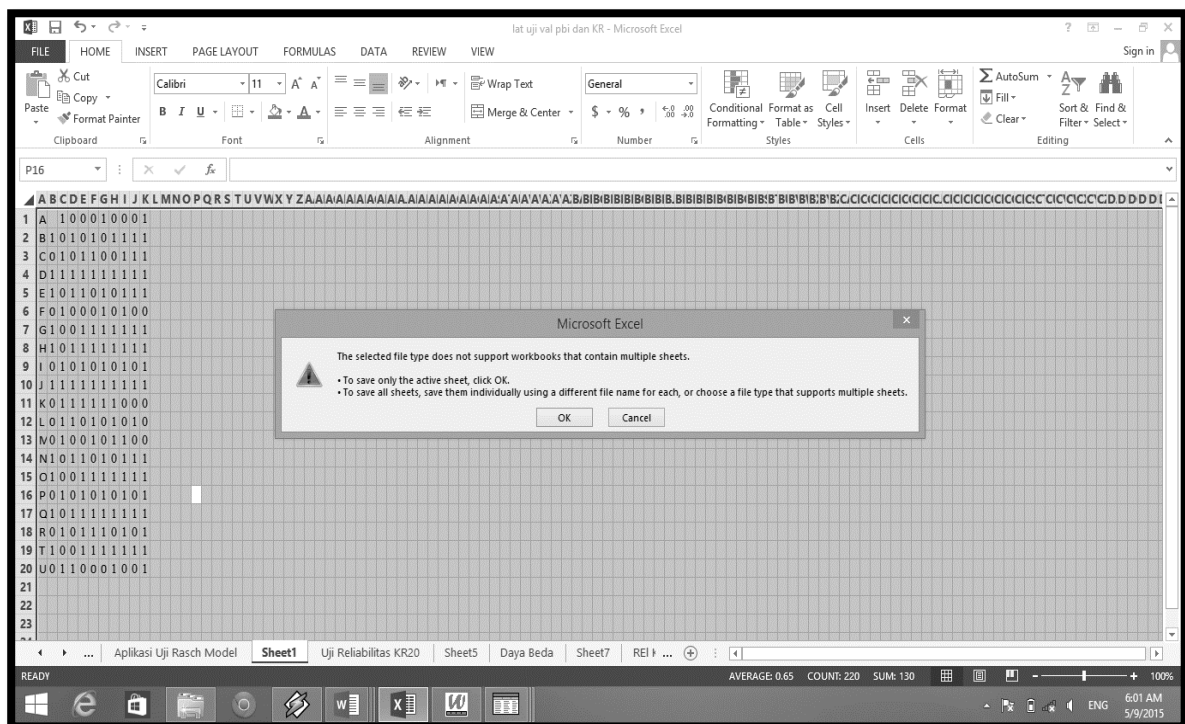


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

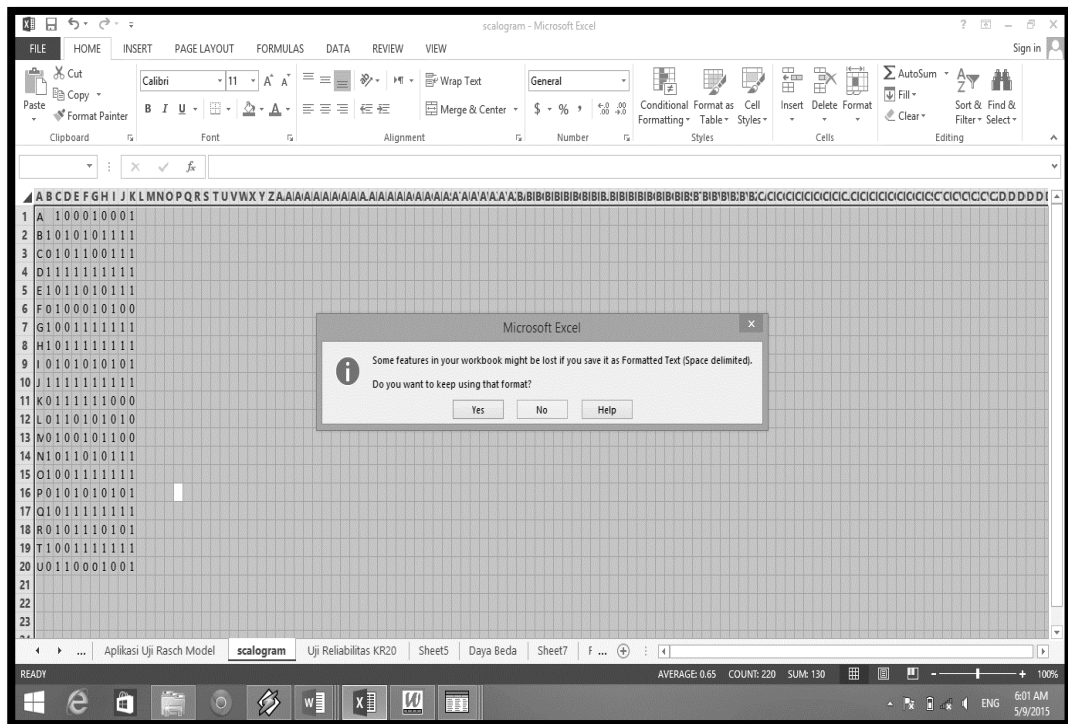
Langkah berikutnya adalah menyimpan berkas data dalam bentuk file *.prn. Caranya adalah dengan menekan tombol Save As dan Pilihlah format '**Formatted Text (space delimited)**' dan beri nama scalogram sehingga nanti Anda mendapatkan berkas: scalogram.prn



Klik Ok



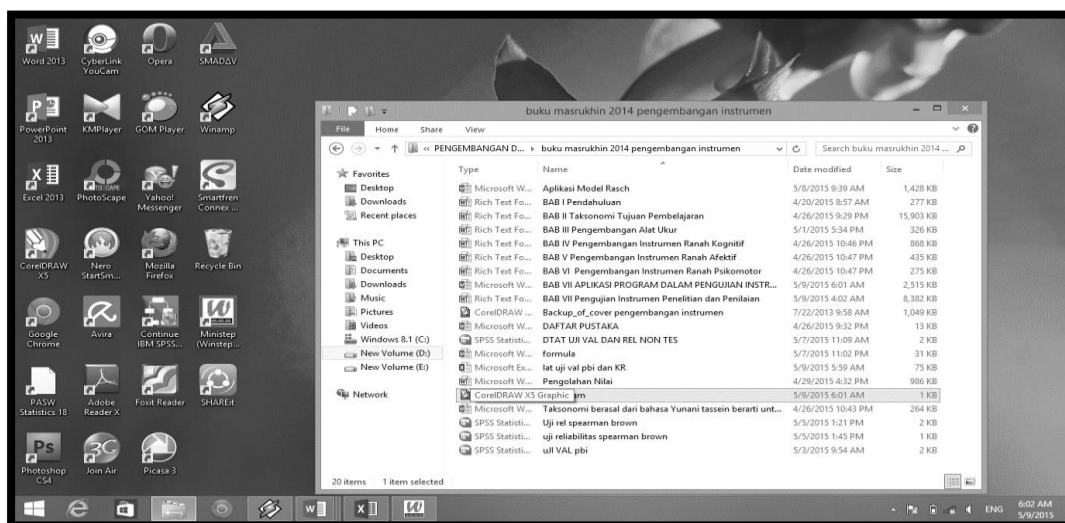
Klik yes.



Berkas scalogram.prn dapat dibuka dengan program Notepad, tutup (Close) semua berkas, tahapan penyiapan berkas data dari MS Excel selesai.

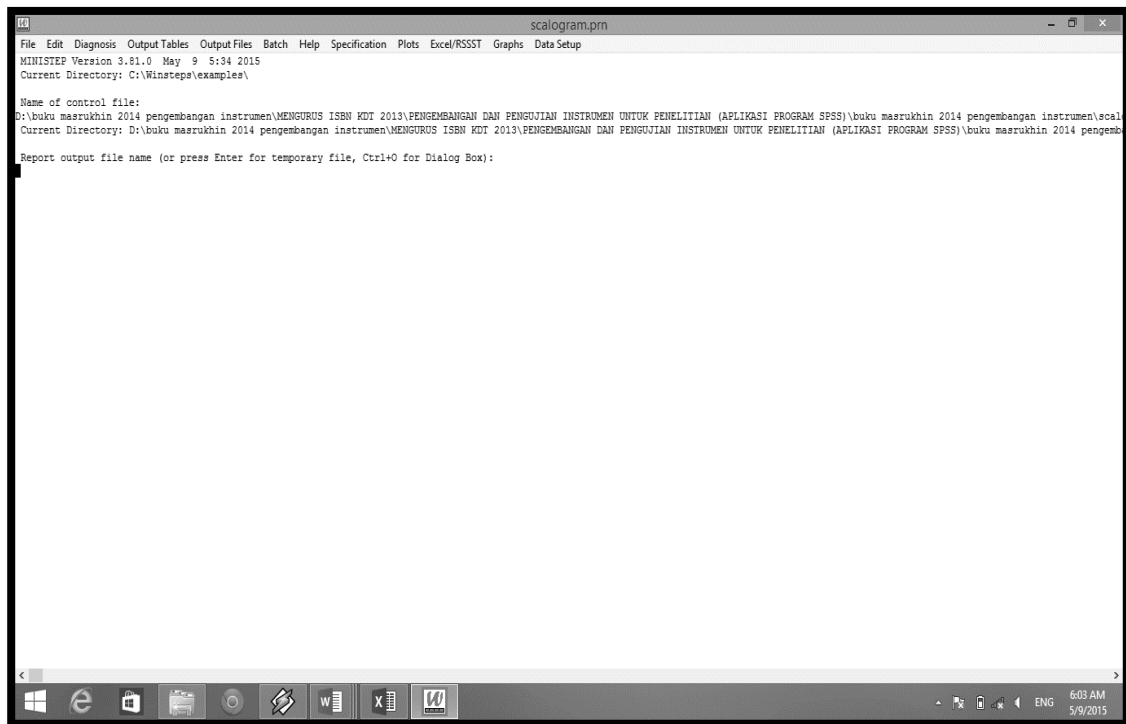
c) Penyiapan Berkas Data dalam Ministep

Setelah berkas data siap dalam bentuk *.prn (yaitu: **scalogram.prn**) maka perlu penyiapan berkas tersebut dalam perangkat lunak Ministep. Cara yang mudah untuk dilakukan adalah dengan menggeser file*.prn tadi (**scalogram.prn**) “dimasukkan” ke ikon Ministep, seperti tampilan berikut:



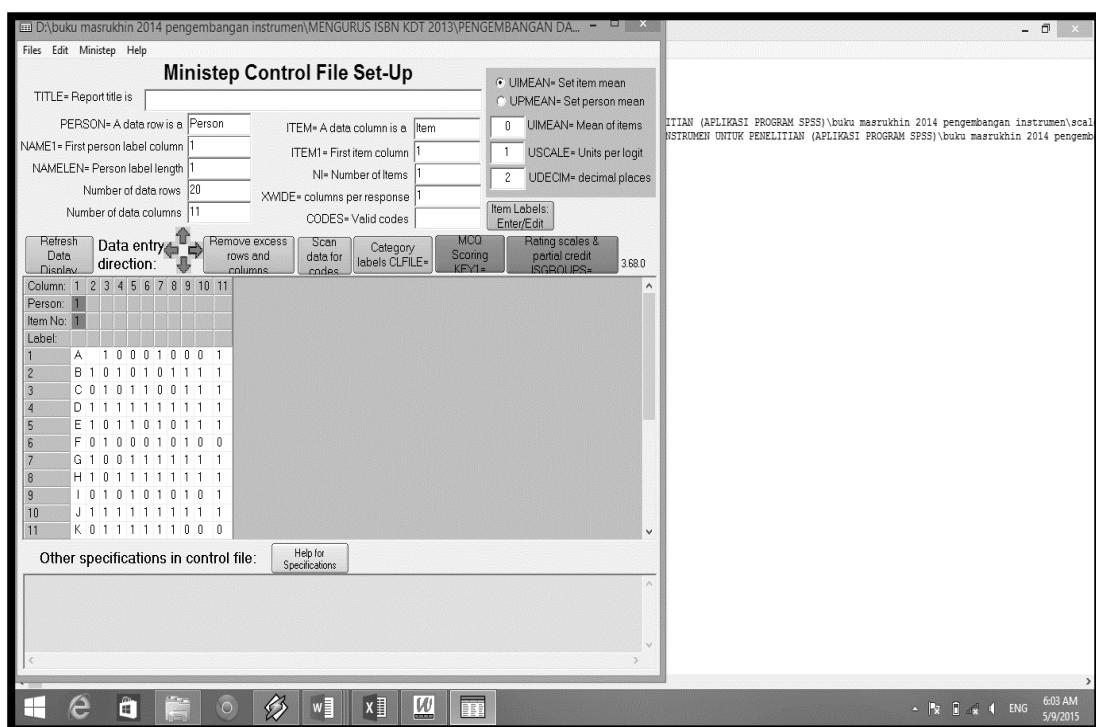
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

Akan muncul tampilan seperti berikut ini:



Langkah selanjutnya adalah mengeklik tombol **Data Setup** (ada di bagian kanan atas layar komputer) yaitu melakukan konfigurasi data sesuai dengan format Ministep.

Tampilan yang muncul adalah seperti berikut ini:



Keterangan untuk mengisi data untuk Person (responden) yang ada di berkas data:

Ministep C

TITLE= Report title is

PERSON= A data row is a Person

NAME1= First person label column 1

NAMELEN= Person label length 1

Number of data rows 20

Number of data columns 11

Refresh Data Display Remove rows and columns

Data entry direction:

Column:	1	2	3	4	5	6	7	8	9	10	11
Person:	1										
Item No:	1										
Label:											
1	A	1	0	0	0	1	0	0	0	1	
2	B	1	0	1	0	1	0	1	1	1	1
3	C	0	1	0	1	1	0	0	1	1	1
4	D	1	1	1	1	1	1	1	1	1	1
5	E	1	0	1	1	0	1	0	1	1	1
6	F	0	1	0	0	0	1	0	1	0	0
7	G	1	0	0	1	1	1	1	1	1	1
8	H	1	0	1	1	1	1	1	1	1	1
9	I	0	1	0	1	0	1	0	1	0	1
10	J	1	1	1	1	1	1	1	1	1	1
11	K	0	1	1	1	1	1	0	0	0	

Other specifications in control file

Keterangan untuk mengisi data item yang ada di berkas data:

Control File Set-Up

ITEM= A data column is a Item

ITEM1= First item column 1

NI= Number of items 1

XWIDE= columns per response 1

CODES= Valid codes

UMEAN= Set item mean 0

UMEAN= Set person mean 1

USCALE= Units per logit 2

UDECM= decimal places

Item Labels Enter/Edit

Scan data for codes

Category labels CLFILE

MOQ Scoring KEY1

Rating scales & partial credit ISGRN IPS

Help for Specifications

Untuk berkas data scalogram.prn maka yang perlu diperbaiki adalah bagian aitem,yaitu:

First item column= diisi dengan angka 2 (kolom pertama data dimulai)

Number of Items= diisi dengan angka 10 (banyaknya aitem dalam berkas scalogram)

Tampilan layar computer menjadi seperti berikut ini:

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a Person

NAME1= First person label column 1

NAMELEN= Person label length 1

Number of data rows 20

Number of data columns 11

ITEM= A data column is a Item

ITEM1= First item column 2

NI= Number of Items 10

XWIDE= columns per response 1

CODES= Valid codes

UIMEAN= Set item mean

UPMEAN= Set person mean

0 UIMEAN= Mean of items

1 USCALE= Units per logit

2 UDECIM= decimal places

Item Labels: Enter/Edit

Refresh Data Display

Data entry direction: Person

Remove excess rows and columns

Scan data for codes

Category labels CLFILE=

MCQ Scoring KEY1=

Rating scales & partial credit ISGROUPS=

368.0

Column:	1	2	3	4	5	6	7	8	9	10	11
Person:	1										
Item No:	1										
Label:											
1	A	1	0	0	0	1	0	0	0	1	
2	B	1	0	1	0	1	0	1	1	1	
3	C	0	1	0	1	1	0	0	1	1	
4	D	1	1	1	1	1	1	1	1	1	
5	E	1	0	1	1	0	1	0	1	1	
6	F	0	1	0	0	0	1	0	1	0	
7	G	1	0	0	1	1	1	1	1	1	
8	H	1	0	1	1	1	1	1	1	1	
9	I	0	1	0	1	0	1	0	1	0	
10	J	1	1	1	1	1	1	1	1	1	
11	K	0	1	1	1	1	1	0	0	0	

Other specifications in control file:

Help for Specifications

Langkah berikutnya adalah menyiapkan label untuk item :

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a Person

NAME1= First person label column 1

NAMELEN= Person label length 1

Number of data rows 20

Number of data columns 11

ITEM= A data column is a Item

ITEM1= First item column 2

NI= Number of Items 10

XWIDE= columns per response 1

CODES= Valid codes

UIMEAN= Set item mean

UPMEAN= Set person mean

0 UIMEAN= Mean of items

1 USCALE= Units per logit

2 UDECIM= decimal places

Item Labels: Enter/Edit

Refresh Data Display

Data entry direction: Person

Remove excess rows and columns

Scan data for codes

Category labels CLFILE=

MCQ Scoring KEY1=

Rating scales & partial credit ISGROUPS=

368.0

Column:	1	2	3	4	5	6	7	8	9	10	11
Person:	1										
Item No:	1										
Label:											
1	A	1	0	0	0	1	0	0	0	1	
2	B	1	0	1	0	1	0	1	1	1	
3	C	0	1	0	1	1	0	0	1	1	
4	D	1	1	1	1	1	1	1	1	1	
5	E	1	0	1	1	0	1	0	1	1	
6	F	0	1	0	0	0	1	0	1	0	
7	G	1	0	0	1	1	1	1	1	1	
8	H	1	0	1	1	1	1	1	1	1	
9	I	0	1	0	1	0	1	0	1	0	
10	J	1	1	1	1	1	1	1	1	1	
11	K	0	1	1	1	1	1	0	0	0	

Other specifications in control file:

Help for Specifications

Item Labels: Enter/Edit

Edit

NI = Number of Items= 10

Item Labels OK

Number	Label
1	a
2	b
3	c
4	d
5	e
6	f
7	g
8	h
9	i
10	j

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a Person

NAME1= First person label column 1

NAMELEN= Person label length 1

Number of data rows 20

Number of data columns 11

ITEM= A data column is a Item

ITEM1= First item column 2

NI= Number of Items 10

XWIDE= columns per response 1

CODES= Valid codes

UIMEAN= Set item mean

UPMEAN= Set person mean

0 UIMEAN= Mean of items

1 USCALE= Units per logit

2 UDECIM= decimal places

Item Labels: Enter/Edit

Refresh Data Display

Data entry direction: Person

Remove excess rows and columns

Scan data for codes

Category labels CLFILE=

MOO Scoring KEY1=

Rating scales & partial credit ISGROUPS=

3680

Column:	1	2	3	4	5	6	7	8	9	10	11
Person:	1	2	3	4	5	6	7	8	9	10	11
Item No:	1	2	3	4	5	6	7	8	9	10	
Label:	a	b	c	d	e	f	g	h	i	j	
1	A	1	0	0	0	1	0	0	0	1	
2	B	1	0	1	0	1	0	1	1	1	
3	C	0	1	0	1	1	0	0	1	1	
4	D	1	1	1	1	1	1	1	1	1	
5	E	1	0	1	1	0	1	0	1	1	
6	F	0	1	0	0	0	1	0	1	0	
7	G	1	0	0	1	1	1	1	1	1	
8	H	1	0	1	1	1	1	1	1	1	
9	I	0	1	0	1	0	1	0	1	0	
10	J	1	1	1	1	1	1	1	1	1	
11	K	0	1	1	1	1	1	0	0	0	

Other specifications in control file:

Help for Specifications

Setelah memberikan label untuk aitem, langkah selanjutnya adalah pengecekan kode: Scan data for Code.

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a Person

NAME1= First person label column 1

NAMELEN= Person label length 1

Number of data rows 20

Number of data columns 11

ITEM= A data column is a Item

ITEM1= First item column 2

NI= Number of Items 10

XWIDE= columns per response 1

CODES= Valid codes

UIMEAN= Set item mean

UPMEAN= Set person mean

0 UIMEAN= Mean of items

1 USCALE= Units per logit

2 UDECIM= decimal places

Item Labels: Enter/Edit

Refresh Data Display

Data entry direction: Person

Remove excess rows and columns

Scan data for codes

Category labels CLFILE=

MOO Scoring KEY1=

Rating scales & partial credit ISGROUPS=

3680

Column:	1	2	3	4	5	6	7	8	9	10	11
Person:	1	2	3	4	5	6	7	8	9	10	11
Item No:	1	2	3	4	5	6	7	8	9	10	
Label:	a	b	c	d	e	f	g	h	i	j	
1	A	1	0	0	0	1	0	0	0	1	
2	B	1	0	1	0	1	0	1	1	1	
3	C	0	1	0	1	1	0	0	1	1	
4	D	1	1	1	1	1	1	1	1	1	
5	E	1	0	1	1	0	1	0	1	1	
6	F	0	1	0	0	0	1	0	1	0	
7	G	1	0	0	1	1	1	1	1	1	
8	H	1	0	1	1	1	1	1	1	1	
9	I	0	1	0	1	0	1	0	1	0	
10	J	1	1	1	1	1	1	1	1	1	
11	K	0	1	1	1	1	1	0	0	0	

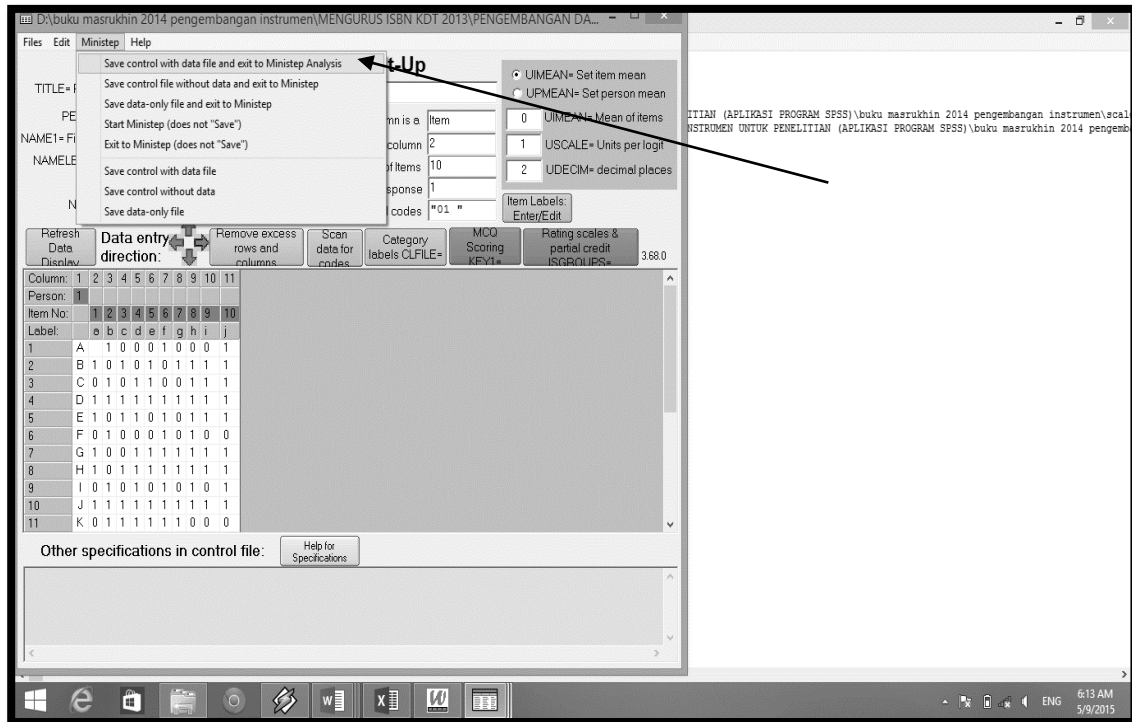
Other specifications in control file:

Help for Specifications

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

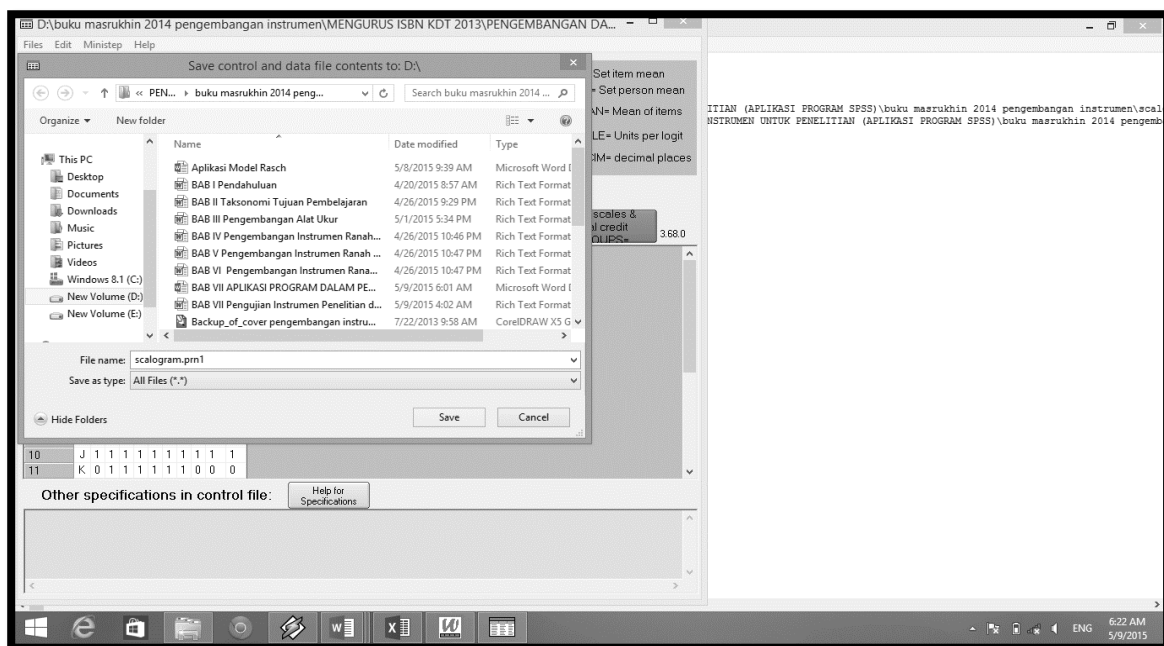
Pemindaian (scan) data adalah hal yang penting untuk mengetahui data seperti apa yang akan diolah Ministep, untuk contoh diatas ada tiga jenis, yaitu "0", "1" dan "..." (spasi atau data hilang). Apabila terdapat ketidaksesuaian jenis data, proses *scan* akan memberi tahu ada jenis *data entry* yang salah dimasukkan, dan Ministep tidak bias dipaksakan untuk menganalisis data.

Apabila konfigurasi data sudah selesai, fail bias disimpan dan dilanjutkan dengan pengolahan data dengan pemodelan Rasch.

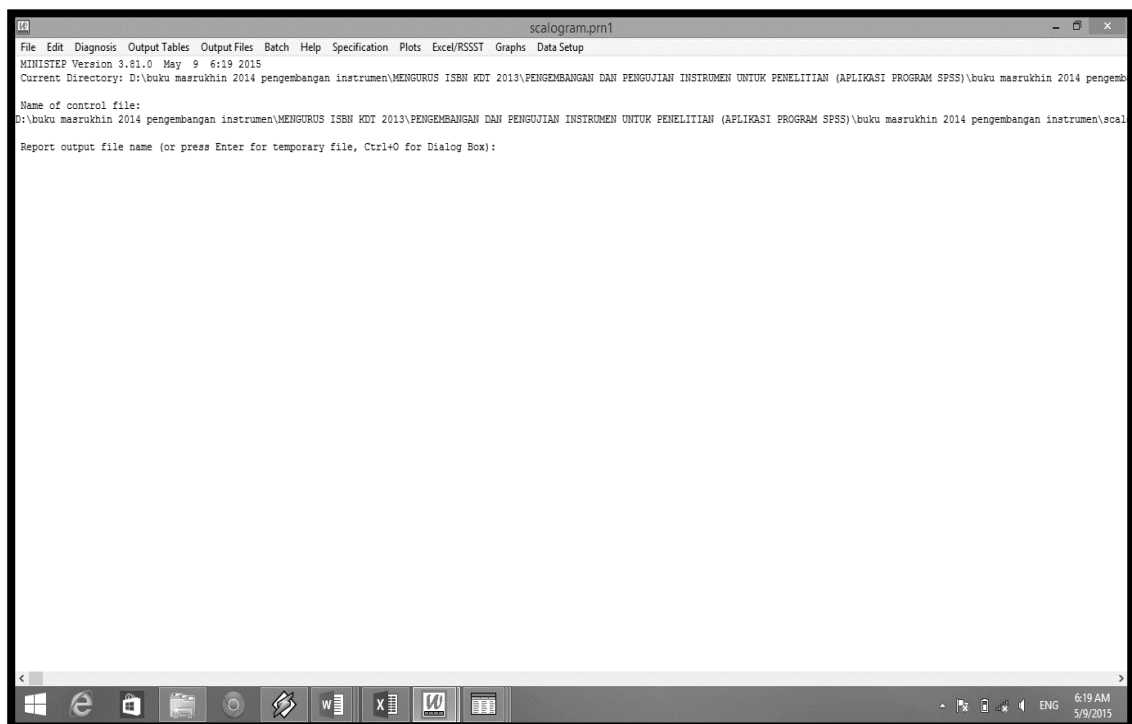
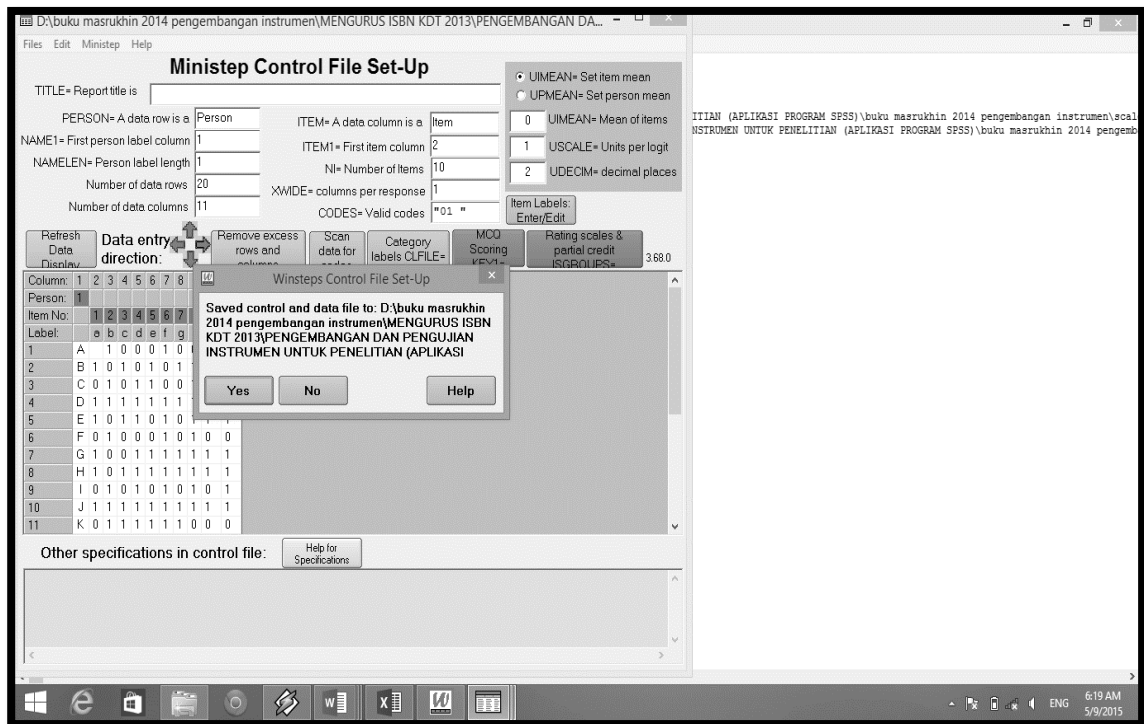


Berikan nama fail untuk data yang siap dianalisis ini, contoh **scalogram.prn**.

Klik tombol **Save** dan tekan tombol **Enter** dua kali.



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



d) Pengolahan Data Pemodelan Rasch dengan Ministep

Apabila penyiapan data untuk pengolahan oada Ministep berhasil, akan muncul tampilan. Keterangan:

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

scalogram.prn1

File Edit Diagnosis Output Tables Output Files Batch Help Specification Plots Excel/RSSST Graphs Data Setup

PROX ACTIVE COUNT EXTREME 5 RANGE MAX LOGIT CHANGE
 ITERATION Person Item CATS Person Item MEASURES STRUCTURE

1 20 10 2 2.18 .80 2.1972
 2 18 10 2 2.28 1.00 .2170
 3 18 10 2 2.33 1.00 .0480

Probing data connection: to skip out: Ctrl+F - to bypass: subset=no
 Processing unanchored persons ...

Data fully connected. No subsets found
 (Control: \scalogram.prn1 Output: \ZOU439WS.TXT
 JMLE MAX SCORE MAX LOGIT LEAST CONVERGED CATEGORY STRUCTURE
 ITERATION RESIDUAL* CHANGE Person Item CAT RESIDUAL CHANGE

1 .17 .3414 4* 3
 2 .04 .0980 4* 3

Calculating Fit Statistics
 Standardized Residuals N(0,1) Mean: .01 S.D.: .98
 Time for estimation: 0:0:0.571
 Processing Table 0
 D:\buku masrukhin 2014 pengembangan instrumen\MENGGURUS ISBN KDT 2013\PENGEMBAN

Person 20 INPUT 20 MEASURED INFIT OUTFIT
 TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD
 MEAN 6.5 9.9 .89 .88 1.00 .1 .95 .01
 S.D. 2.1 .2 1.28 .33 .22 1.1 .28 1.1
 REAL RMSE .94 TRUE SD .86 SEPARATION .91 Person RELIABILITY .45

Item 10 INPUT 10 MEASURED INFIT OUTFIT
 TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD
 MEAN 13.0 19.9 .00 .58 1.00 -.1 .95 -.2
 S.D. 2.1 .3 .60 .07 .45 1.9 .59 1.8
 REAL RMSE .58 TRUE SD .15 SEPARATION .26 Item RELIABILITY .06

Output written to D:\buku masrukhin 2014 pengembangan instrumen\MENGGURUS ISBN KDT 2013\PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN (APLIKASI PROGRAM SPSS)\buku masrukhin 2014 peng
 CODES= 01
 Measures constructed: use "Diagnosis" and "Output Tables" menus

Beberapa pilihan yang penting untuk analisis data sebagai berikut:

Tabel 3.1. Summary Statistics

scalogram.prn1

File Edit Diagnosis Output Tables Output Files Batch Help Specification Plots Excel/RSSST Graphs Data Setup

PROX Request Subtables 1. Variable maps 20. Score table
 ITERATION Person 3.2+ Rating (partial credit) scale 2.2 General Keyform 21. Probability curves
 1 2 2. Measure forms (all) 2.5 Category Averages 29. Empirical curves
 2 1 3.1 Summary statistics 22. Scalograms

10. Item (column): fit order 6. Person (row): fit order 40. Person Keyforms: fit order
 13. Item: measure 17. Person: measure 37. Person Keyforms: measure
 14. Item: entry 18. Person: entry 38. Person Keyforms: entry
 15. Item: alphabetical 19. Person: alphabetical 39. Person Keyforms: alphabetical
 25. Item: displacement 42. Person: displacement 41. Person Keyforms: unexpected
 26. Item: correlation 43. Person: correlation
 11. Item: responses 7. Person: responses 32. Control variable list
 9. Item: outfit plot 5. Person: outfit plot 33. Person-Item: DGF: DIF & DPF
 8. Item: infit plot 4. Person: infit plot 34. Comparison of two statistics
 12. Item: map 16. Person: map 35. Person Paired Agreement
 23. Item: dimensionality 24. Person: dimensionality 36. Person diagnostic PKMAPs
 27. Item: subtotals 28. Person: subtotals
 30. Item: DIF, between/within 31. Person: DPF, between/within

Standardized Residuals N(0,1) Mean: .01 S.D.: .98
 Time for estimation: 0:0:0.391
 Processing Table 0
 D:\buku masrukhin 2014 pengembangan instrumen\MENGGURUS ISBN KDT 2013\PENGEMBAN

Person 20 INPUT 20 MEASURED INFIT OUTFIT
 TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD
 MEAN 6.5 9.9 .89 .88 1.00 .1 .95 .01
 S.D. 2.1 .2 1.28 .33 .22 1.1 .28 1.1
 REAL RMSE .94 TRUE SD .86 SEPARATION .91 Person RELIABILITY .45

Item 10 INPUT 10 MEASURED INFIT OUTFIT
 TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD
 MEAN 13.0 19.9 .00 .58 1.00 -.1 .95 -.2
 S.D. 2.1 .3 .60 .07 .45 1.9 .59 1.8
 REAL RMSE .58 TRUE SD .15 SEPARATION .26 Item RELIABILITY .06

Output written to D:\buku masrukhin 2014 pengembangan instrumen\MENGGURUS ISBN KDT 2013\PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN (APLIKASI PROGRAM SPSS)\buku masrukhin 2014 peng
 CODES= 01
 Measures constructed: use "Diagnosis" and "Output Tables" menus
 Processing Table 3.1

03-377WS - Notepad

File Edit Format View Help

INPUT: 20 Person 10 Item REPORTED: 20 Person 10 Item 2 CATS MINISTEP 3.81.0

SUMMARY OF 18 MEASURED (NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	6.1	9.9	.59	.75	1.00	.1	.95	.0
S.D.	1.9	.2	.95	.12	.22	1.1	.28	1.1
MAX.	9.0	10.0	2.33	1.07	1.55	2.7	1.58	2.6
MIN.	3.0	9.0	-.91	.66	.70	-1.8	.66	-1.8

REAL RMSE .79 TRUE SD .53 SEPARATION .68 Person RELIABILITY .31
 MODEL RMSE .76 TRUE SD .57 SEPARATION .75 Person RELIABILITY .36
 S.E. OF Person MEAN = .23

MAXIMUM EXTREME SCORE: 2 Person 10.0%

SUMMARY OF 20 MEASURED (EXTREME AND NON-EXTREME) Person

	TOTAL SCORE	COUNT	MEASURE	MODEL ERROR	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD
MEAN	6.5	9.9	.89	.86				
S.D.	2.1	.2	1.28	.34				
MAX.	10.0	10.0	3.59	1.83				
MIN.	3.0	9.0	-.91	.66				

REAL RMSE .94 TRUE SD .86 SEPARATION .91 Person RELIABILITY .45
 MODEL RMSE .92 TRUE SD .88 SEPARATION .95 Person RELIABILITY .48
 S.E. OF Person MEAN = .29

Person RAW SCORE-TO-MEASURE CORRELATION = .98
 CRONBACH ALPHA (KR-20) Person RAW SCORE "TEST" RELIABILITY = .57

Nilai rata-rata logit untuk Pearson

Nilai Reliabilitas untuk Pearson

Nilai Alpha Cronbach

1. Person measure= 0,59 logit menunjukkan rata-rata kemampuan aitem

2. Nilai Person *Reliability* dan Item *Reliability*:

< 0,67: Lemah 0,67-0,80: Cukup 0,81-0,90: Bagus
 0,91-0,94: Bagus sekali > 0,94: Istimewa

Nilai Pearson Reliabilitas sebesar 0,31 masuk dalam kategori LEMAH

3. Nilai *alpha cronbach*:

< 0,5:Buruk 0,5-0,6: Jelek 0,6-0,7:Cukup
 0,7-0,8: Bagus > 0,8: Bagus Sekali

Nilai Alpha Cronbach sebesar 0,57 masuk dalam kategori JELEK

Tabel 13: Item Measure

Tabel berikut memberikan informasi mengenai aitem:

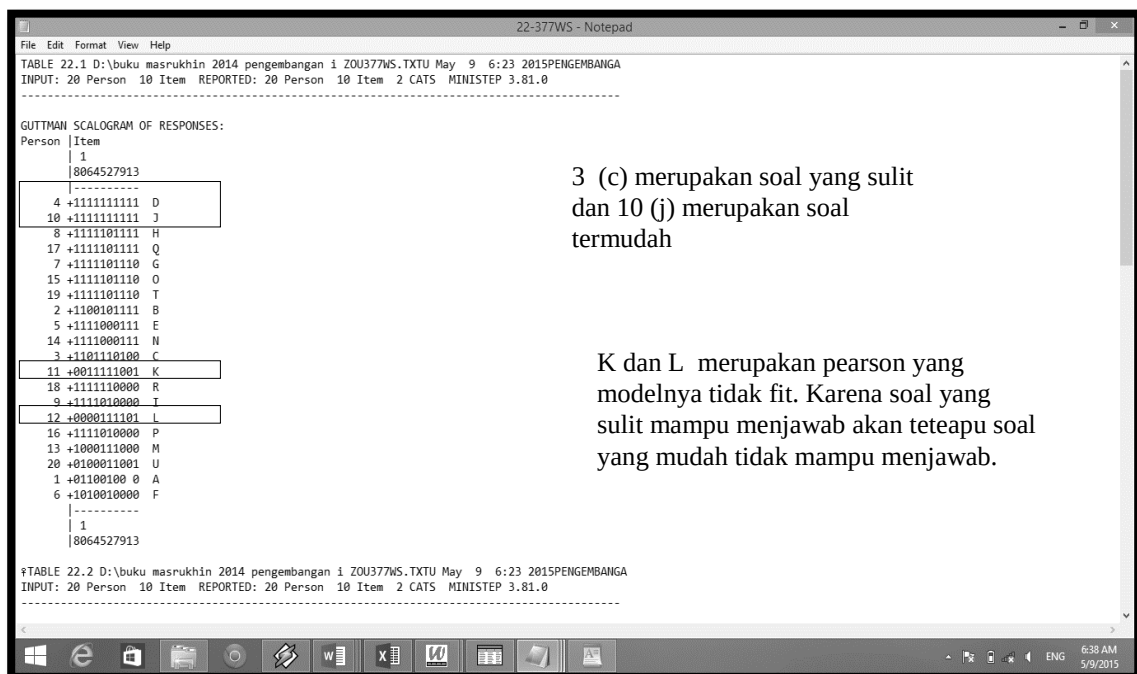
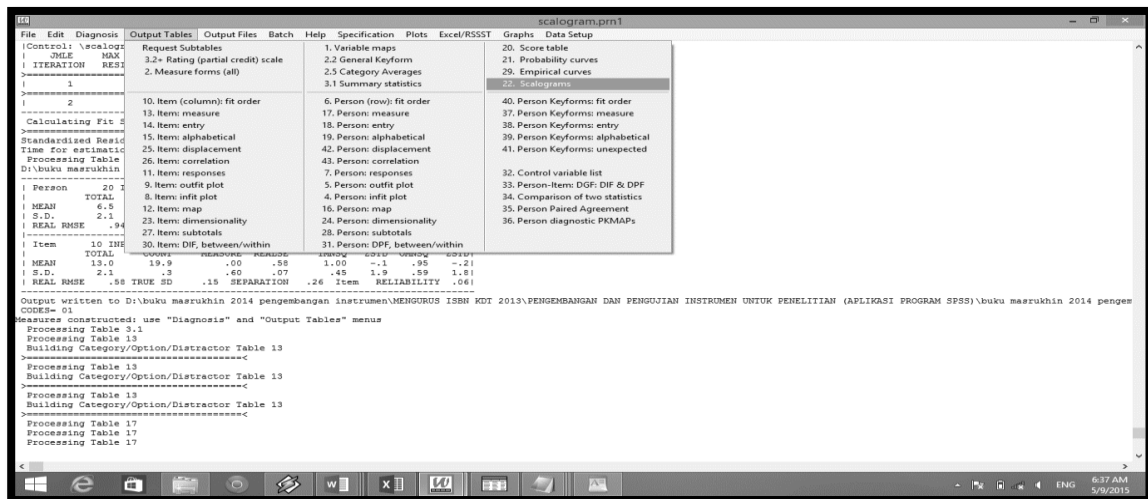
The screenshot shows the SPSS 20.0.0.993 interface with the 'Data' menu open. The menu options are as follows:

- File Edit Diagnosis
- 1 PROX
- 1 ITERATION PERM
- 1 1
- 1 2
- 1 3
- 1 4
- 1 5
- 1 6
- 1 7
- 1 8
- 1 9
- 1 10
- 1 11
- 1 12
- 1 13
- 1 14
- 1 15
- 1 16
- 1 17
- 1 18
- 1 19
- 1 20
- 1 21
- 1 22
- 1 23
- 1 24
- 1 25
- 1 26
- 1 27
- 1 28
- 1 29
- 1 30
- 1 31
- 1 32
- 1 33
- 1 34
- 1 35
- 1 36
- 1 37
- 1 38
- 1 39
- 1 40
- 1 41
- 1 42
- 1 43
- 1 44
- 1 45
- 1 46
- 1 47
- 1 48
- 1 49
- 1 50
- 1 51
- 1 52
- 1 53
- 1 54
- 1 55
- 1 56
- 1 57
- 1 58
- 1 59
- 1 60
- 1 61
- 1 62
- 1 63
- 1 64
- 1 65
- 1 66
- 1 67
- 1 68
- 1 69
- 1 70
- 1 71
- 1 72
- 1 73
- 1 74
- 1 75
- 1 76
- 1 77
- 1 78
- 1 79
- 1 80
- 1 81
- 1 82
- 1 83
- 1 84
- 1 85
- 1 86
- 1 87
- 1 88
- 1 89
- 1 90
- 1 91
- 1 92
- 1 93
- 1 94
- 1 95
- 1 96
- 1 97
- 1 98
- 1 99
- 1 100
- 1 101
- 1 102
- 1 103
- 1 104
- 1 105
- 1 106
- 1 107
- 1 108
- 1 109
- 1 110
- 1 111
- 1 112
- 1 113
- 1 114
- 1 115
- 1 116
- 1 117
- 1 118
- 1 119
- 1 120
- 1 121
- 1 122
- 1 123
- 1 124
- 1 125
- 1 126
- 1 127
- 1 128
- 1 129
- 1 130
- 1 131
- 1 132
- 1 133
- 1 134
- 1 135
- 1 136
- 1 137
- 1 138
- 1 139
- 1 140
- 1 141
- 1 142
- 1 143
- 1 144
- 1 145
- 1 146
- 1 147
- 1 148
- 1 149
- 1 150
- 1 151
- 1 152
- 1 153
- 1 154
- 1 155
- 1 156
- 1 157
- 1 158
- 1 159
- 1 160
- 1 161
- 1 162
- 1 163
- 1 164
- 1 165
- 1 166
- 1 167
- 1 168
- 1 169
- 1 170
- 1 171
- 1 172
- 1 173
- 1 174
- 1 175
- 1 176
- 1 177
- 1 178
- 1 179
- 1 180
- 1 181
- 1 182
- 1 183
- 1 184
- 1 185
- 1 186
- 1 187
- 1 188
- 1 189
- 1 190
- 1 191
- 1 192
- 1 193
- 1 194
- 1 195
- 1 196
- 1 197
- 1 198
- 1 199
- 1 200
- 1 201
- 1 202
- 1 203
- 1 204
- 1 205
- 1 206
- 1 207
- 1 208
- 1 209
- 1 210
- 1 211
- 1 212
- 1 213
- 1 214
- 1 215
- 1 216
- 1 217
- 1 218
- 1 219
- 1 220
- 1 221
- 1 222
- 1 223
- 1 224
- 1 225
- 1 226
- 1 227
- 1 228
- 1 229
- 1 230
- 1 231
- 1 232
- 1 233
- 1 234
- 1 235
- 1 236
- 1 237
- 1 238
- 1 239
- 1 240
- 1 241
- 1 242
- 1 243
- 1 244
- 1 245
- 1 246
- 1 247
- 1 248
- 1 249
- 1 250
- 1 251
- 1 252
- 1 253
- 1 254
- 1 255
- 1 256
- 1 257
- 1 258
- 1 259
- 1 260
- 1 261
- 1 262
- 1 263
- 1 264
- 1 265
- 1 266
- 1 267
- 1 268
- 1 269
- 1 270
- 1 271
- 1 272
- 1 273
- 1 274
- 1 275
- 1 276
- 1 277
- 1 278
- 1 279
- 1 280
- 1 281
- 1 282
- 1 283
- 1 284
- 1 285
- 1 286
- 1 287
- 1 288
- 1 289
- 1 290
- 1 291
- 1 292
- 1 293
- 1 294
- 1 295
- 1 296
- 1 297
- 1 298
- 1 299
- 1 300
- 1 301
- 1 302
- 1 303
- 1 304
- 1 305
- 1 306
- 1 307
- 1 308
- 1 309
- 1 310
- 1 311
- 1 312
- 1 313
- 1 314
- 1 315
- 1 316
- 1 317
- 1 318
- 1 319
- 1 320
- 1 321
- 1 322
- 1 323
- 1 324
- 1 325
- 1 326
- 1 327
- 1 328
- 1 329
- 1 330
- 1 331
- 1 332
- 1 333
- 1 334
- 1 335
- 1 336
- 1 337
- 1 338
- 1 339
- 1 340
- 1 341
- 1 342
- 1 343
- 1 344
- 1 345
- 1 346
- 1 347
- 1 348
- 1 349
- 1 350
- 1 351
- 1 352
- 1 353
- 1 354
- 1 355
- 1 356
- 1 357
- 1 358</

Keterangan:

- Dari hasil analisis tersebut terlihat bahwa *aitem-b* = (1) outfit mean square (MNSQ) sebesar -2,66, (2) outfit Z standard (ZSTD) sebesar +4,5, (3) Pt-Mean Corr. Sebesar -0,38 dan *aitem-a* = outfit Z Standard (ZSTD) sebesar -2,4. Masuk ke dalam kategori *outlier* atau *misfit* dengan data yang ada, aitem harus direvisi atau dibuang.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

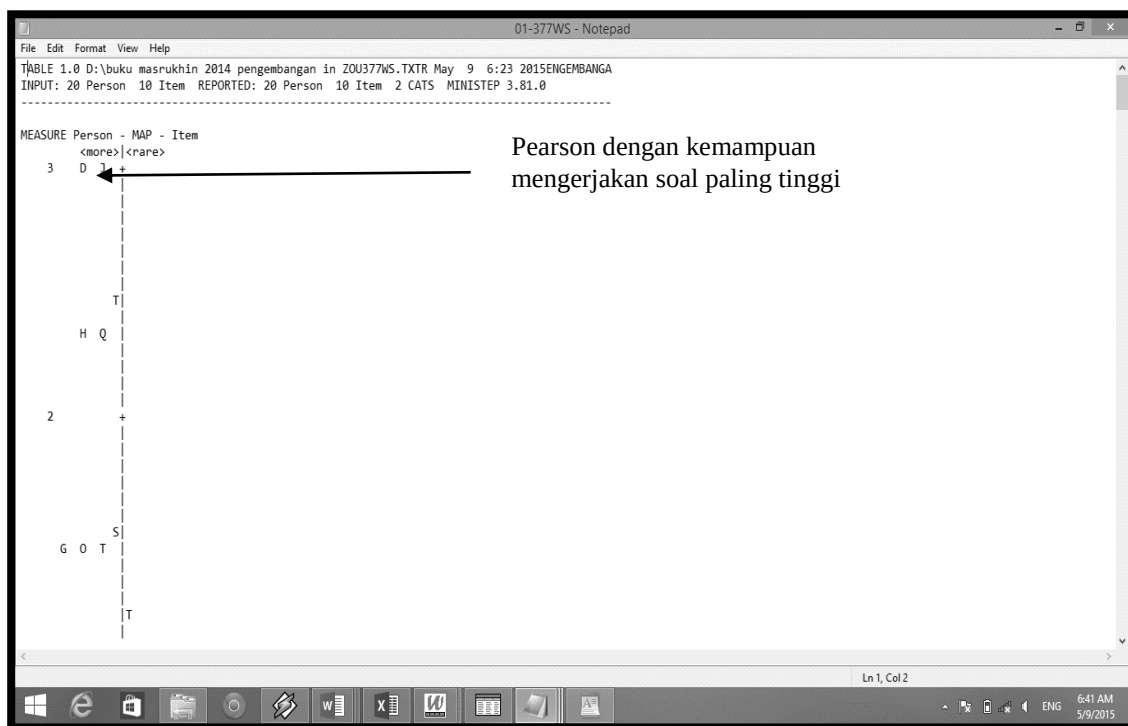
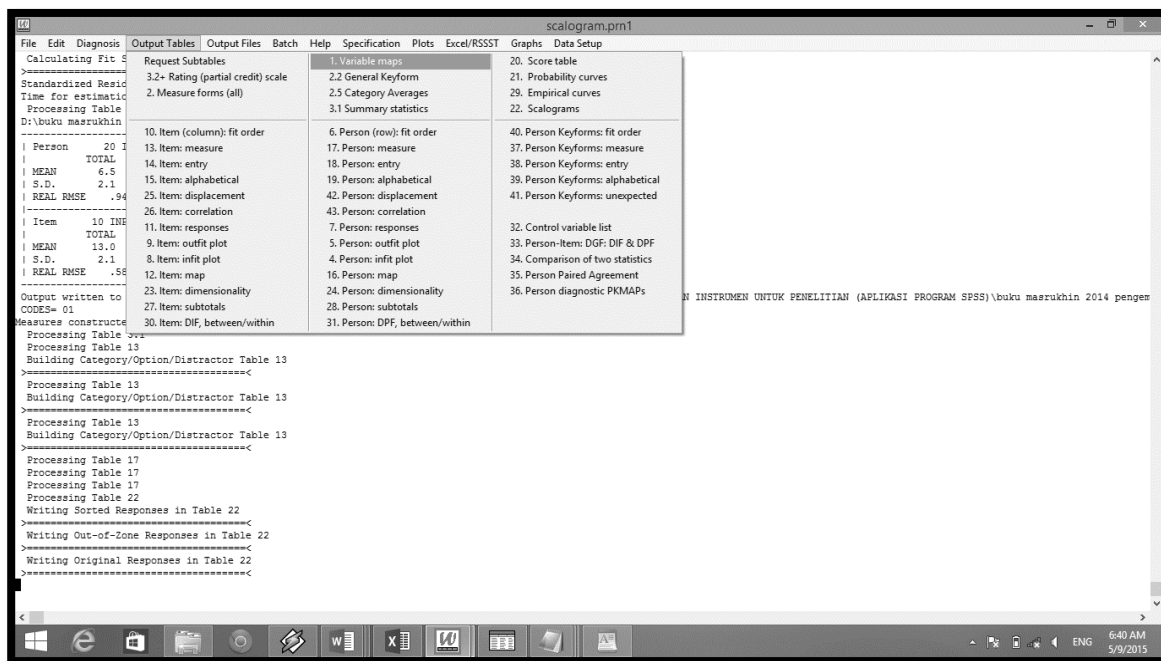


Analisis dari table *person measure* mendapati bahwa alasan person D dan J sesuai (*fit*), sedangkan K dan L tidak fit adalah dia tidak mampu mengerjakan soal yang mudah dan akan tetapi ia mampu mengerjakan soal sulit. Analisis pemodelan Rasch mampu meramalkan dan mendeteksi kemampuan responden dengan lebih tepat.

Tabel 1: variable Maps

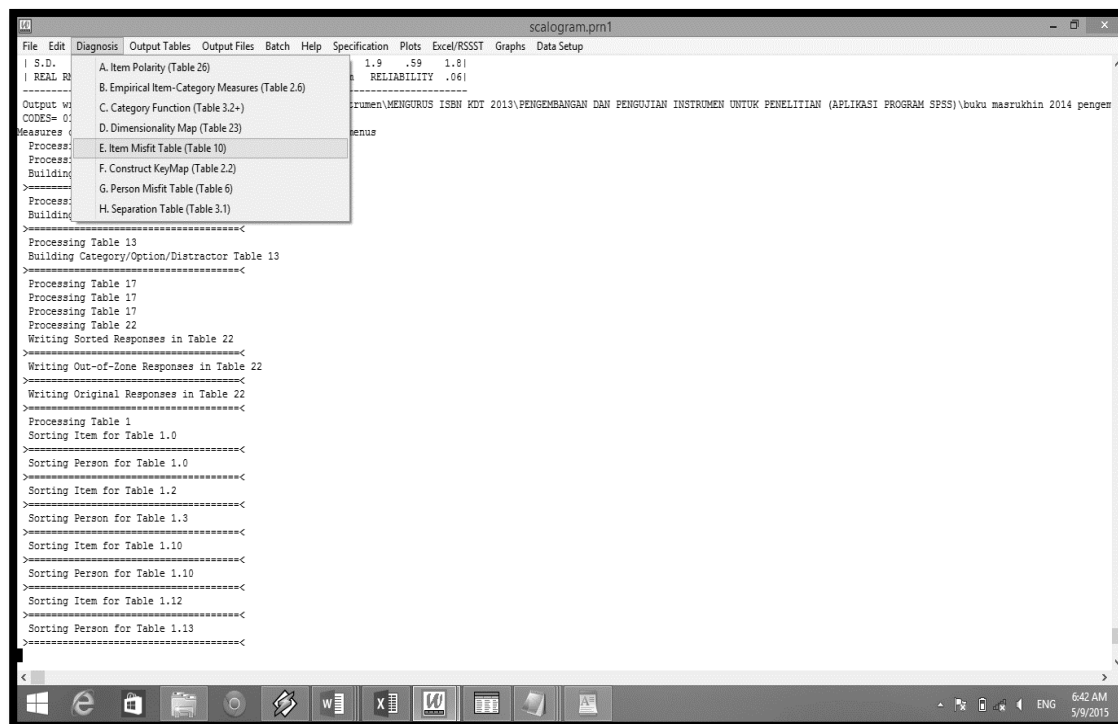
Tabel berikut menggambarkan person ((bagian kiri) dan item (sebelah kanan)

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Cara lain untuk memeriksa adanya ketidaksesuaian (*misfit*), baik untuk aitem maupun person, adalah dengan menggunakan menu **Diagnosis** dari menu utama.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



10-377WS - Notepad

File Edit Format View Help

TABLE 10.1 D:\buku masrukhin 2014 pengembangan i ZOU377WS.TXTU May 9 6:23 2015PENGEMBANGA
INPUT: 20 Person 10 Item REPORTED: 20 Person 10 Item 2 CATS MINISTEP 3.81.0

Person: REAL SEP.: .91 REL.: .45 ... Item: REAL SEP.: .26 REL.: .06

Item STATISTICS: MISFIT ORDER

ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MEASURE	MODEL S.E.	INFIT MNSQ	ZSTD	OUTFIT MNSQ	ZSTD	PTMEASURE-A CORR.	EXP.	OBS%	EXP%	Item
2	12	20	.31	.52	2.24	4.9	2.66	4.5	A-.38	.47	16.7	67.9	b
3	10	20	.84	.52	1.12	.6	1.07	.3	B-.45	.51	66.7	68.4	c
6	15	20	-.55	.57	1.07	.4	.93	.0	C-.36	.38	61.1	73.9	f
7	12	20	.31	.52	.99	.0	.90	-.3	D-.49	.47	61.1	67.9	g
8	16	20	-.89	.60	.91	-.2	.71	-.4	E-.43	.35	72.2	77.8	h
10	16	20	-.89	.60	.90	-.2	.70	-.4	e-.44	.35	83.3	77.8	j
5	13	20	.03	.53	.87	-.6	.79	-.7	d-.54	.44	77.8	68.6	e
4	14	20	-.25	.54	.72	-1.3	.64	-1.0	c-.60	.42	88.9	71.0	d
9	12	20	.31	.52	.64	-2.1	.58	-1.8	b-.71	.47	83.3	67.9	i
1	10	19	.80	.53	.56	-2.6	.52	-2.4	a-.79	.50	100.0	67.4	a
MEAN	13.0	19.9	.00	.54	1.00	-1.1	.95	-.2			71.1	70.9	
S.D.	2.1	.3	.60	.03	.45	1.9	.59	1.8			21.6	3.9	

Item yang paling misfit
berdasarkan kriteria outfit
MNSQ, outfit ZSTD dan Pt-
Measure Corr.

Item yang paling fit berdasarkan
kriteria outfit MNSQ, outfit
ZSTD dan Pt-Measure Corr.

TABLE 10.3 D:\buku masrukhin 2014 pengembangan i ZOU377WS.TXTU May 9 6:23 2015PENGEMBANGA
INPUT: 20 Person 10 Item REPORTED: 20 Person 10 Item 2 CATS MINISTEP 3.81.0

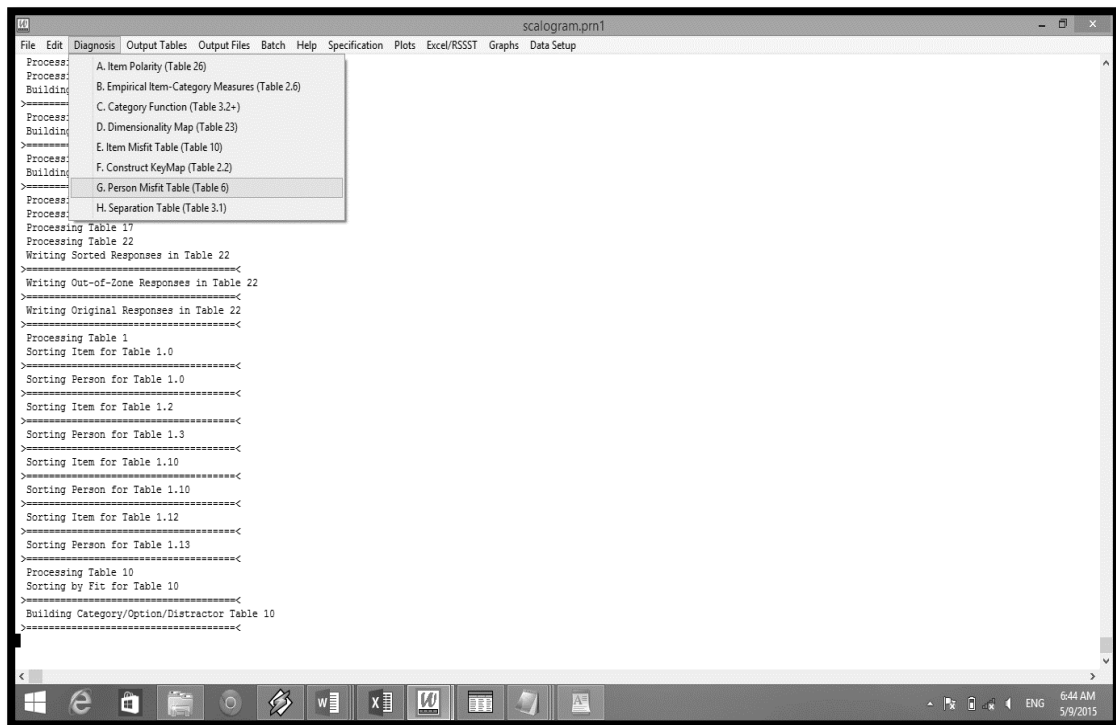
Item CATEGORY/OPTION/DISTRACTOR FREQUENCIES: MISFIT ORDER

ENTRY NUMBER	DATA CODE	SCORE VALUE	DATA COUNT	%	AVERAGE ABILITY	S.E. MEAN	OUTF MNSQ	PTMA CORR.	Item

Seperti terlihat pada table, aitem yang tidak sesuai (*misfit*) diurutkan dari atas ke bawah, dan item nomor 2(b) menunjukkan aitem yang paling tidak *fit*. Satu aitem a yang ditampilkan menunjukkan sangat sesuai (*fit*) dengan pola yang ada.

Yang kedua dengan memilih **Person Misfit (tabel 6)**

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



06-377WS - Notepad

File Edit Format View Help

TABLE 6.1 D:\buku masrukhin 2014 pengembangan in ZOU377WS.TXT May 9 6:23 2015ENGEMBANGA
INPUT: 20 Person 10 Item REPORTED: 20 Person 10 Item 2 CATS MINISTEP 3.81.0

Person: REAL SEP.: .91 REL.: .45 ... Item: REAL SEP.: .26 REL.: .06

Person STATISTICS: MISFIT ORDER

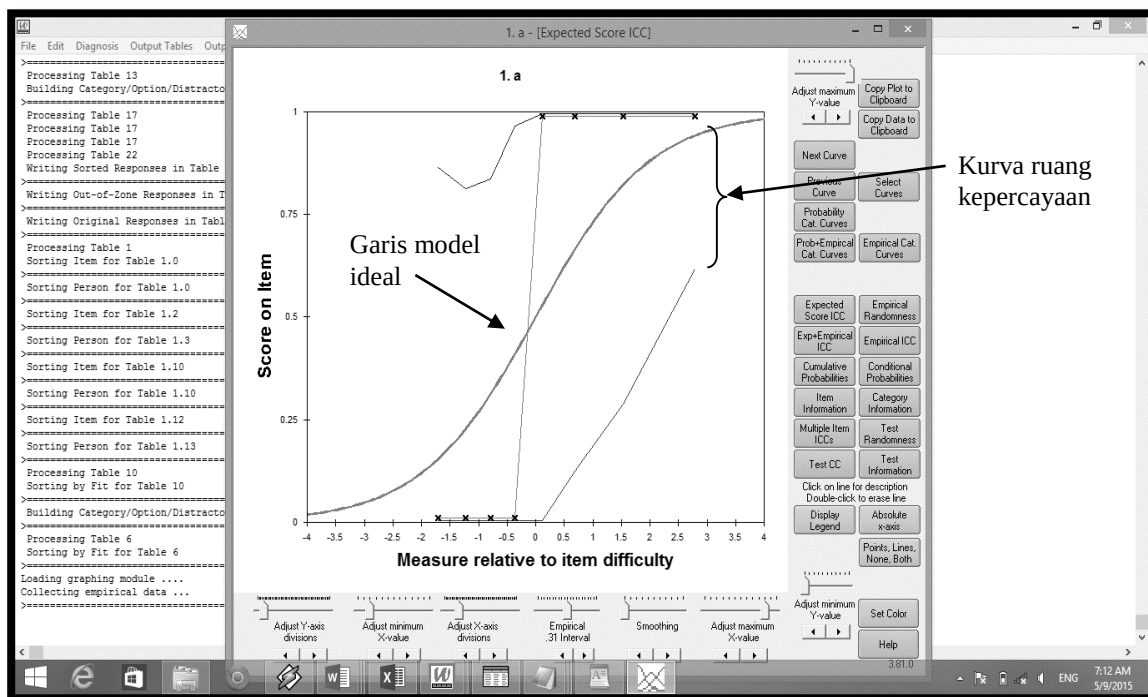
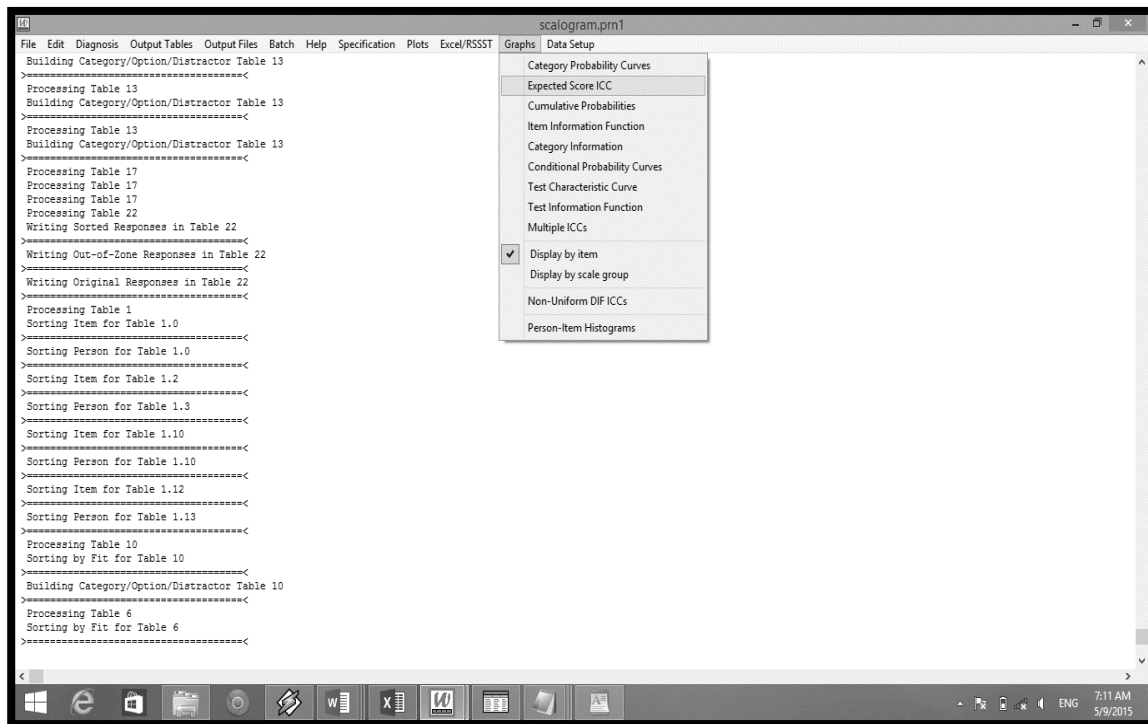
ENTRY NUMBER	TOTAL SCORE	TOTAL COUNT	MODEL MEASURE	S.E.	INFINIT MNSQ	OUTFIT ZSTD	PTMEASURE-A CORR.	EXP.	OBSV EXP%	EXACT MATCH	Person	
12	5	10	.00	.66	1.55	2.7	1.58	2.6	A-.60	.28	10.0 62.4	L
11	6	10	.44	.67	1.28	1.3	1.42	1.6	B-.23	.28	60.0 63.7	K
20	4	10	-.44	.67	1.29	1.3	1.34	1.3	C-.19	.28	50.0 65.1	U
2	7	10	.92	.71	1.27	.9	1.31	.9	D-.18	.25	70.0 70.0	B
13	4	10	-.44	.67	1.14	.7	1.10	.5	E .08	.28	50.0 65.1	M
5	7	10	.92	.71	1.04	.2	.93	.0	F .24	.25	70.0 70.0	E
14	7	10	.92	.71	1.04	.2	.93	.0	G .24	.25	70.0 70.0	N
8	9	10	2.33	1.07	1.02	.3	.86	.2	H .17	.16	90.0 90.0	H
17	9	10	2.33	1.07	1.02	.3	.86	.2	I .17	.16	90.0 90.0	Q
1	3	9	-.83	.73	.94	-.1	.92	-.1	J .36	.26	66.7 67.3	A
6	3	10	-.91	.72	.92	-.2	.87	-.2	K .42	.27	70.0 70.0	F
3	6	10	.44	.67	.88	-.6	.87	-.5	L .47	.28	80.0 63.7	C
7	8	10	1.49	.81	.87	-.1	.70	-.4	M .48	.22	80.0 80.0	G
15	8	10	1.49	.81	.87	-.1	.70	-.4	N .48	.22	80.0 80.0	O
19	8	10	1.49	.81	.87	-.1	.70	-.4	P .48	.22	80.0 80.0	T
9	5	10	.00	.66	.70	-1.8	.68	-1.8	Q .76	.28	90.0 62.4	I
16	5	10	.00	.66	.70	-1.8	.68	-1.8	R .76	.28	90.0 62.4	P
18	6	10	.44	.67	.70	-1.6	.66	-1.5	a .77	.28	80.0 63.7	R
MEAN	6.5	9.9	.89	.86	1.00	.1	.95	.0			70.9 70.9	
S.D.	2.1	.2	1.28	.34	.22	1.1	.28	1.1			19.1 8.9	

Item yang paling misfit
berdasarkan kriteria outfit
MNSQ, outfit ZSTD dan Pt-
Measure Corr.

TABLE 6.4 D:\buku masrukhin 2014 pengembangan in ZOU377WS.TXT May 9 6:23 2015ENGEMBANGA
INPUT: 20 Person 10 Item REPORTED: 20 Person 10 Item 2 CATS MINISTEP 3.81.0

6:44 AM
5/9/2015

Pilihan lain untuk melihat ketidaksesuaian (*misfit*) aitem adalah dengan menggunakan grafik pada menu **Graphs** dan pilih **Expected Score ICC**.



Respons dari aitem tertentu (dalam grafik di atas aitem a) di tampilan bersamaan dengan garis ideal regresi logistik (berwarna merah) dan kurva ruang kepercayaan (warna biru sisi atas dan bawah), Respons yang *misfit* akan berada diluar kurva ruang kepercayaan.

e) Penjelasan tentang Infit, Outfit, Mean-square, dan Standardized

Dalam pemodelan Rasch diproses seberapa akurat data sesuai dengan model ideal, baik itu dalam bentuk data dikotomi atau politomi, ini tidak lain bentuk dari statistic kesesuaian.

Infit berarti *inlier-sensitive* atau *information-weightedfit*, tidak lain adalah kesensitifan pola respons terhadap aitem sasaran pada responden (*person*), atau sebaliknya. Kondisi *infit* memang sulit diperiksa karena memang sesuai dengan model yang ada.

Outfit berarti *outlier-sensitive fit*, maksudnya adalah mengukur kesensitifan pola respons terhadap aitem dengan tingkat kesulitan tertentu pada responden (*person*), atau sebaliknya. Contoh yang mudah adalah respons yang salah dari person, seperti tidak bisa mengerjakan soal yang mudah padahal mempunyai kemampuan tinggi (*careless*); ataupun bisa mengerjakan soal yang susah dengan kemampuan rendah (*lucky guess*). Keduanya relatif mudah dideteksi.

Mean-square fit statistic, memperlihatkan ukuran keacakan yaitu jumlah distorsi dalam sistem pengukuran. Nilai yang diharapkan adalah antara 0,5-1,5; Jika nilai kurang dari nilai tersebut mengindikasikan hal itu terlalu mudah ditebak (*data overfit the model*); sedangkan nilai yang lebih besar mengindikasikan tidak mudah diprediksi (*data underfit the model*). Secara statistik *mean-square* adalah nilai statistic chi-kuadrat yang dibagi dengan derajat bebas; dan nilainya selalu positif.

Nilai Mean-square	Implikasi pada pengukuran
>2,0	Menurunkan kualitas sistem pengukuran
1,5-2,0	Kurang bagus untuk pembuatan instrumen, tetapi tidak menurunkan kualitas
0,5-1,5	Kondisi yang baik untuk pengukuran
<0,5	Kurang produktif untuk pengukuran, namun tidak menurunkan kualitas; kemungkinan bias menyebabkan kesalahan dengan reliabilitas yang tinggi

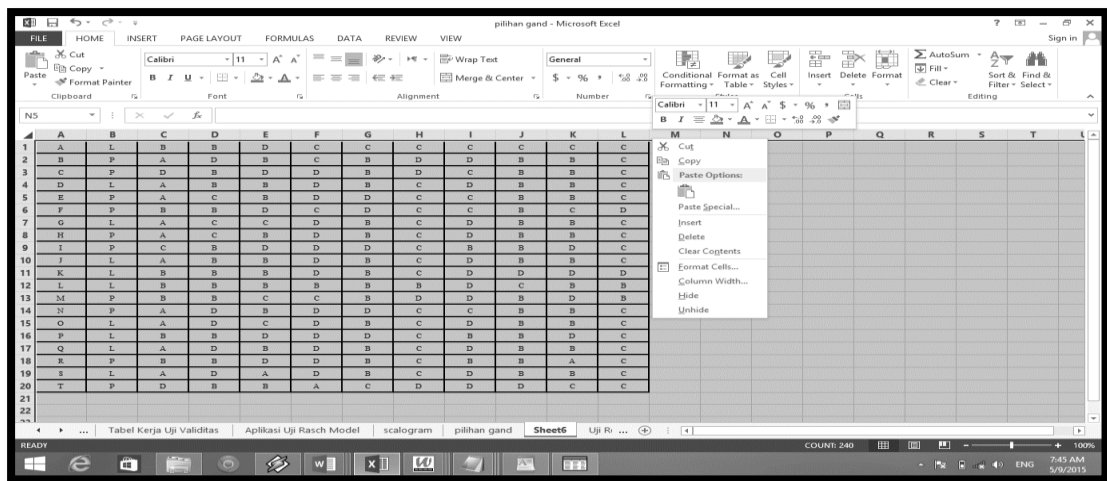
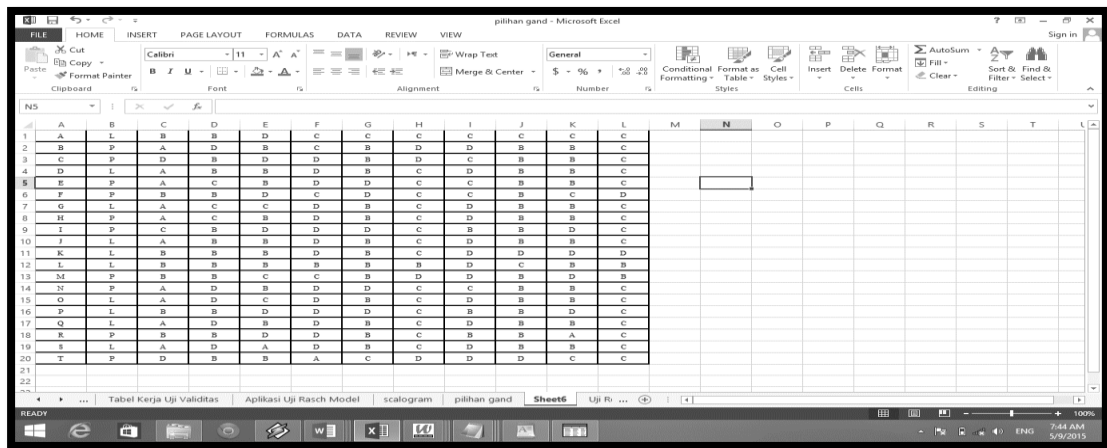
Standardized fit statistic (ZSTD), adalah uji-t untuk hipotesis, “apakah data sesuai (*fit*) dengan model?” Hasilnya adalah nilai-z yaitu penyimpangan unit. Ini menjelaskan ketidakmungkinan dari data, yaitu signifikasinya jika data memang sesuai dengan model. Tabel berikut menjelaskan nilai Z.

Nilai ZSTD	Implikasi pada pengukuran
$\geq 3,0$	Data tidak diharapkan jika sesuai dengan model (secara sempurna) Namun, dengan sampel yang besar, ketidaksesuaiannya mungkin lebih kecil
2,0-2,9	Data tampak tidak dapat diprediksi
-1,9-1,9	Data mempunyai perkiraan yang logis
≤ -2	Data terlalu mudah diprediksi

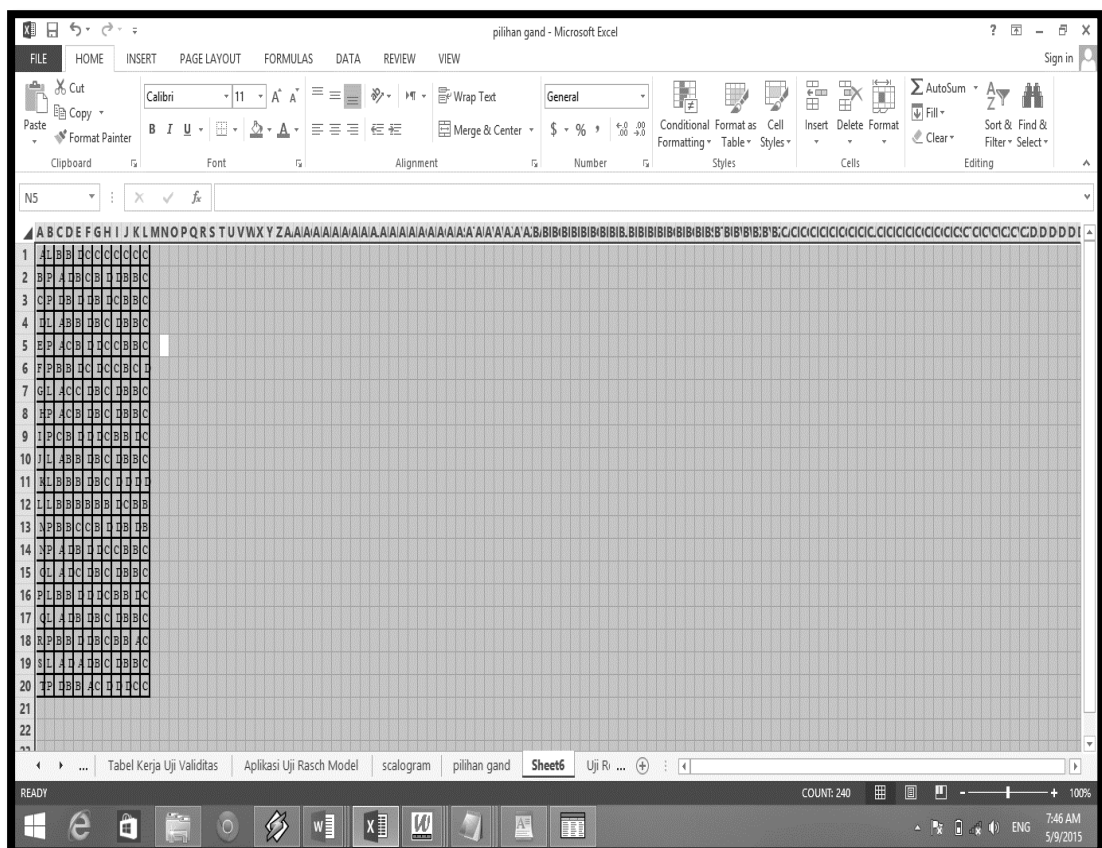
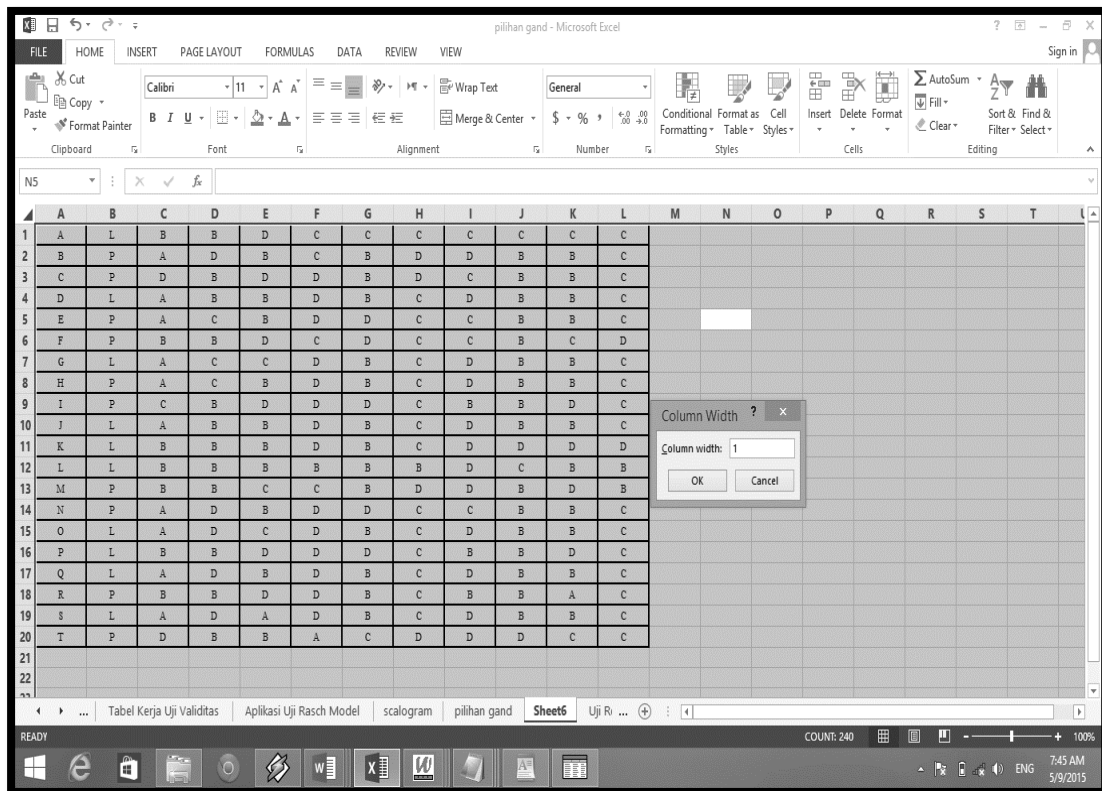
f) Penyiapan Berkas Data Pilihan Ganda (Multiple Choices)

- Siapkan berkas data *.prn dalam Microsoft Excel:

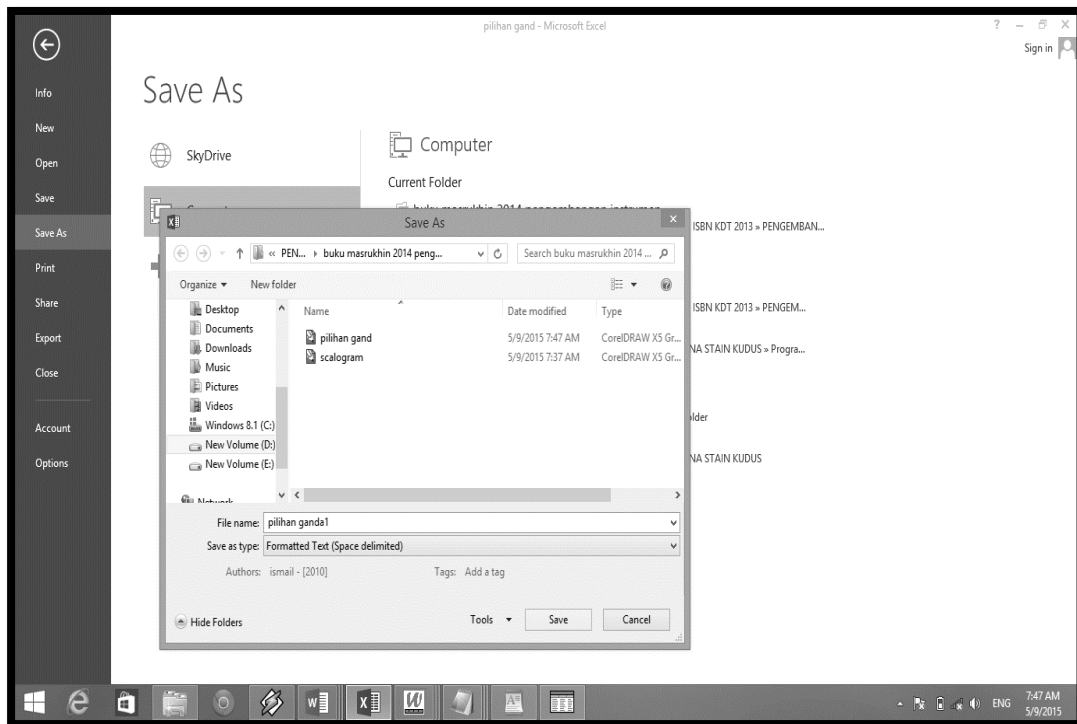
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



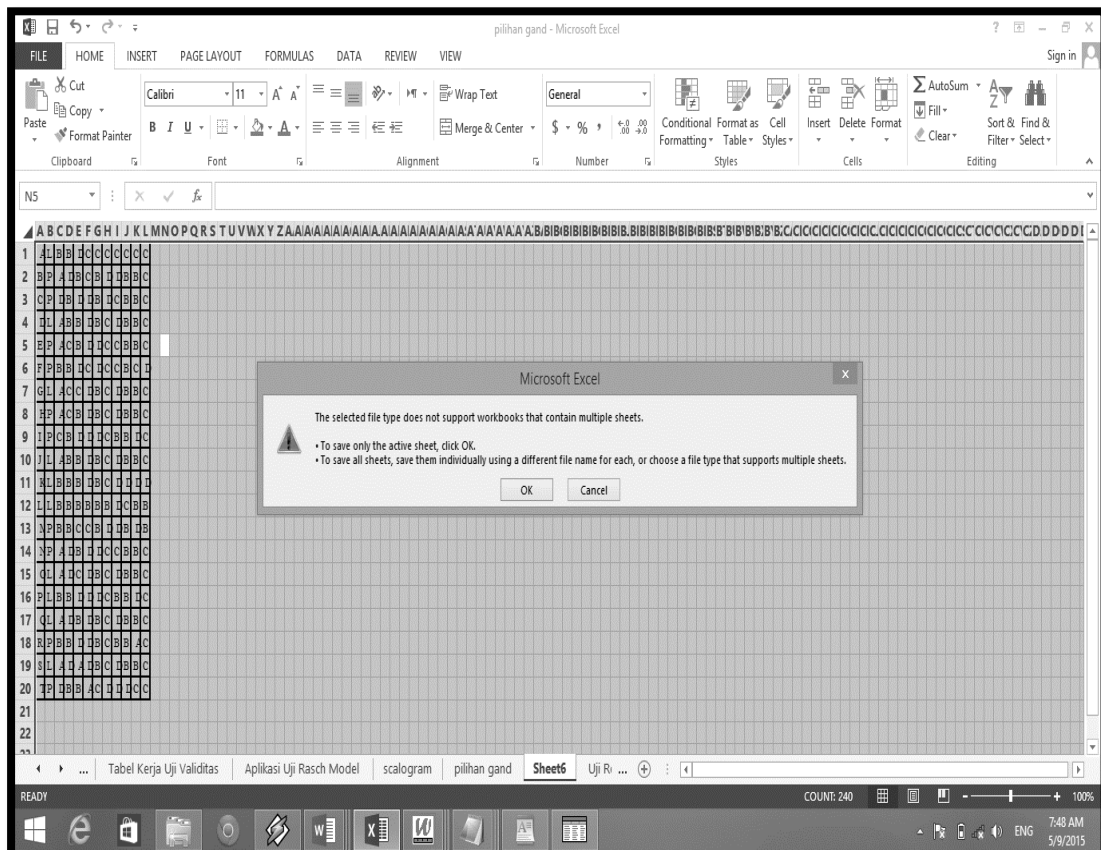
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

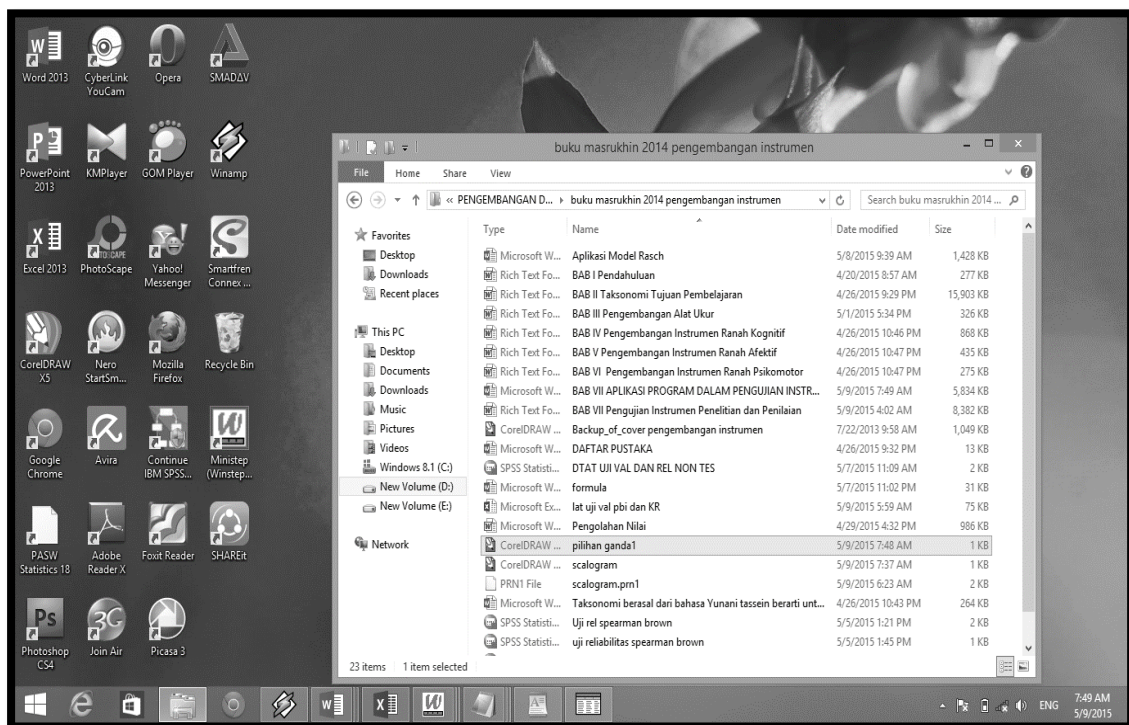
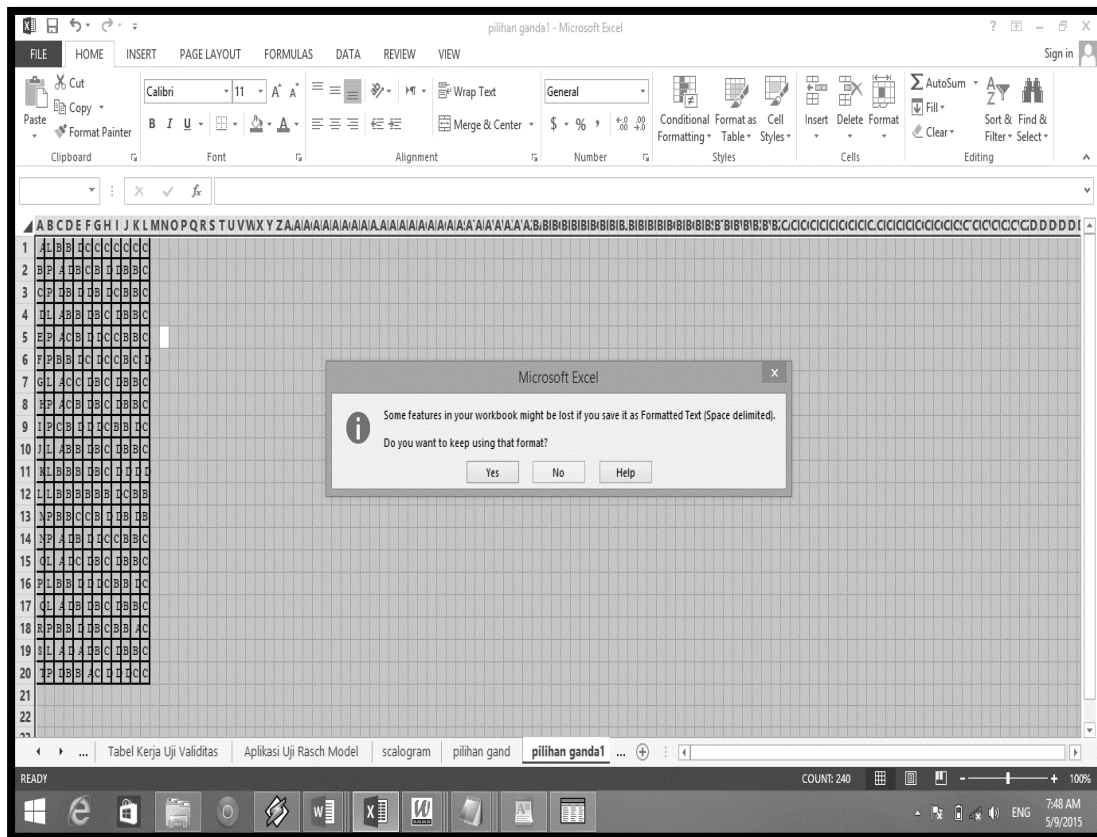


Klik Ok

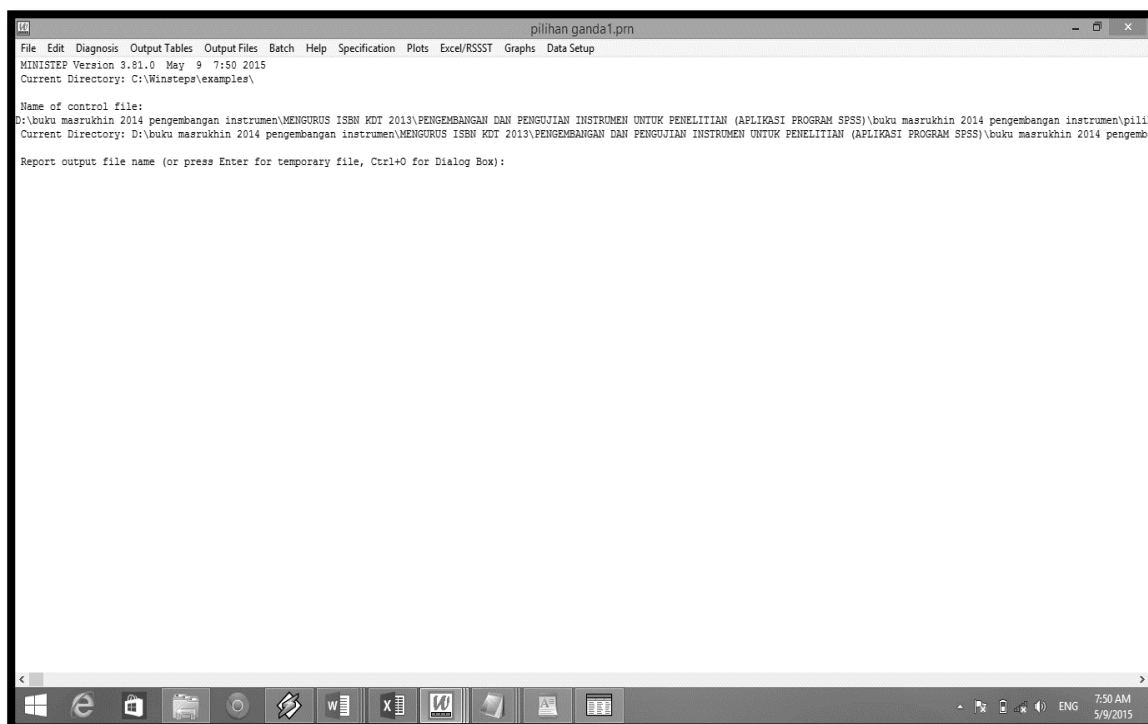
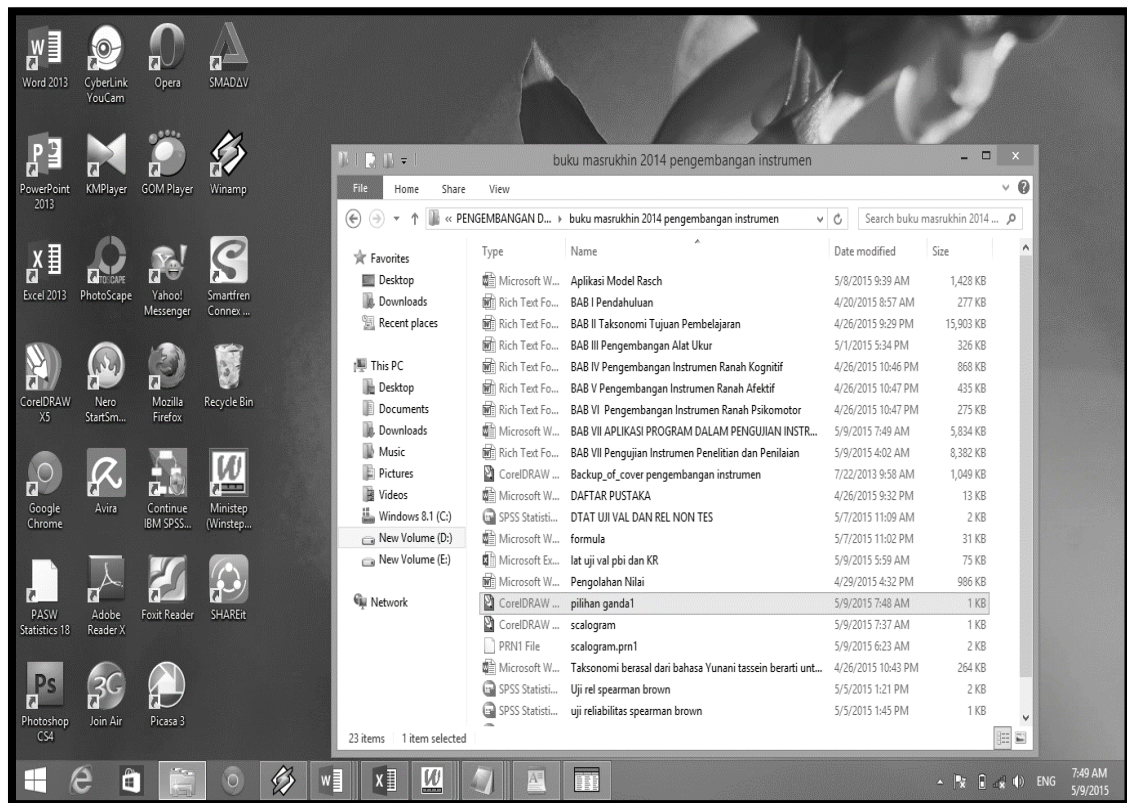


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

Klik Yes



b. Geser berkas (File) **pilihganda.prn** ke dalam program Ministep.



c. Spesifikasi berkas data untuk Ministep.

First person column = 1 (mulai dari kolom pertama)

Person label length = 2 (dua kolom pertama identitas responden)

First item column = diisi dengan angka 3 (kolom pertama data dimulai)

Number of Items = diisi dengan angka 10 (banyaknya aitem dalam berkas *scalogram*)

The screenshot shows the 'Ministep Control File Set-Up' window. The 'PERSON' section has 'PERSON' as the data row, '1' for NAME1, and '1' for NAMELEN. The 'ITEM' section has 'Item' as the data column, '1' for ITEM1, and '1' for NI. The 'Number of data rows' is 20, and 'Number of data columns' is 12. The 'Data entry direction' is set to 'Person'. The 'Other specifications in control file' section is empty. The 'Item Labels' section shows 'Enter/Edit' and '3.68.0'.

The screenshot shows the 'Ministep Control File Set-Up' window with updated settings. The 'PERSON' section remains the same. The 'ITEM' section now has '3' for ITEM1 and '10' for NI. The 'Number of data rows' is 20, and 'Number of data columns' is 12. The 'Data entry direction' is set to 'Person'. The 'Other specifications in control file' section is empty. The 'Item Labels' section shows 'Enter/Edit' and '3.68.0'.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

The screenshot displays the 'Ministep Control File Set-Up' application. The main window has several input fields and buttons:

- Title:** Report title is []
- PERSON=** A data row is a [Person]
- NAME1=** First person label column [1]
- NAMLEN=** Person label length [1]
- Number of data rows** [20]
- Number of data columns** [12]
- ITEM=** A data column is a [Item]
- ITEM1=** First item column [3]
- NI=** Number of Items [10]
- XWIDE=** columns per response [1]
- CODES=** Valid codes []
- UIMEAN=** Set item mean []
- UPMEAN=** Set person mean []
- UIMEAN=** Mean of items [0]
- USCALE=** Units per logit [1]
- UDECIM=** decimal places [2]
- Buttons:** Refresh Data Distribution, Data entry direction (arrows), Remove excess rows and columns, Scan data for codes, Category labels CLFILE=, MCQ Scoring KEY1+, Rating scales & partial credit ISGROUPS=, 3680.

A secondary window titled 'Item Labels: Enter/Edit' is open, showing a table with columns 'Number' and 'Label':

Number	Label
1	p1
2	p2
3	p3
4	p4
5	p5
6	p6
7	p7
8	p8
9	p9
10	p10

The background shows a Windows taskbar with icons for Start, Internet Explorer, File Explorer, and other applications. The system clock indicates 7:55 AM on 5/9/2015.

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a Person

NAME= First person label column

NAMELEN= Person label length

Number of data rows

Number of data columns

ITEM= A data column is a Item

ITEM1= First item column

NI= Number of Items

XWIDE= columns per response

CODES= Valid codes

UIMEAN= Set item mean

UPMEAN= Set person mean

UITEM= Mean of items

USCALE= Units per logit

UDECIM= decimal places

Item Labels:
Enter/Edit

Refresh Data Display

Data entry direction:

Remove excess rows and columns

Scan data for codes

Category labels CLFILE=

MCQ Scoring KFV1=

Rating scales & partial credit ISGROUPS=

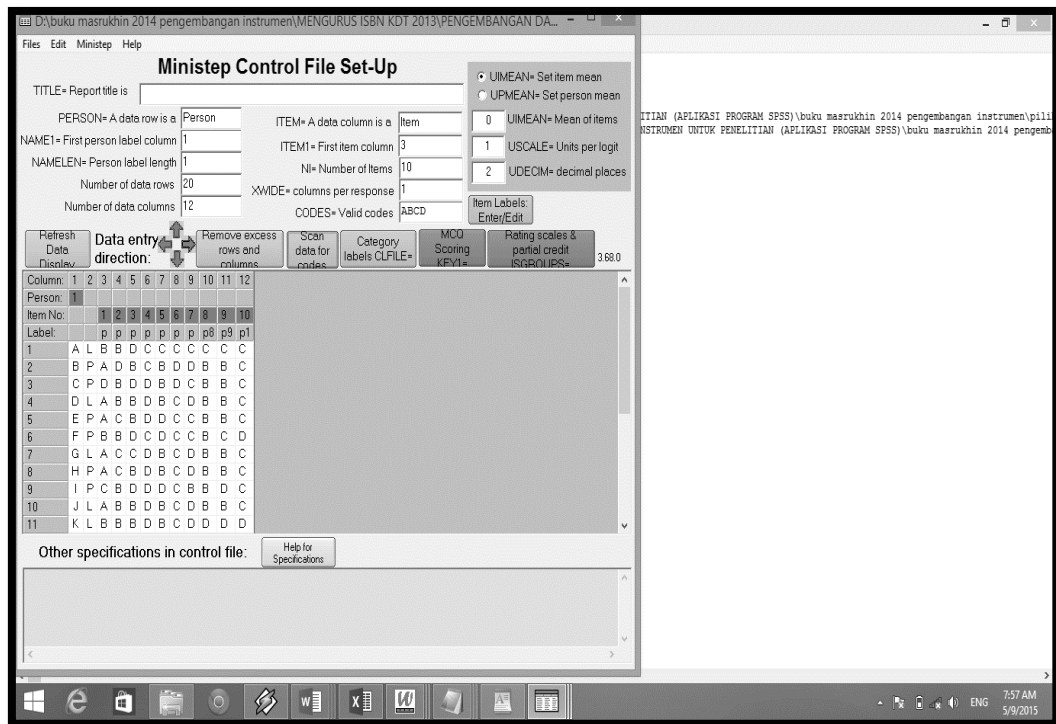
368.0

Column:	1	2	3	4	5	6	7	8	9	10	11	12
Person:	1											
Item No:	1	2	3	4	5	6	7	8	9	10		
Label:		p	p	p	p	p	p	p	p8	p9	p1	
1	A	L	B	B	D	C	C	C	C	C	C	C
2	B	P	A	D	B	C	B	D	B	B	B	C
3	C	P	D	B	D	D	B	D	C	B	B	C
4	D	L	A	B	B	D	B	C	D	B	B	C
5	E	P	A	C	B	D	D	C	C	B	B	C
6	F	P	B	B	D	C	D	C	B	C	B	D
7	G	L	A	C	D	B	B	C	D	B	B	C
8	H	P	A	C	B	D	B	B	C	D	B	B
9	I	P	C	B	D	D	C	B	B	D	D	C
10	J	L	A	B	B	D	B	C	D	B	B	C
11	K	L	B	B	B	D	B	C	D	D	D	D

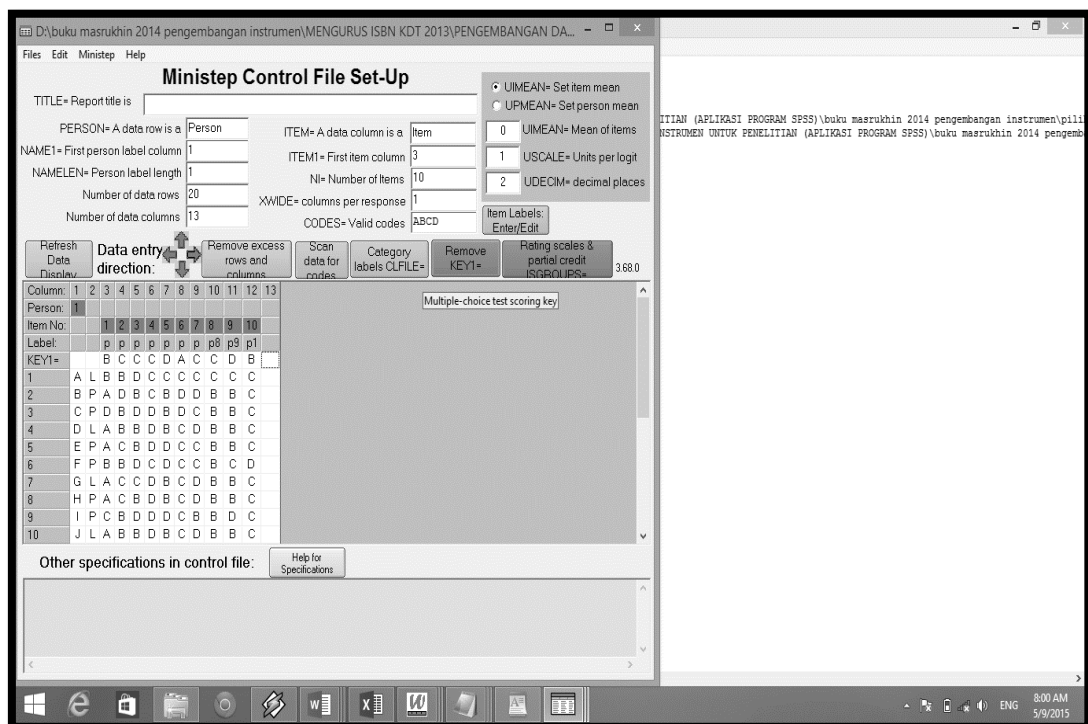
Other specifications in control file:

Help for Specifications

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

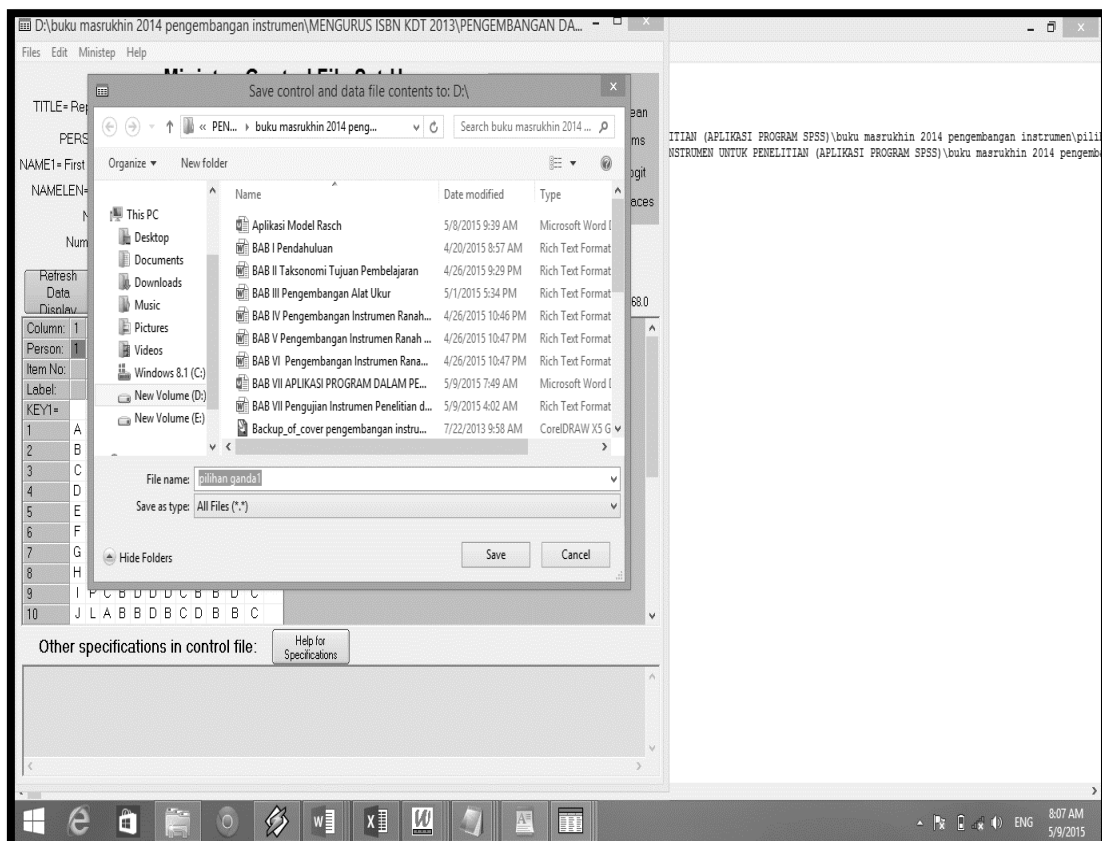
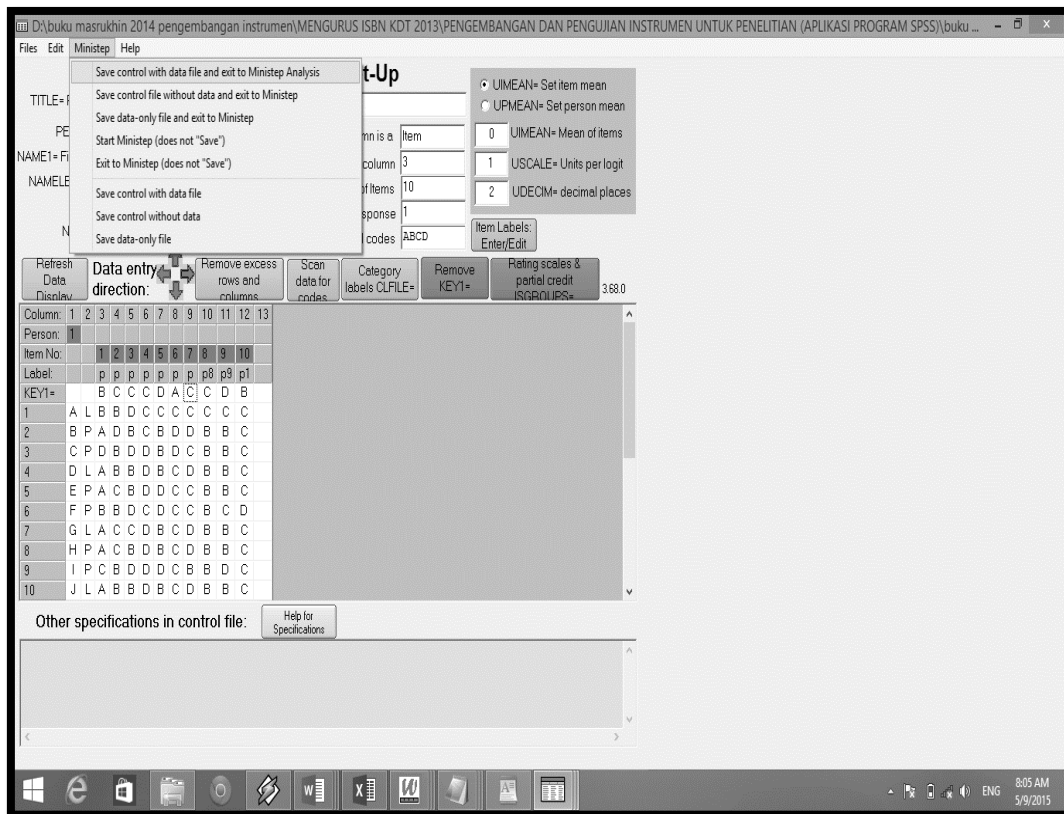


d. Langkah berikutnya adalah mengisi kunci jawaban pilihan ganda tersebut:



e. Apabila sudah lengkap diisi, klik tombol **Ministep** pada menu tersebut, dan simpan (*save*) berkas dengan memberinya nama (misalnya, **pilihan.prn**).

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



D:\buku masruhin 2014 pengembangan instrumen\MENGURUS ISBN KDT 2013\PENGEMBANGAN DA...

Files Edit Ministep Help

Ministep Control File Set-Up

TITLE= Report title is

PERSON= A data row is a

NAME1= First person label column

NAMELEN= Person label length

Number of data rows

Number of data columns

Person

1

1

20

13

ITEM= A data column is a

ITEM1= First item column

N= Number of items

XWIDE= columns per response

CODES= Valid codes

Item

3

10

1

ABCD

UIMEAN= Set item mean

UFMEAN= Set person mean

0

1

2

USCALE= Units per logit

UDECM= decimal places

Item Labels

Enter/Edit

Rating scales & partial credit

IGROUPS= 368.0

Refresh Data Display

Data entry direction:

Remove excess rows and columns

Scan data for codes

Category labels GLFILE=

Remove KEY1=

Winsteps Control File Set-Up

Column: 1 2 3 4 5 6 7 8

Person: 1

Item No: 1

Label

KEY1=

1 A L B B D C C D A

2 B P A D B C B D

3 C P D B D D C B D

4 D L A B B D B C B D

5 E P A C B D D C C B B C

6 F P B C D C D C C B C D

7 G L A C C D B C D B B C

8 H P A C B D B C D B B C

9 I P C B D D D C B B D C

10 J L A B B D C D B B C

Other specifications in control file:

Help for Specifications

Saved control and data file to: D:\buku masruhin 2014 pengembangan instrumen\MENGURUS ISBN KDT 2013\PENGEMBANGAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN (APLIKASI

Yes No Help

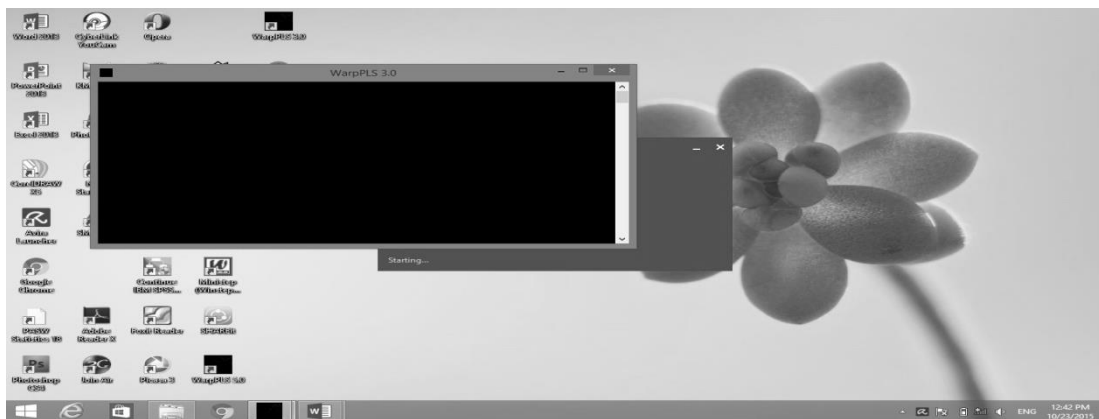
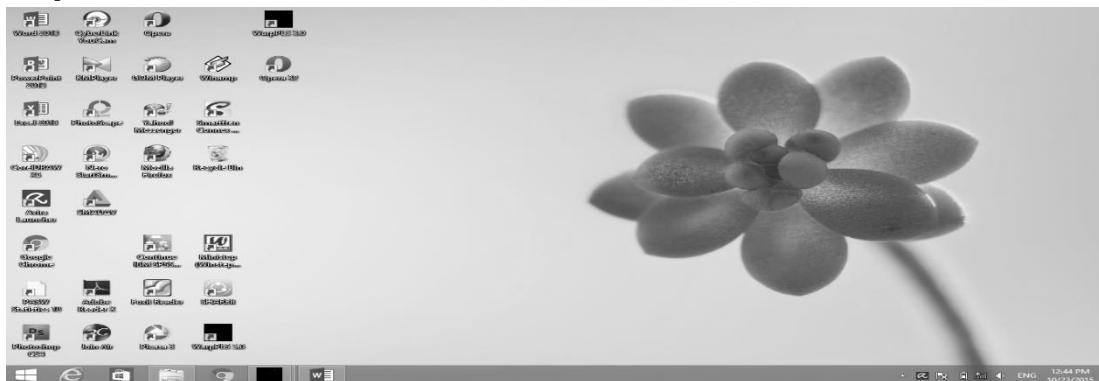
8:09 AM 5/9/2015

- 2.pilihan.prm
- File Edit Diagnosis Output Tables Output Files Batch Help Specification Plots Excel/RSSST Graphs Data Setup
- Control: \pilihan.prm Output: \20U25SWS.TXT
- | PROX ACTIVE COUNT EXTREME 5 RANGE MAX LOGIT CHANGE |
- | ITERATION Person Item CATS Person Item MEASURES STRUCTURE |
- >=====<
- | 1 20 10 2 1.70 .73 -2.1972 |
- >=====<
- | 2 15 9 2 1.84 .89 .2412 |
- Probing data connection: to skip out: Ctrl+F - to bypass: subset=no
- Processing unanchored persons ...
- >=====<
- Consolidating 2 potential subsets pairwise ...
- >=====<
- Data fully connected. No subsets found
- (Control: \pilihan.prm Output: \20U25SWS.TXT
- | JMLE MAX SCORE MAX LOGIT LEAST CONVERGED CATEGORY STRUCTURE |
- | ITERATION RESIDUAL* CHANGE Person Item CAT RESIDUAL CHANGE |
- >=====<
- | 1 -.20 -.3844 4* 6 |
- >=====<
- | 2 -.05 -.1178 4* 6 |
-
- Calculating Fit Statistics
- >=====<
- Standardized Residuals N(0,1) Mean: .00 S.D.: 1.00
- Time for estimation: 0:0:0.600
- Processing Table 0
- D:\buku masrukhin 2014 pengembangan instrumen\MENGURUS ISBN KOT 2013\PEMENGEMBAN
- | Person 20 INPUT 20 MEASURED INFIT OUTFIT |
- | TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD |
- | MEAN 1.8 10.0 -1.82 1.13 1.00 .1 1.00 .1 |
- | S.D. 1.8 .0 1.19 .43 .18 .7 .34 .7 |
- | REAL RMSE 1.21 TRUE SD .00 SEPARATION .00 Person RELIABILITY .00 |
- |-----|
- | Item 10 INPUT 10 MEASURED INFIT OUTFIT |
- | TOTAL COUNT MEASURE REALSE IMNSQ ZSTD OMNSQ ZSTD |
- | MEAN 3.5 20.0 .29 .79 1.00 .1 1.00 .1 |
- | S.D. 1.9 .0 1.04 .35 .15 .5 .33 .6 |
- | REAL RMSE .86 TRUE SD .58 SEPARATION .67 Item RELIABILITY .31 |
- |-----|
- Output written to D:\buku masrukhin 2014 pengembangan instrumen\MENGURUS ISBN KOT 2013\PEMENGEMBAN DAN PENGUJIAN INSTRUMEN UNTUK PENELITIAN (APLIKASI PROGRAM SPSS)\buku masrukhin 2014 pengembangan-38CD
- Measures constructed: use "Diagnosis" and "Output Tables" menus
- < >
- Windows taskbar: 8:10 AM 5/9/2015

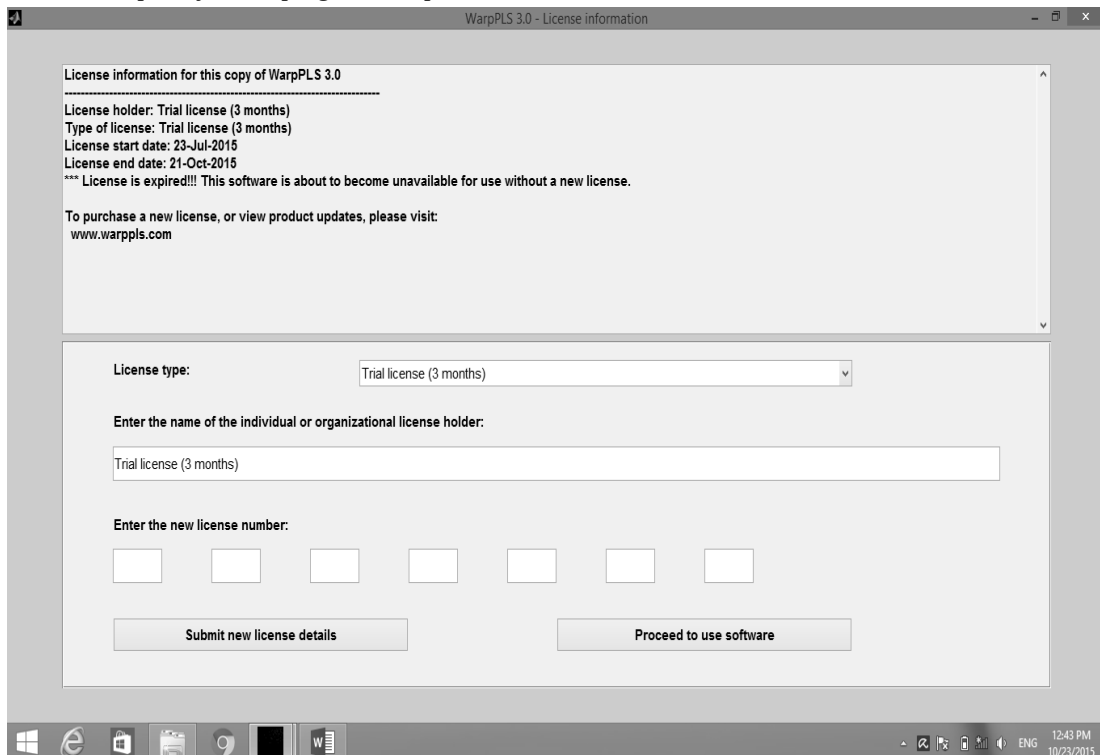
[176]

D. LANGKAH-LANGKAH MENJALANKAN PROGRAM WarpPLS 3.0

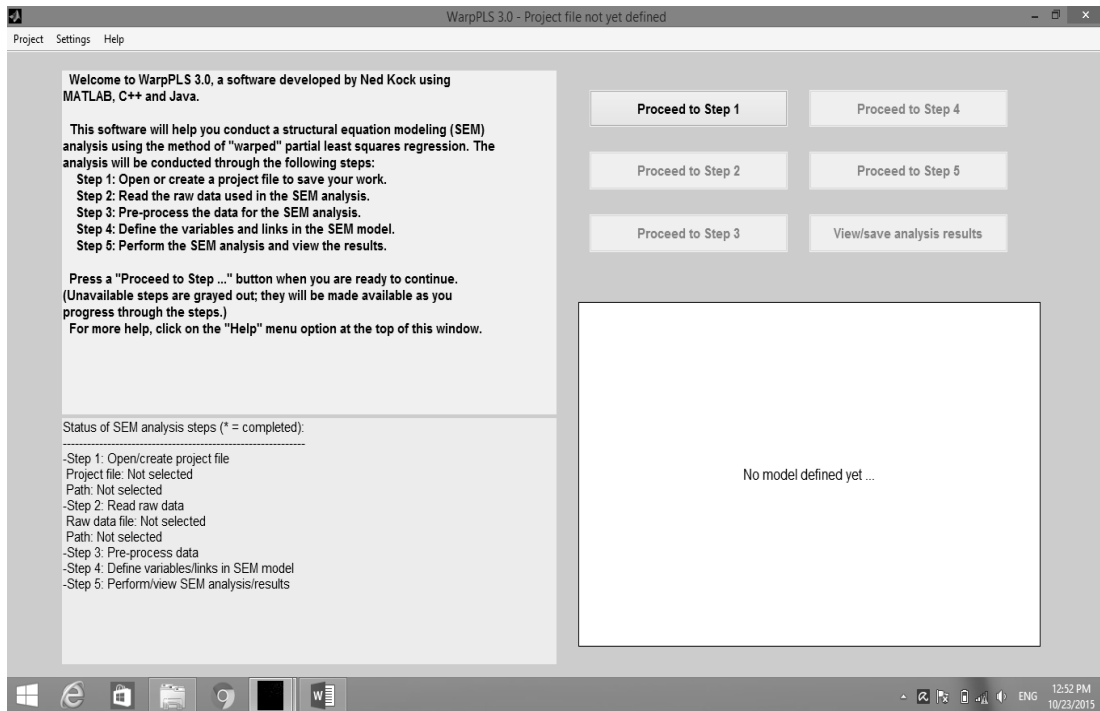
- a. Jalankan program VPLS dengan mengklik *shortcut* Program WarpPLS 3.0 atau 5.0 pada desktop layar komputer.



- b. Muncul tampilan jendela program WarpPLS 3.0



c. Klik *Proceed to use software*

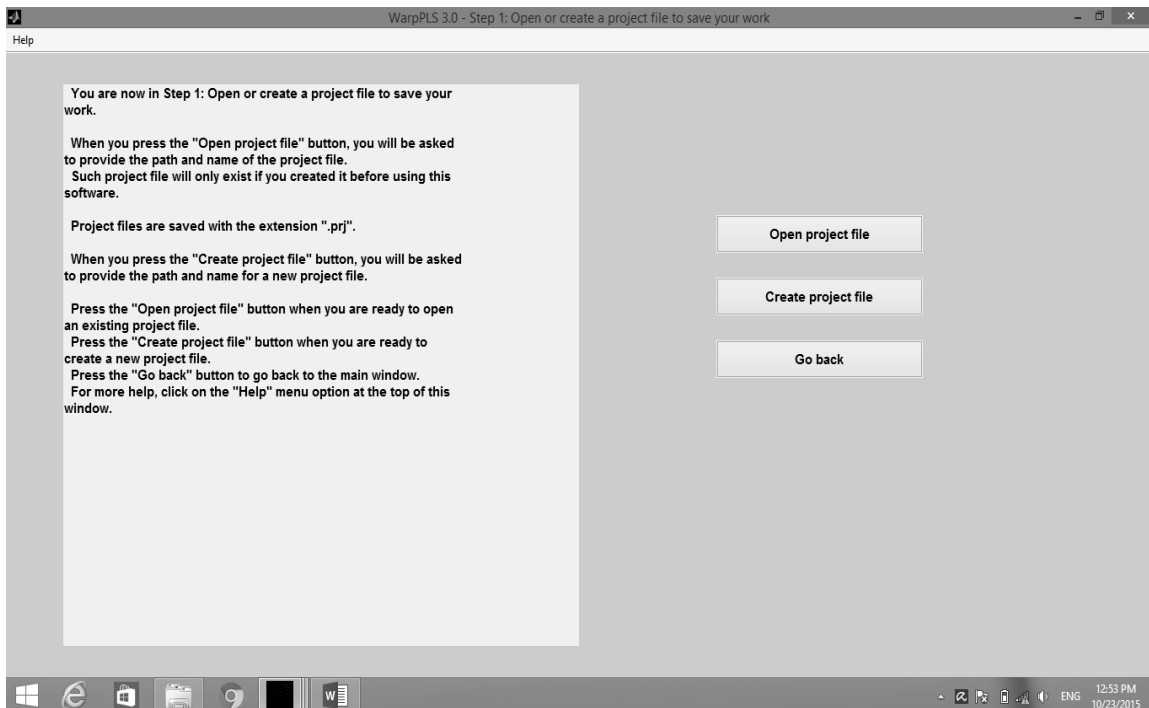


d. WarpPLS akan melakukan analisis melalui lima langkah yang terstruktur (*Proceed to Step 1-proceed to step 5*).

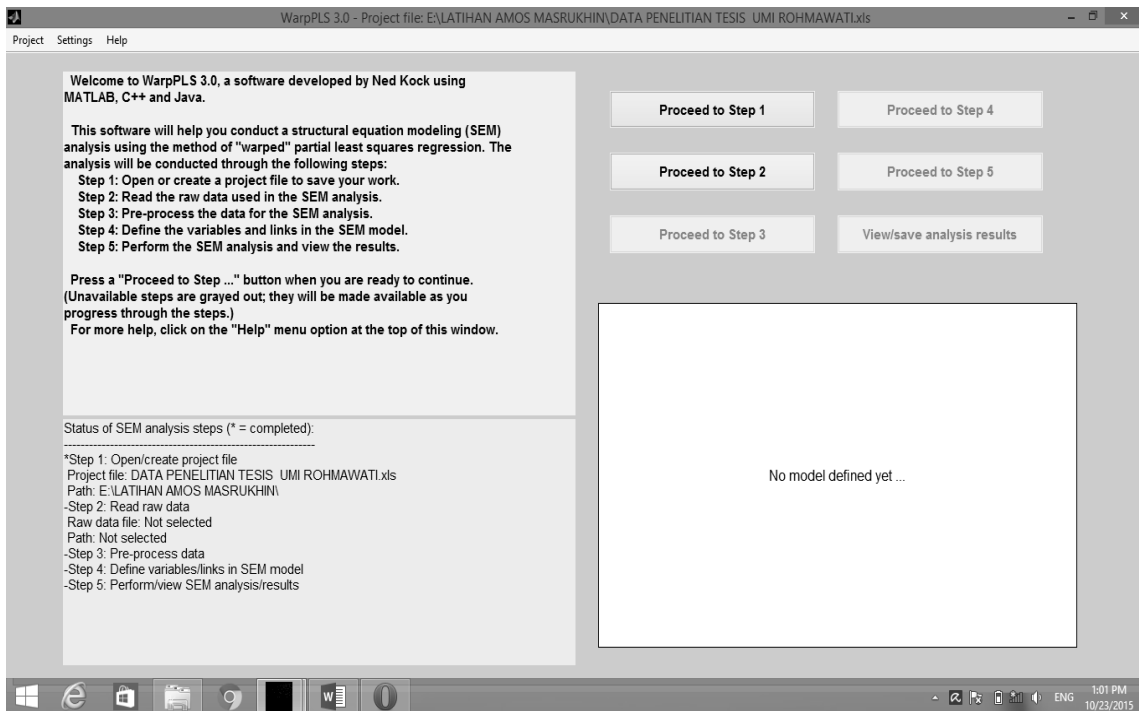
Langkah 1: Membuka atau membuat sebuah file proyek

Setelah *proceed to step 1* diklik, maka akan muncul :

Buatlah proyek atau pekerjaan baru dengan mengklik *Create Project File* dan beri nama file dan simpan di folder komputer anda dengan ekstensi “.prj”.

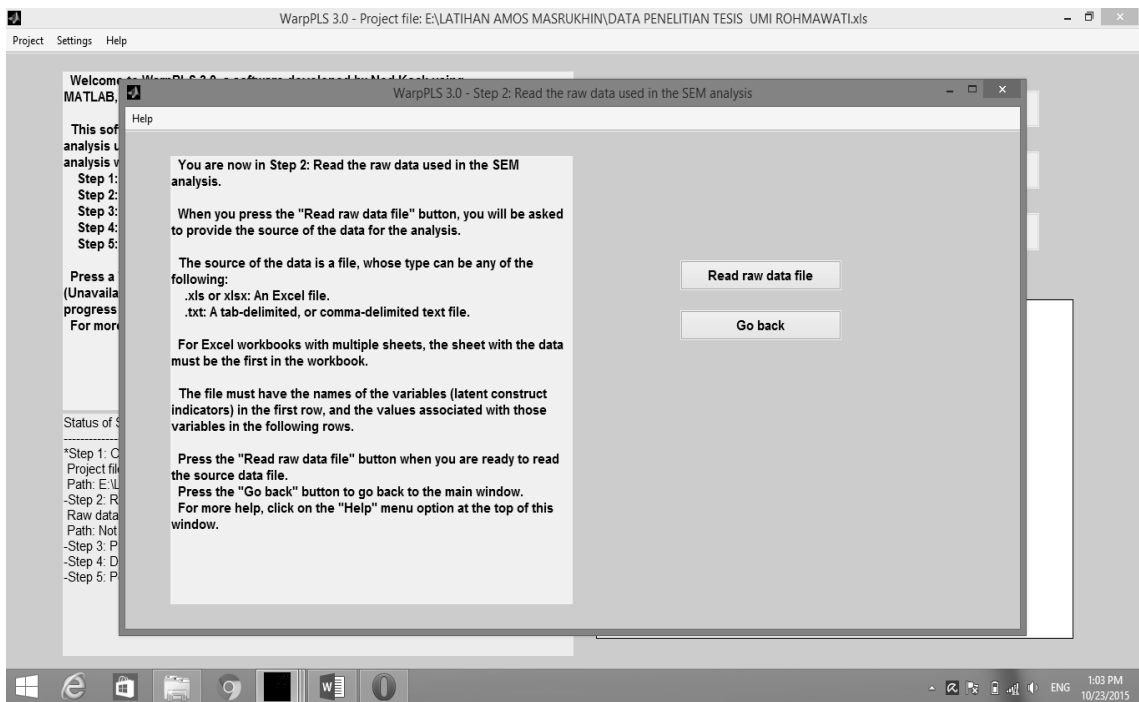


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Langkah 2 : Membaca data mentah yang akan digunakan dalam analisis SEM

Dalam langkah ini, WarpPLS akan membaca data mentah yang akan digunakan. Data mentah dapat dibaca langsung dari file excel (.xls)/ excel workbook atau notepad (.txt). Jika menggunakan data excel, kita harus memastikan bahwa file excel hanya mempunyai satu *worksheet* (yang berisi data). Jika kita telah mengklik *Proceed to Step 2* maka akan muncul:

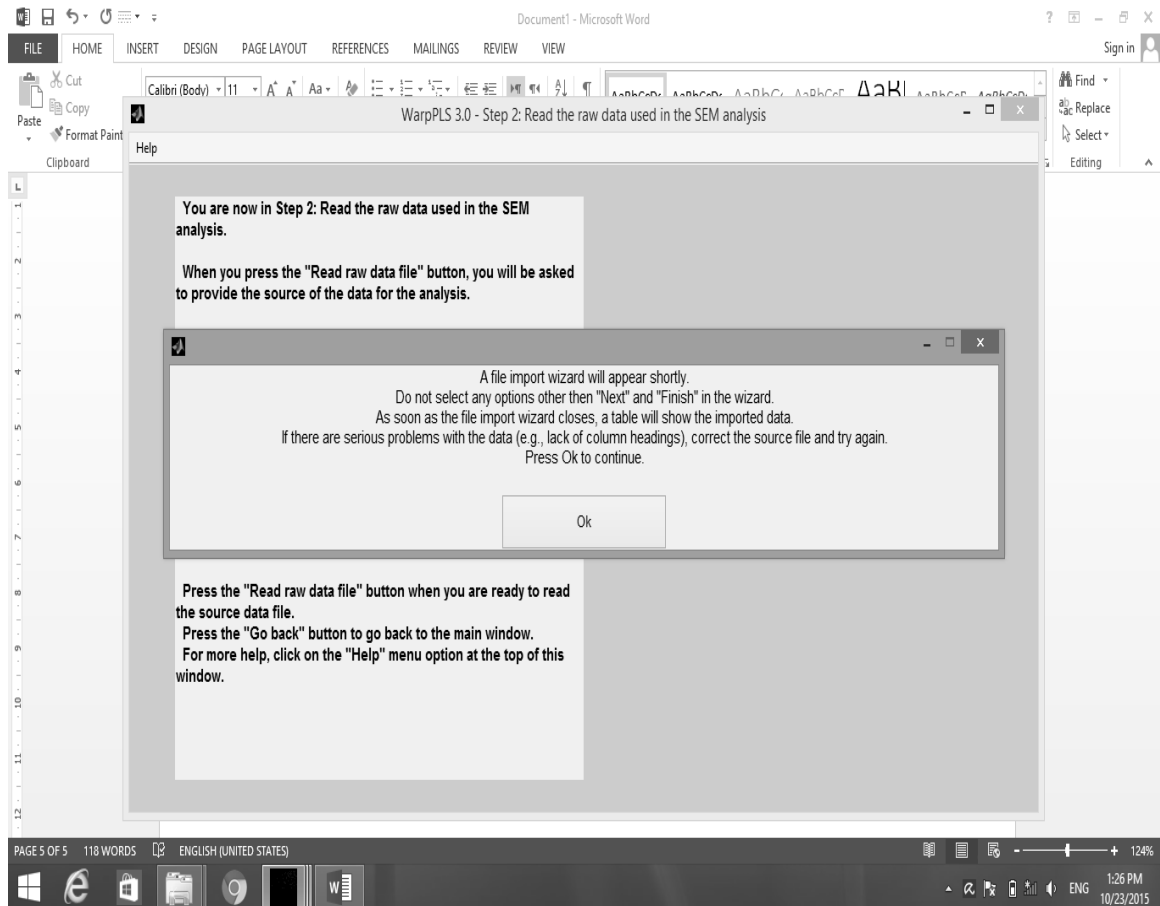


Klik *read raw data file*. Data yang akan digunakan Data WarpPLS.txt atau All files

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

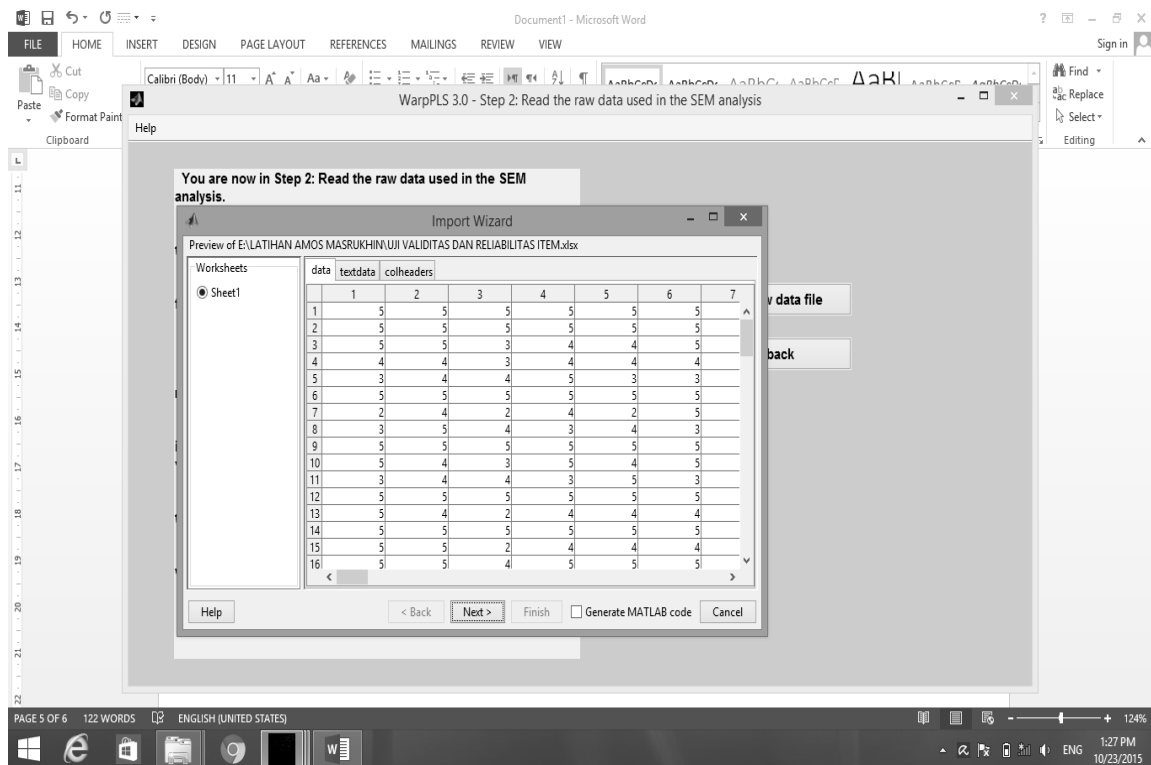


Klik open dan OK.

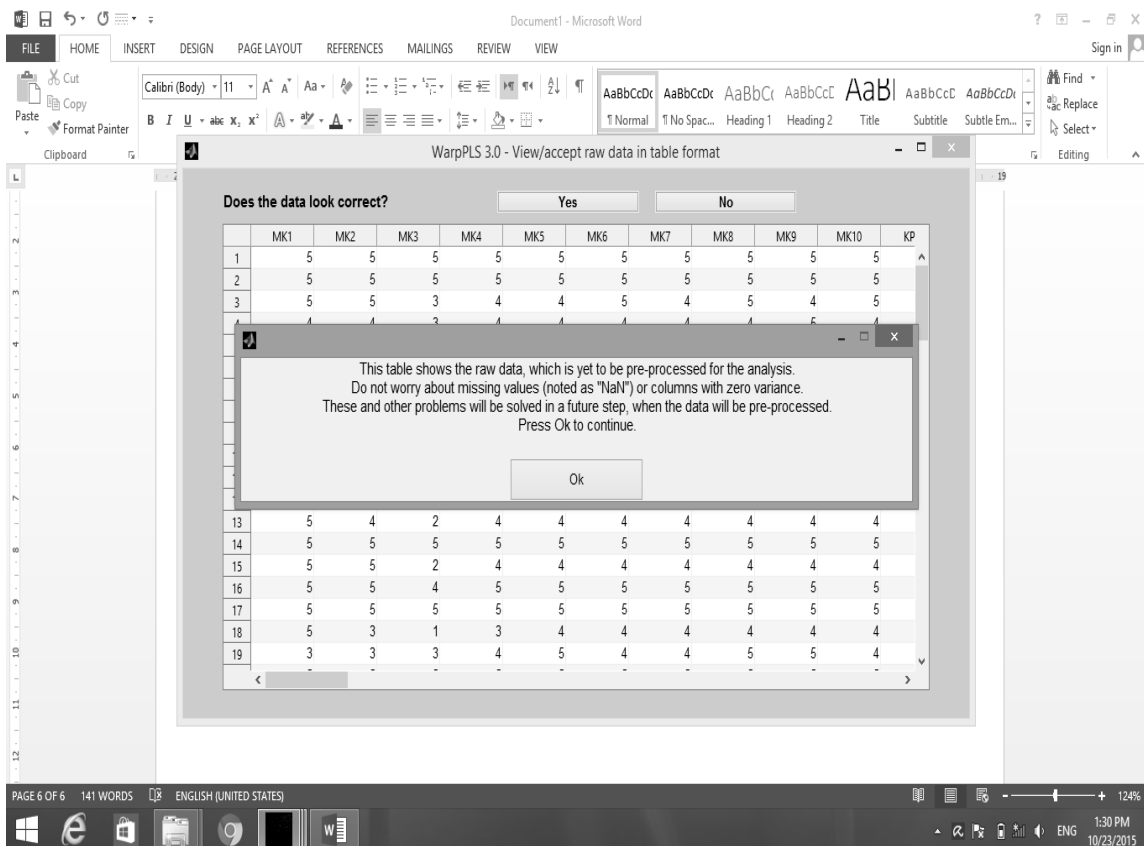


Klik open, maka muncul dengan judul *import Wizard*.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

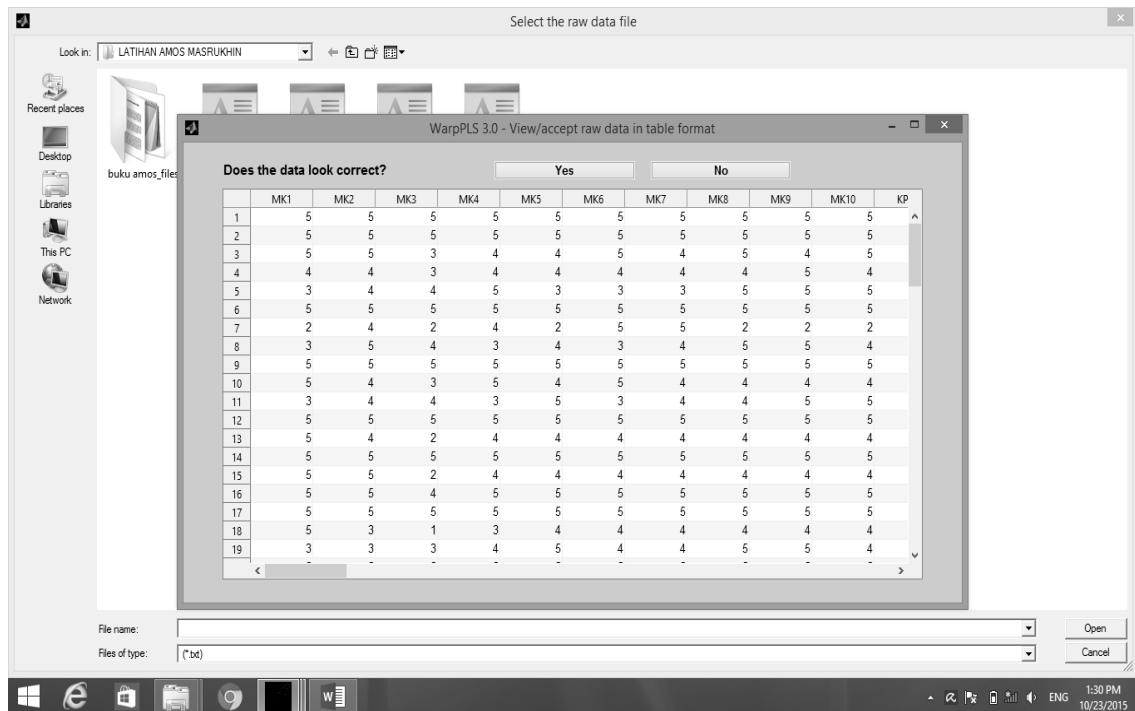


Selanjutnya, klik next, lalu finish sehingga muncul jendela seperti dibawah ini.



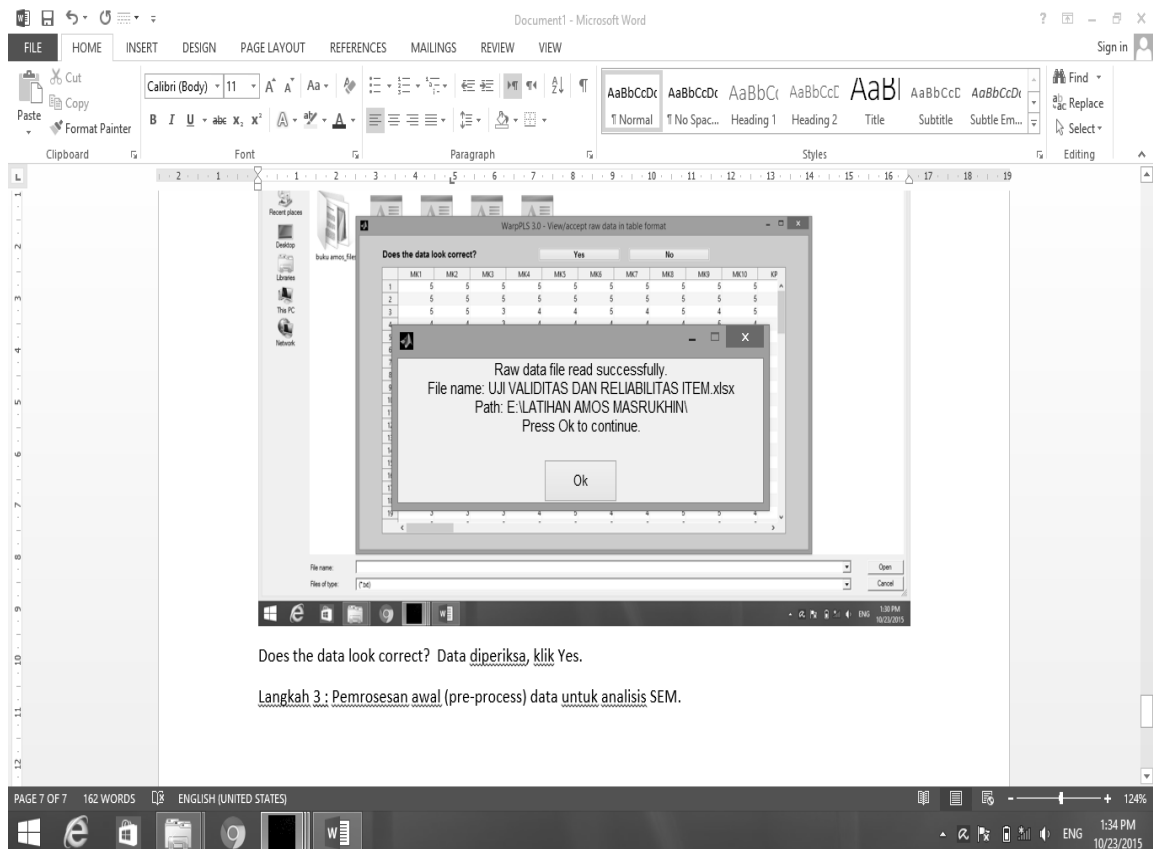
Klik Ok.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Terdapat pertanyaan *Does the data look correct?* Data diperiksa, jika yakin benar maka klik Yes.

Langkah 3 : Pemrosesan awal (pre-process) data untuk analisis SEM.

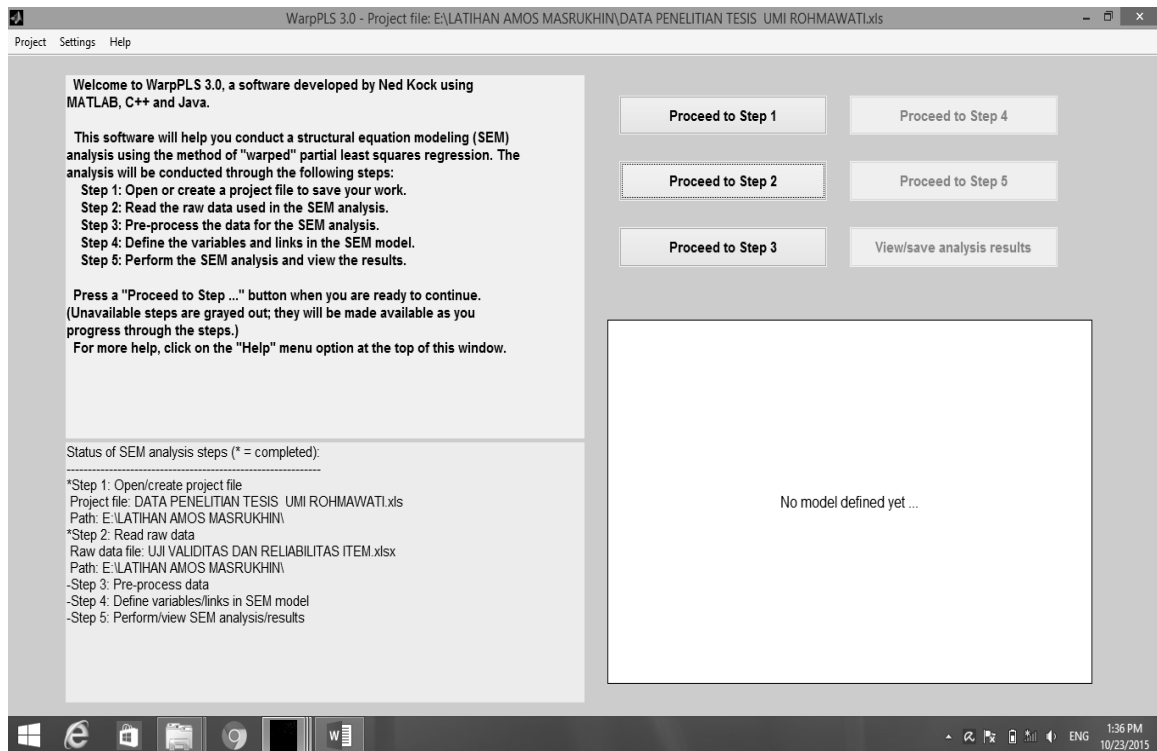


Does the data look correct? Data diperiksa, klik Yes.

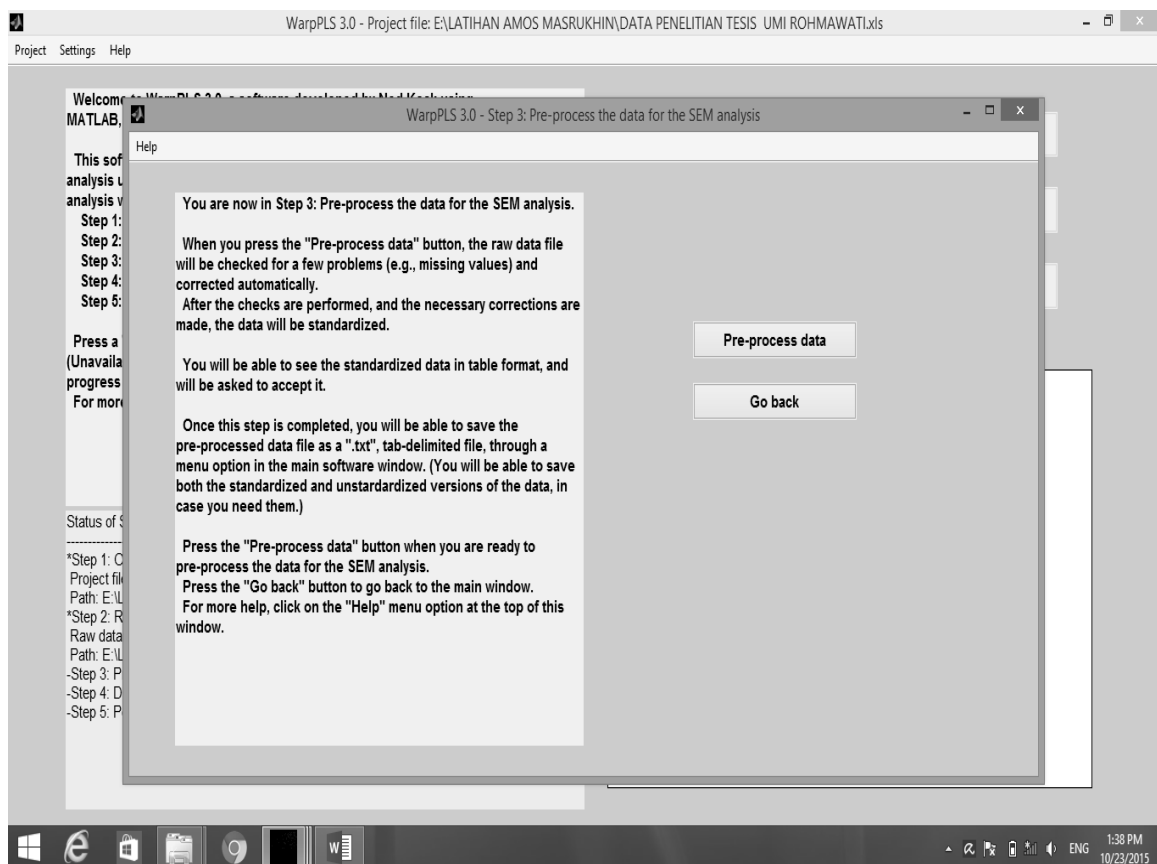
Langkah 3 : Pemrosesan awal (pre-process) data untuk analisis SEM.

Klik OK

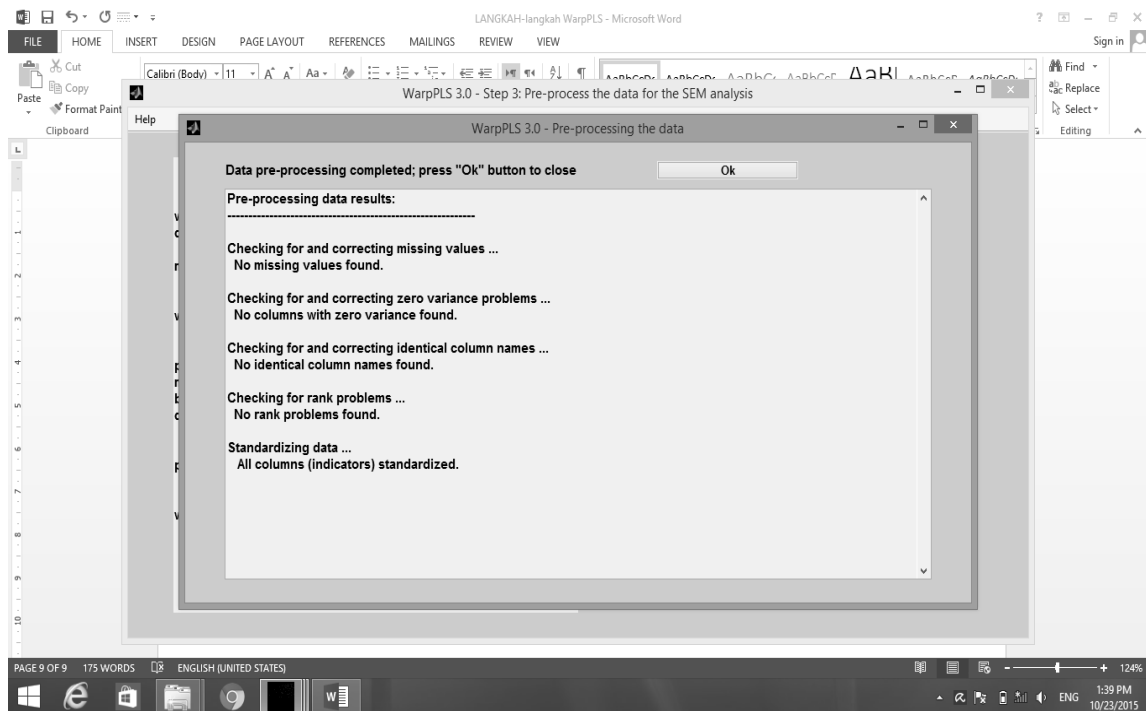
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



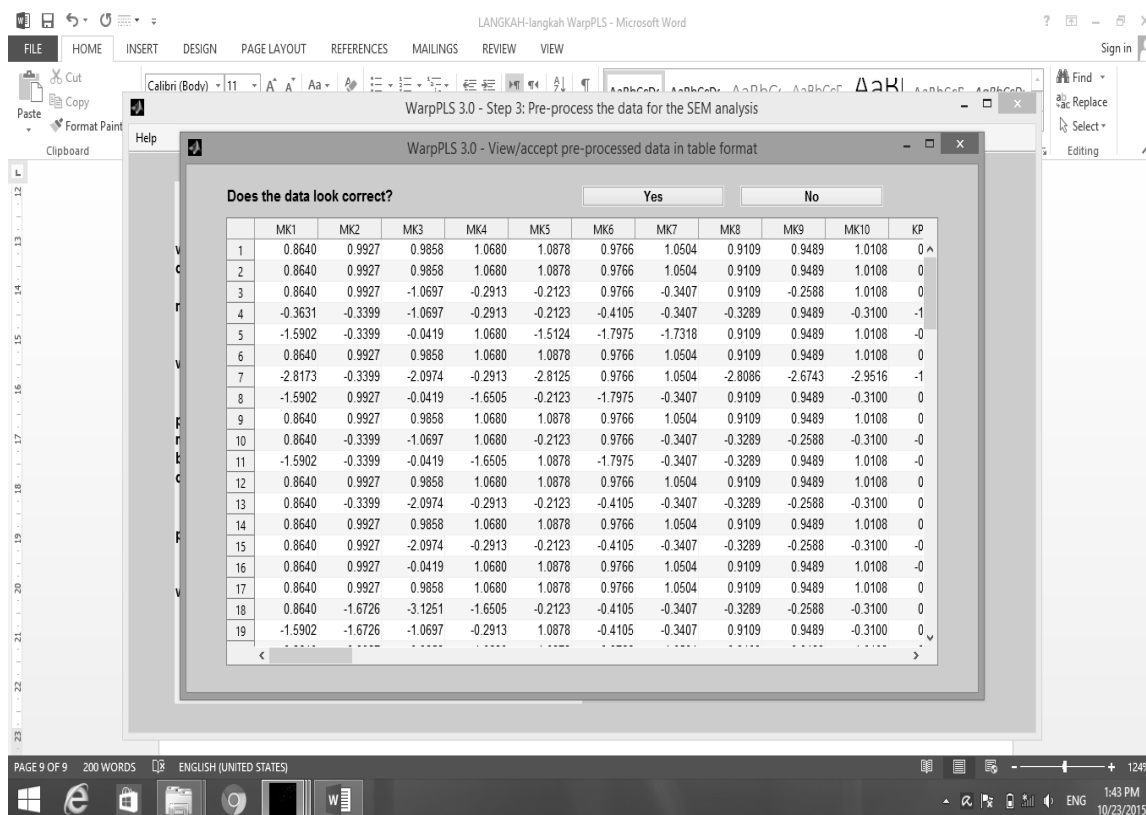
Klik *Pre-process* data sehingga muncul jendela WarpPLS 3.0 –Pre-processing the data.



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

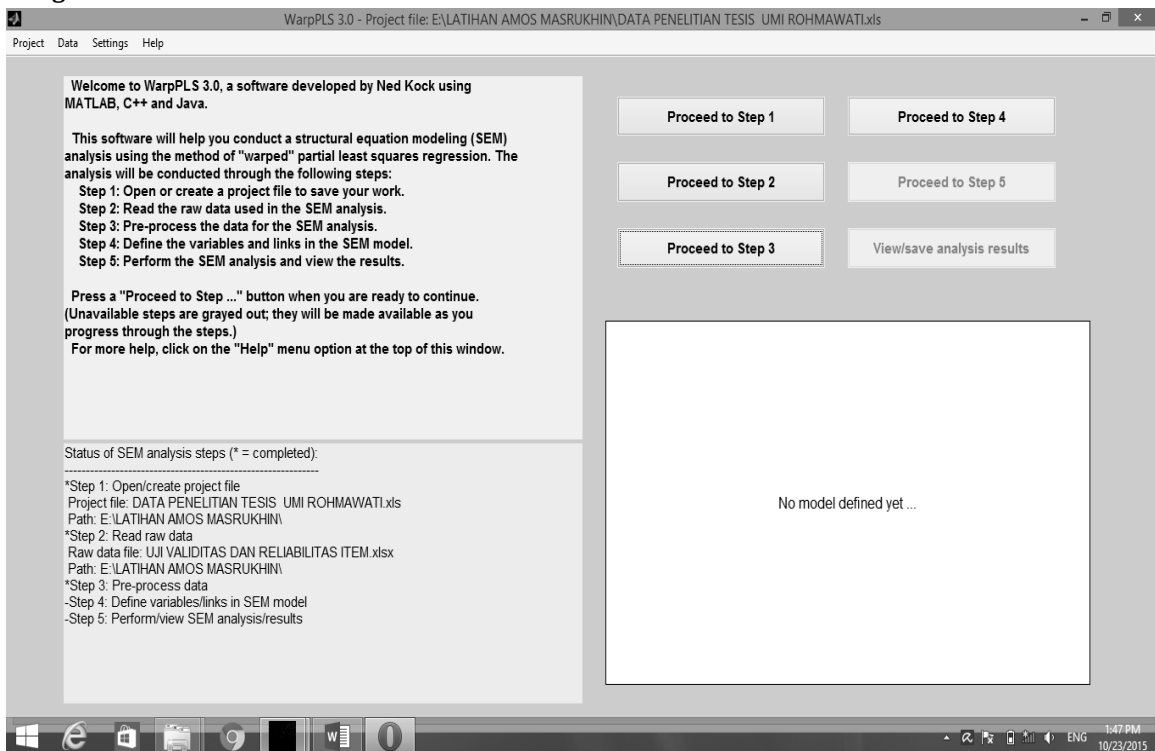


Gambar diatas menunjukkan bahwa tidak ada masalah dalam data seperti *missing values*, *zero variance problem*, dan *rank problems*. Klik OK sehingga muncul data yang distandarisasi.

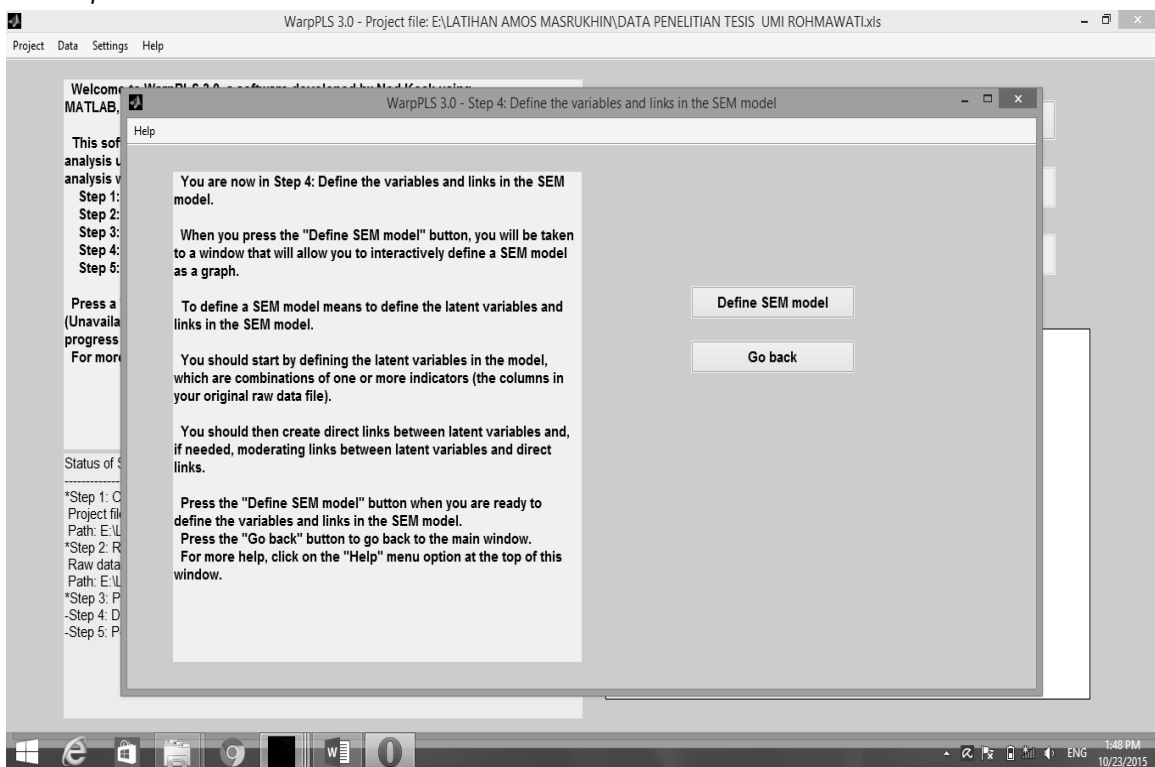


Jika tidak ada *outlier*, dimana *outlier* dengan kriteria -4 s/d 4. Jika data tidak ada yang *outlier*, maka klik yes dan OK.

Langkah 4: Mendefinisikan Model SEM.

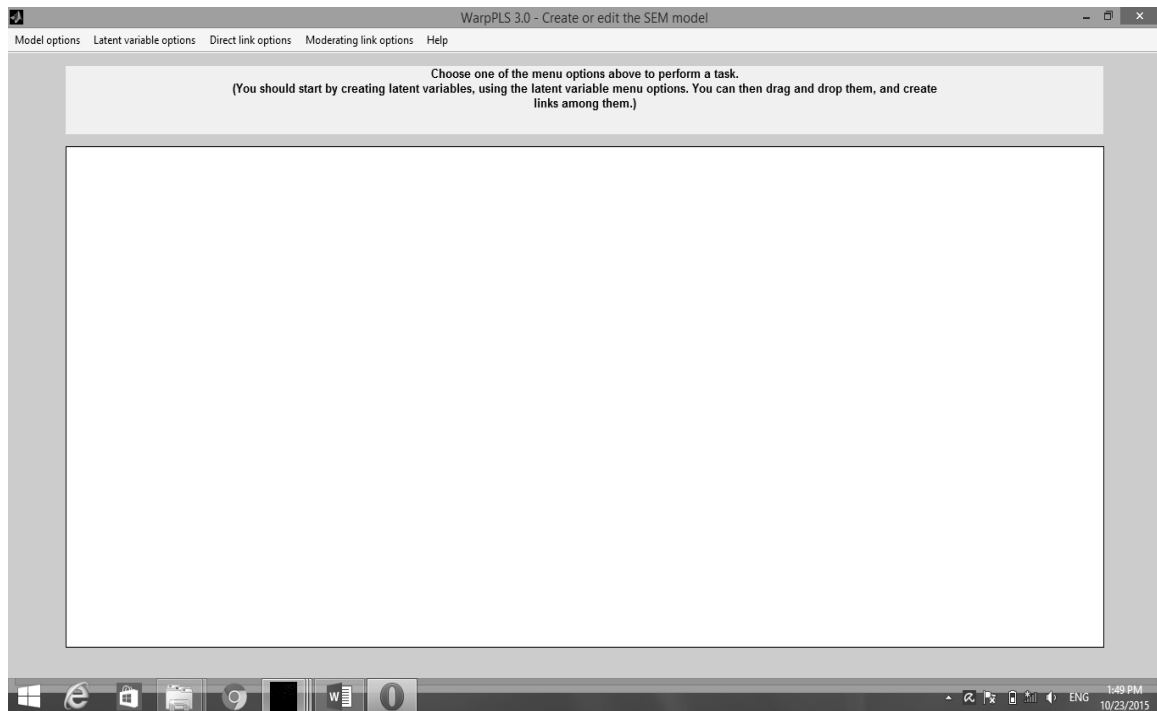


Klik Define SEM Model.

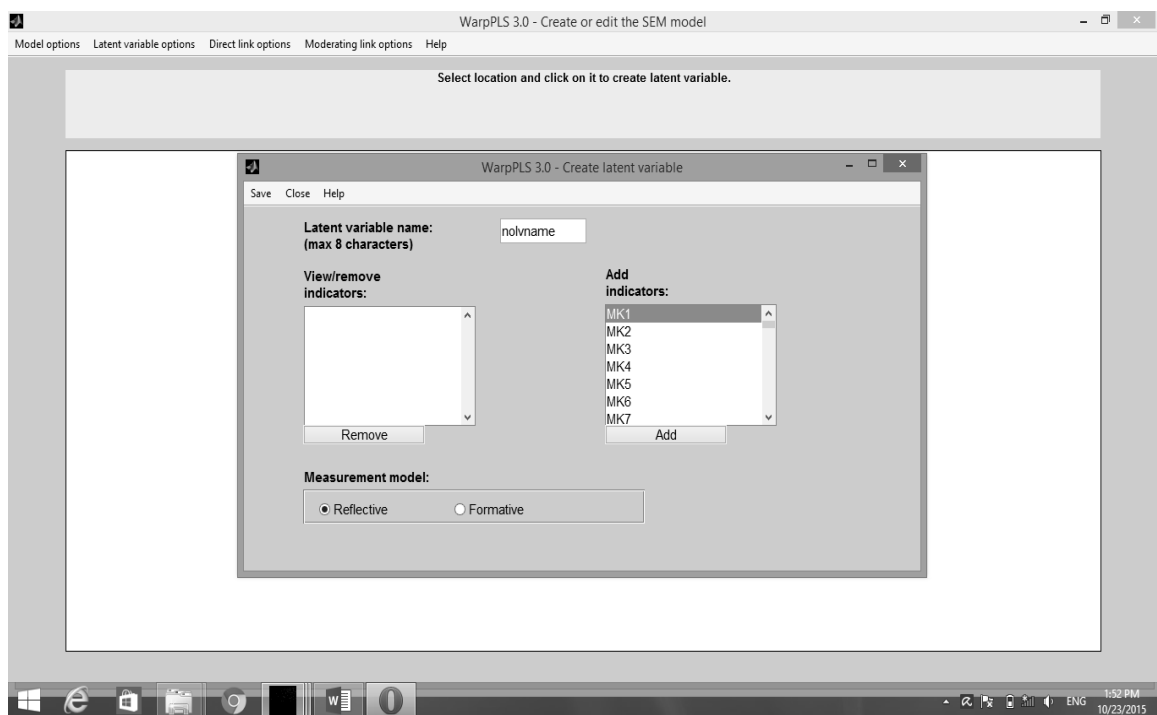


Selanjutnya buatlah gambar model *direct effect* dengan memilih menu *Latent variable options*, lalu *Create latent variable*.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

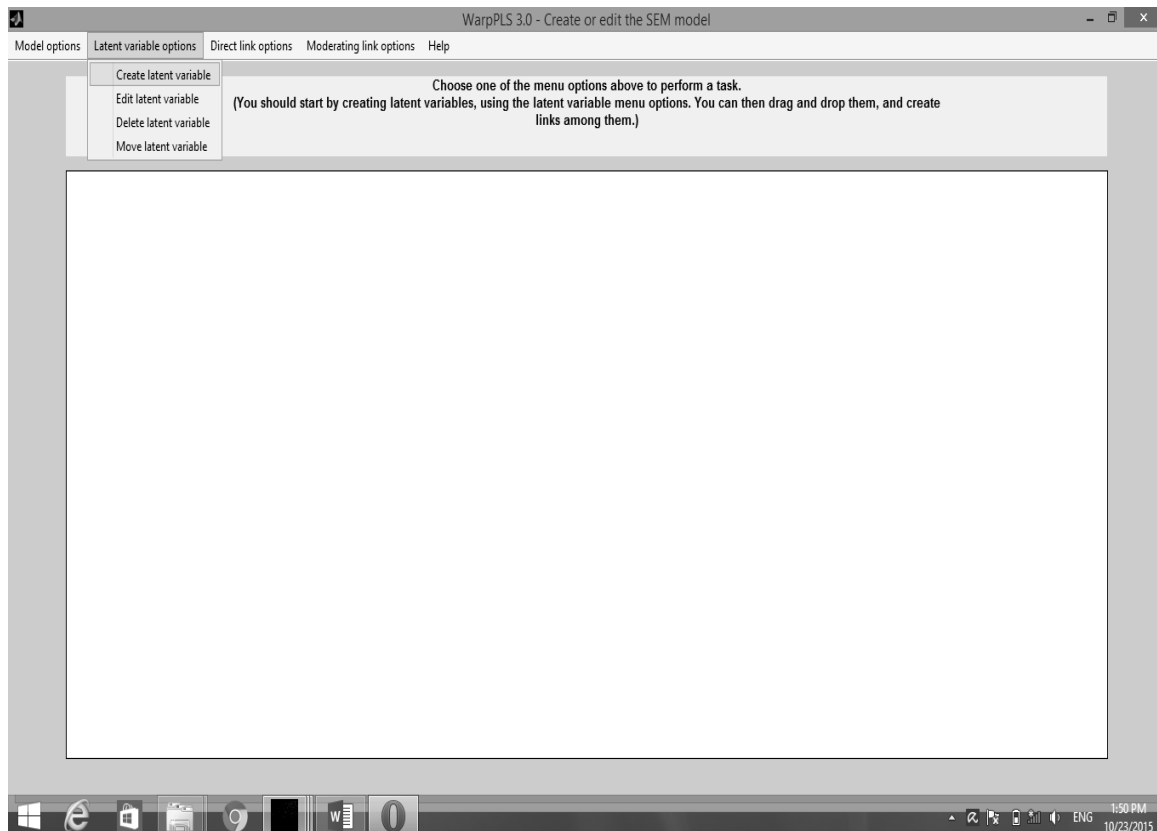
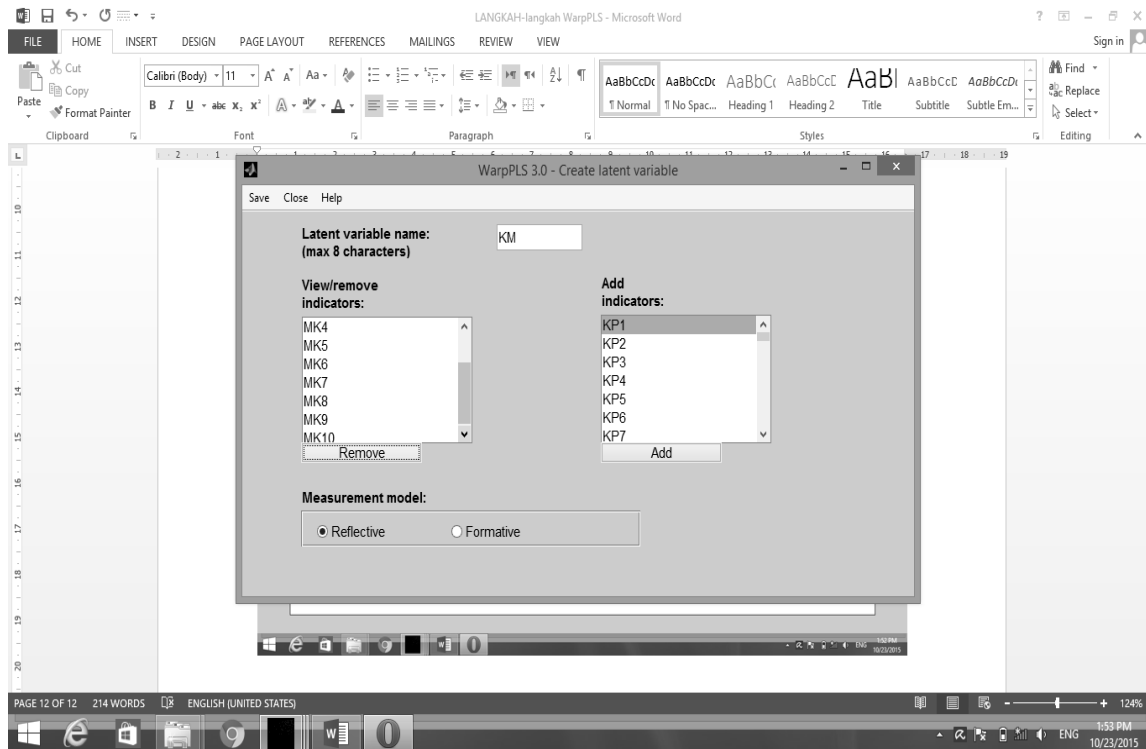


Letakkan *cursor* pada bidang untuk menggambar, lalu klik sehingga muncul jendela seperti di bawah ini.



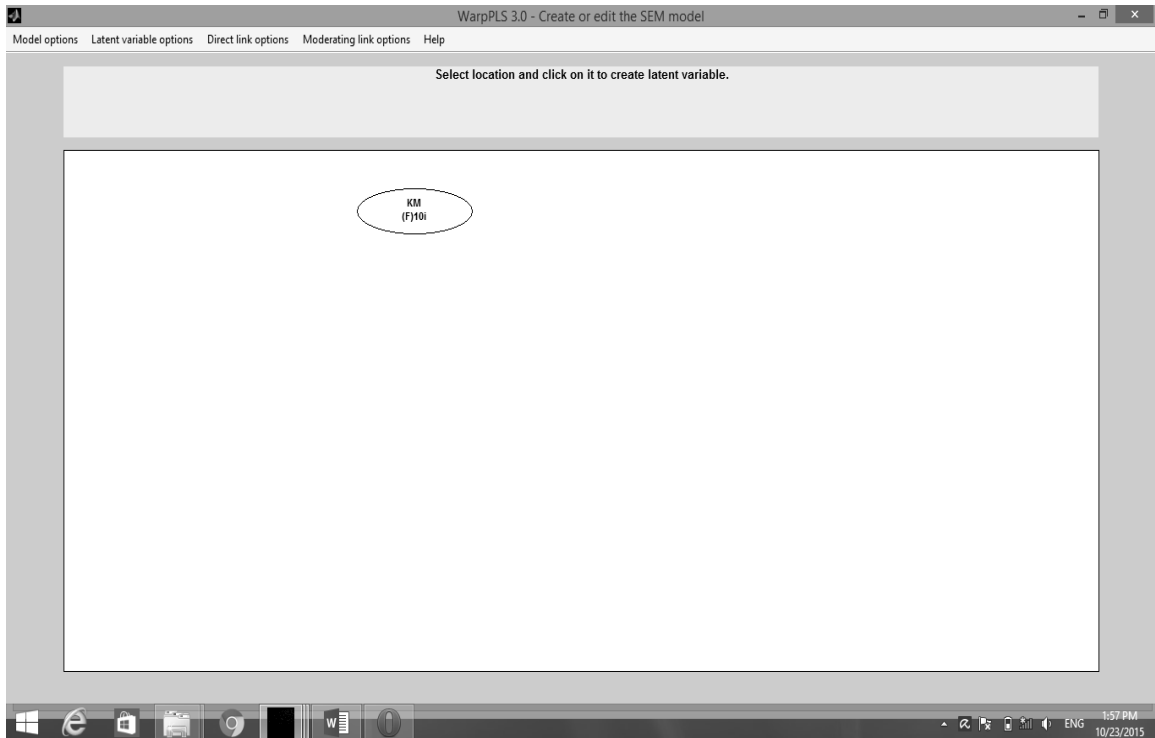
Pada kolom *latent variable name*, ketik KM dan tambahkan indicator yaitu KM1-KM10 dengan measurement model Reflective seperti gambar diatas.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

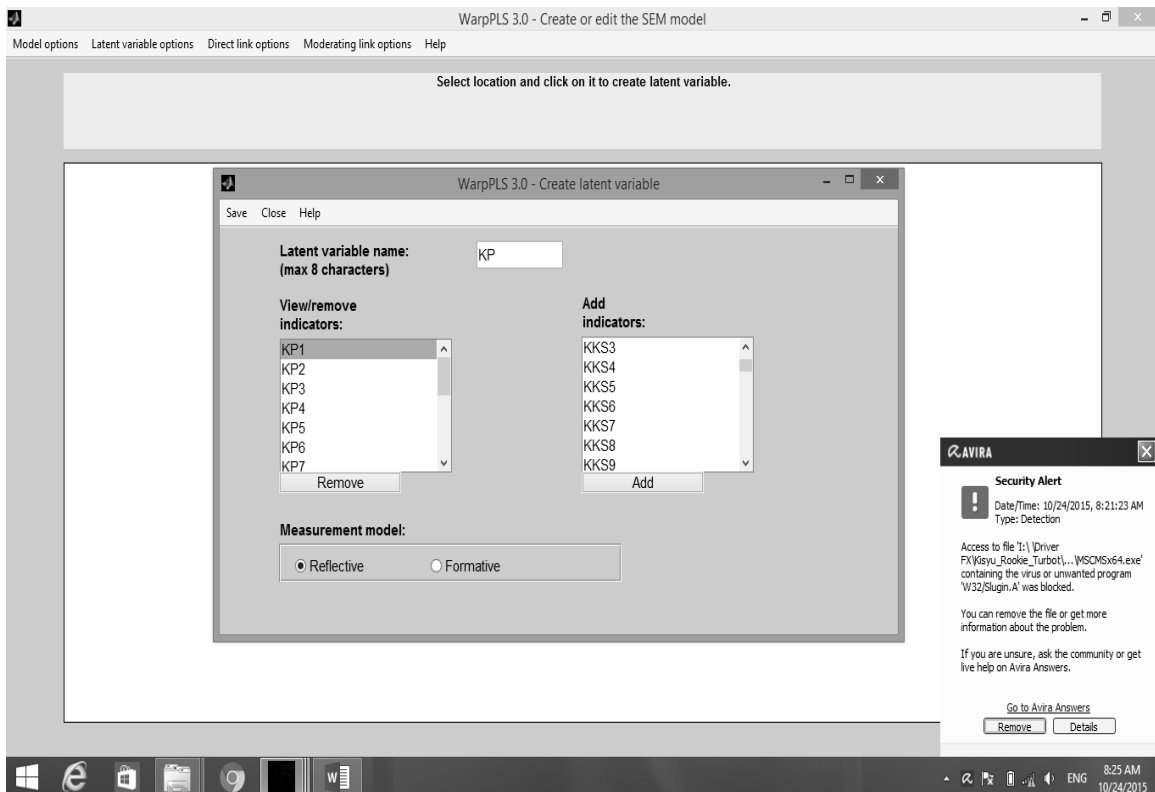


Klik save dan pilih *save latent variable settings*. Selanjutnya diperoleh tampilan seperti gambar dibawah ini.

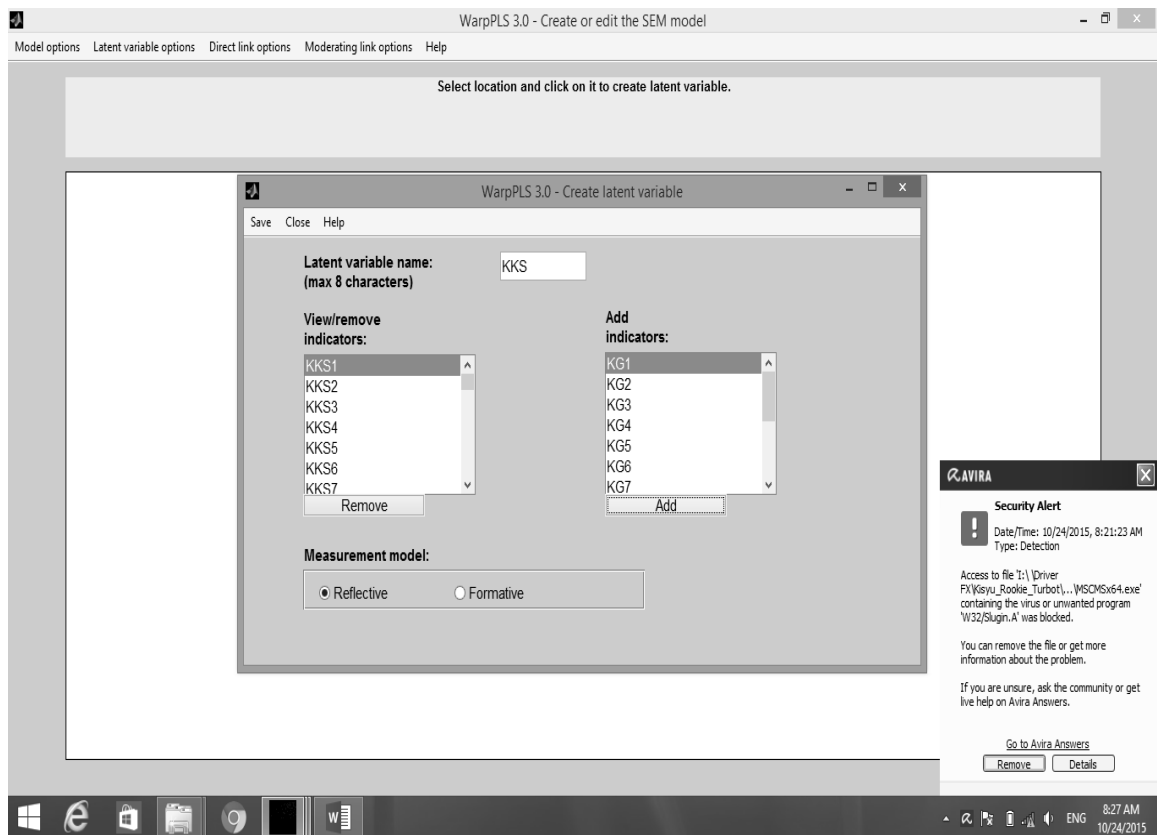
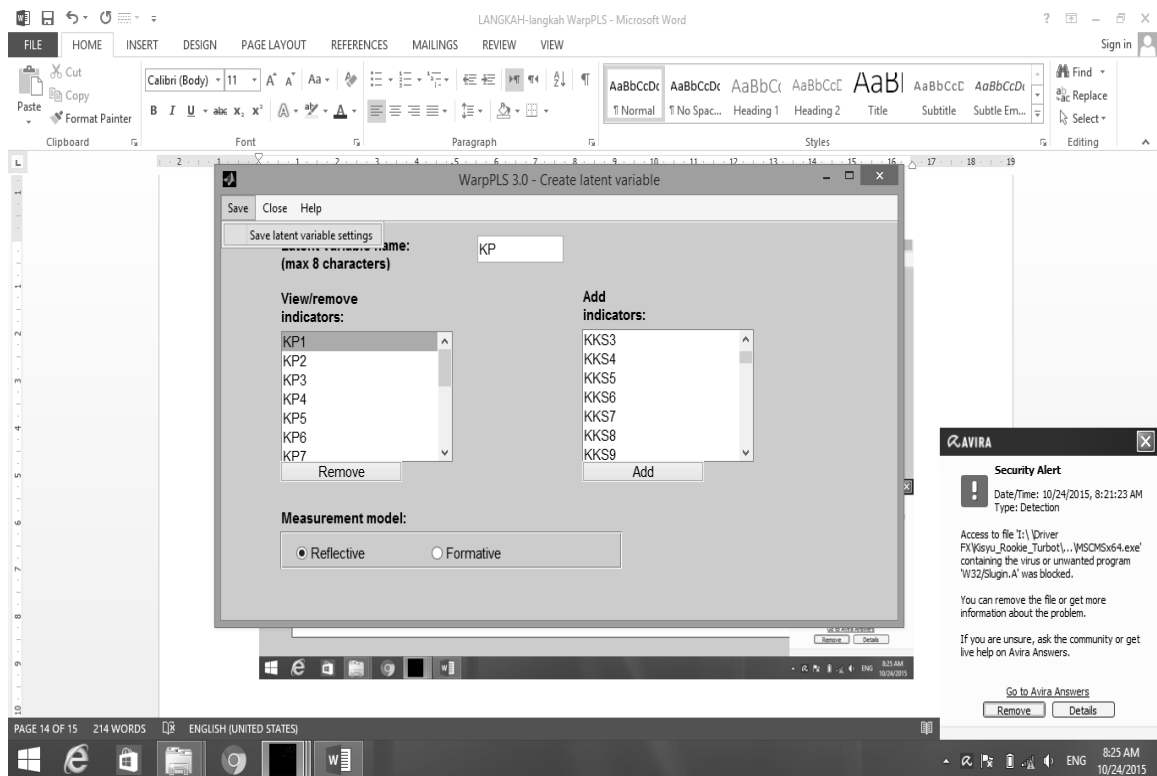
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Ulangi langkah membuat variable laten untuk konstruk motivasi kerja (KM) indikator KM1 sampai KM10.



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

WarpPLS 3.0 - Create latent variable

Save Close Help

Save latent variable settings

Latent variable name:
(max 8 characters)

KM

View/remove indicators:

MK1
MK2
MK3
MK4
MK5
MK6
MK7
MK8
MK9
MK10

Remove

Add indicators:

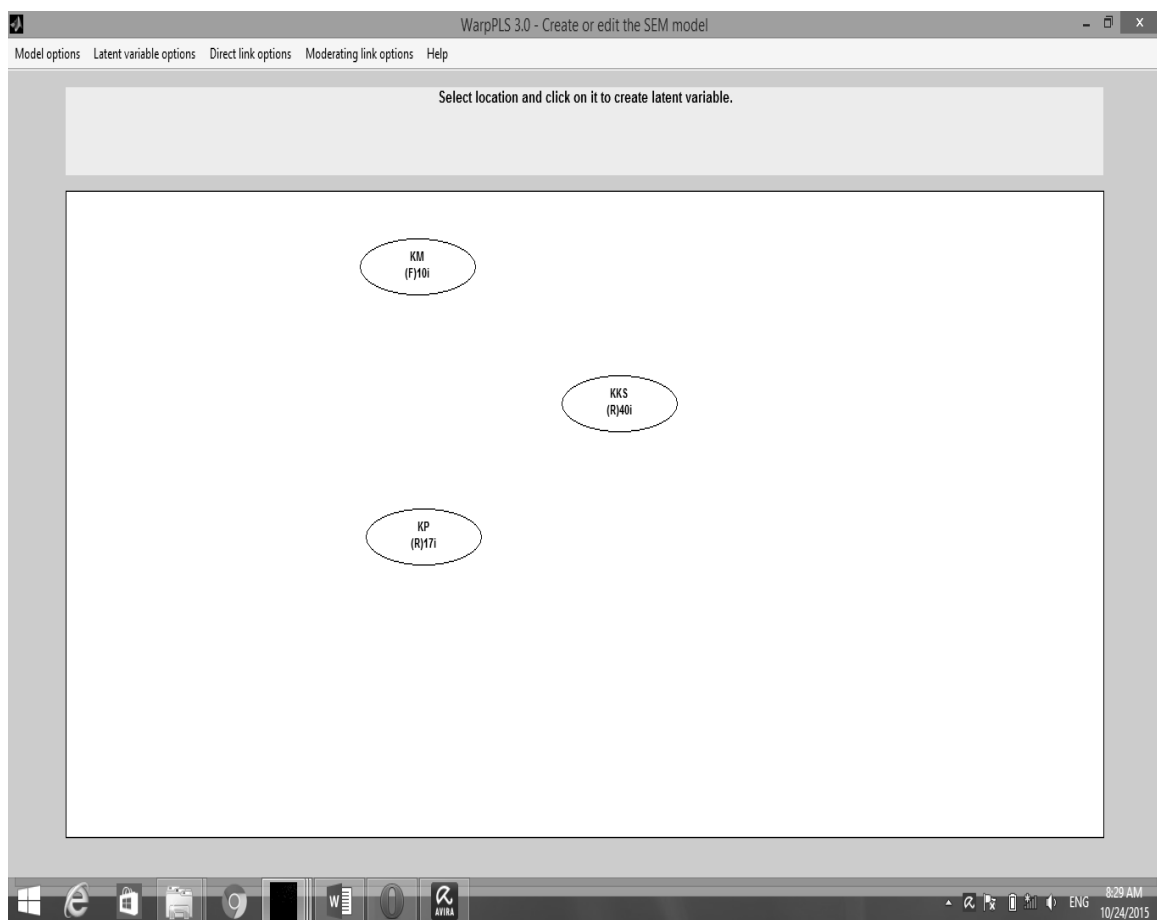
KP1
KP2
KP3
KP4
KP5
KP6
KP7
KP8
KP9
KP10
KP11
KP12

Add

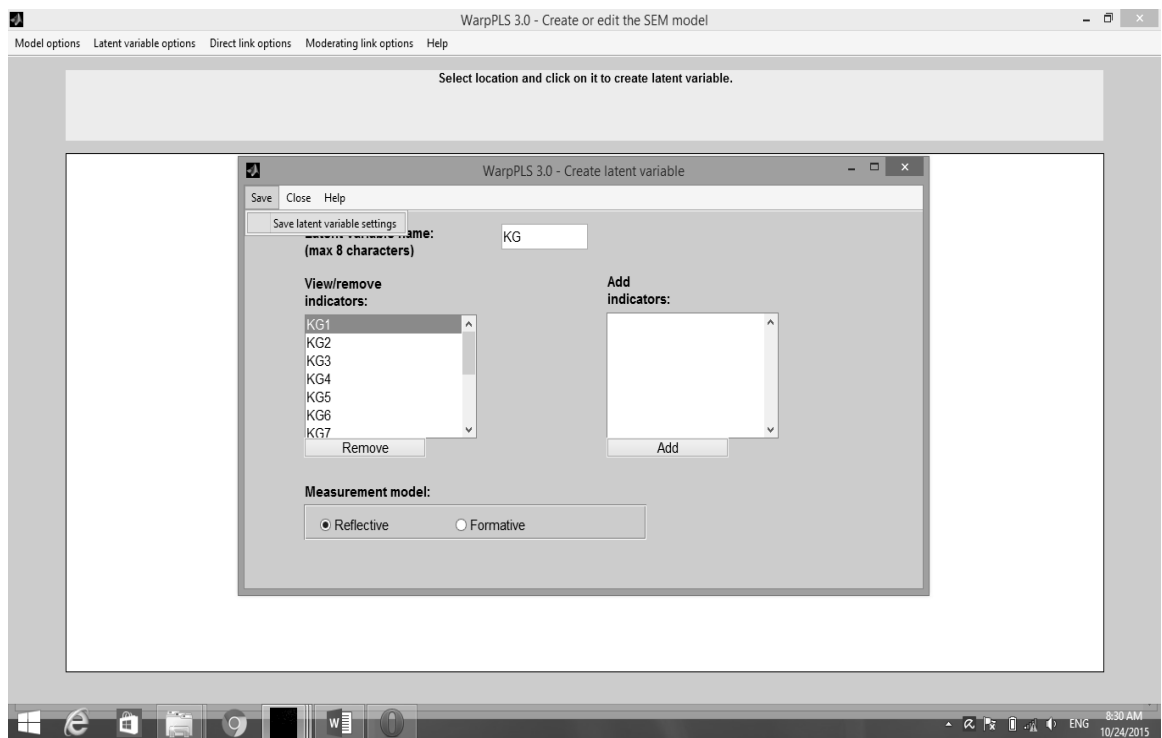
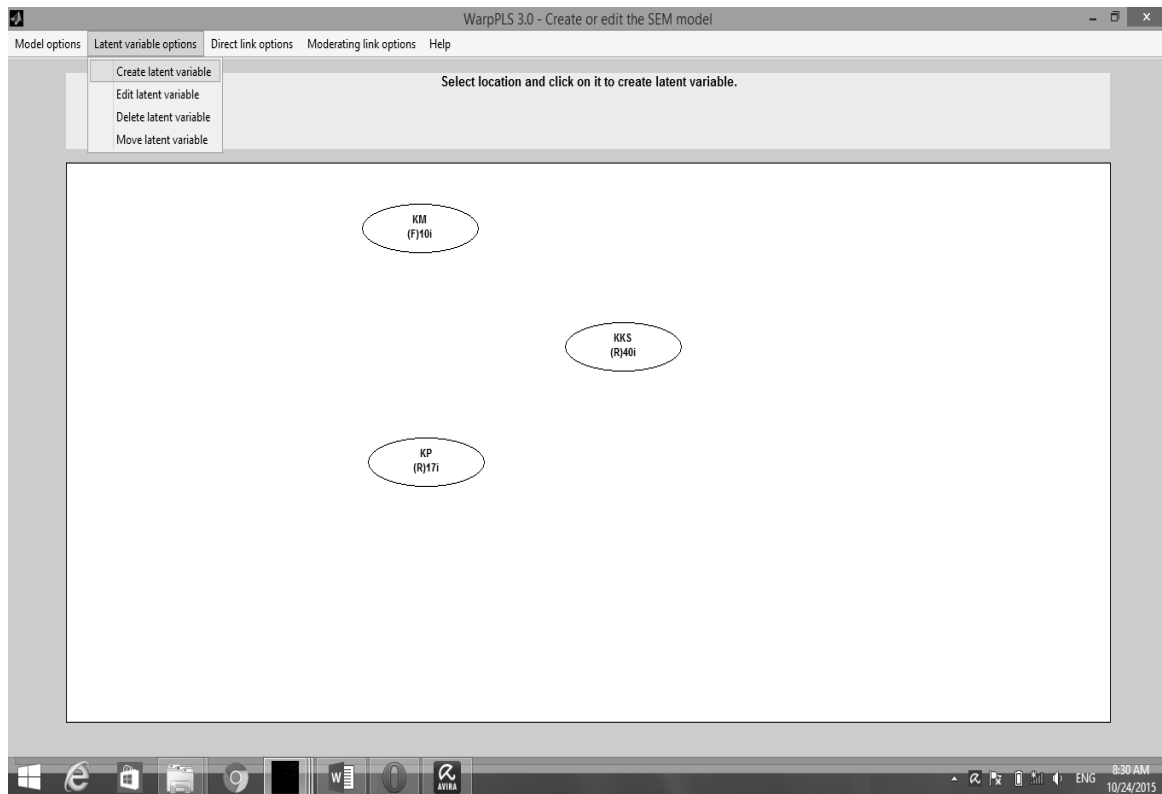
Measurement model:

☐ Reflective ☐ Formative

Windows taskbar: 1:55 PM 10/23/2015

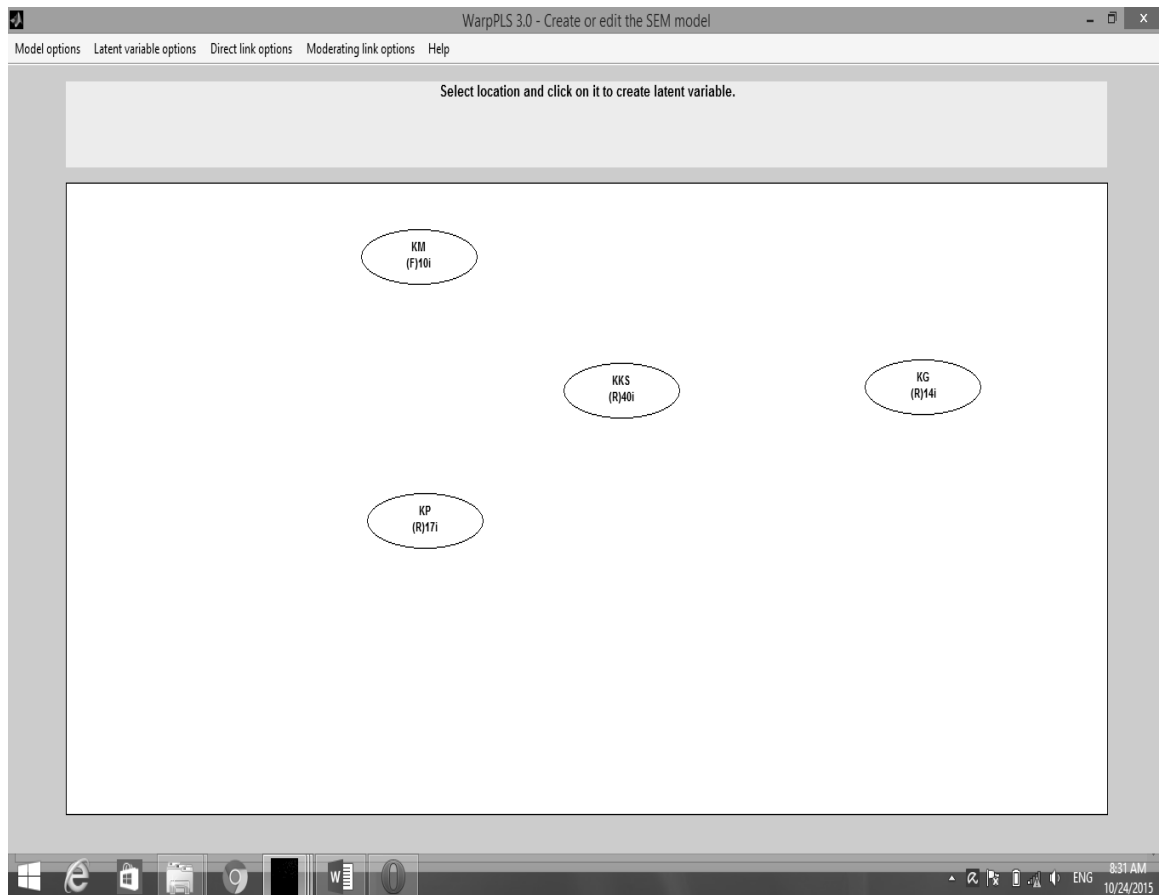


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

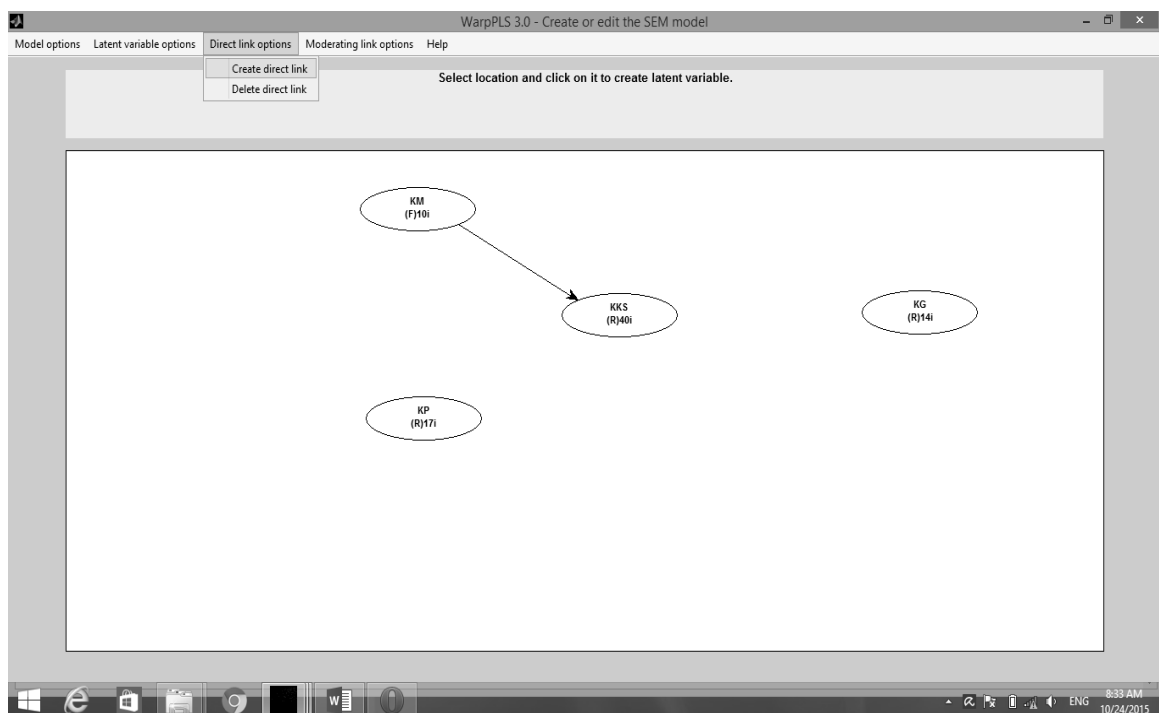


Setelah semua variabel ini dibuat sebagaimana gambar di bawah ini, maka langkah selanjutnya adalah buat jalur pengaruh KM terhadap KKS, KP terhadap KKS, KKS terhadap KG, secara langsung KM terhadap KG dan KP terhadap KG dengan mengklik *Direct link options* dan pilih *Create direct link*

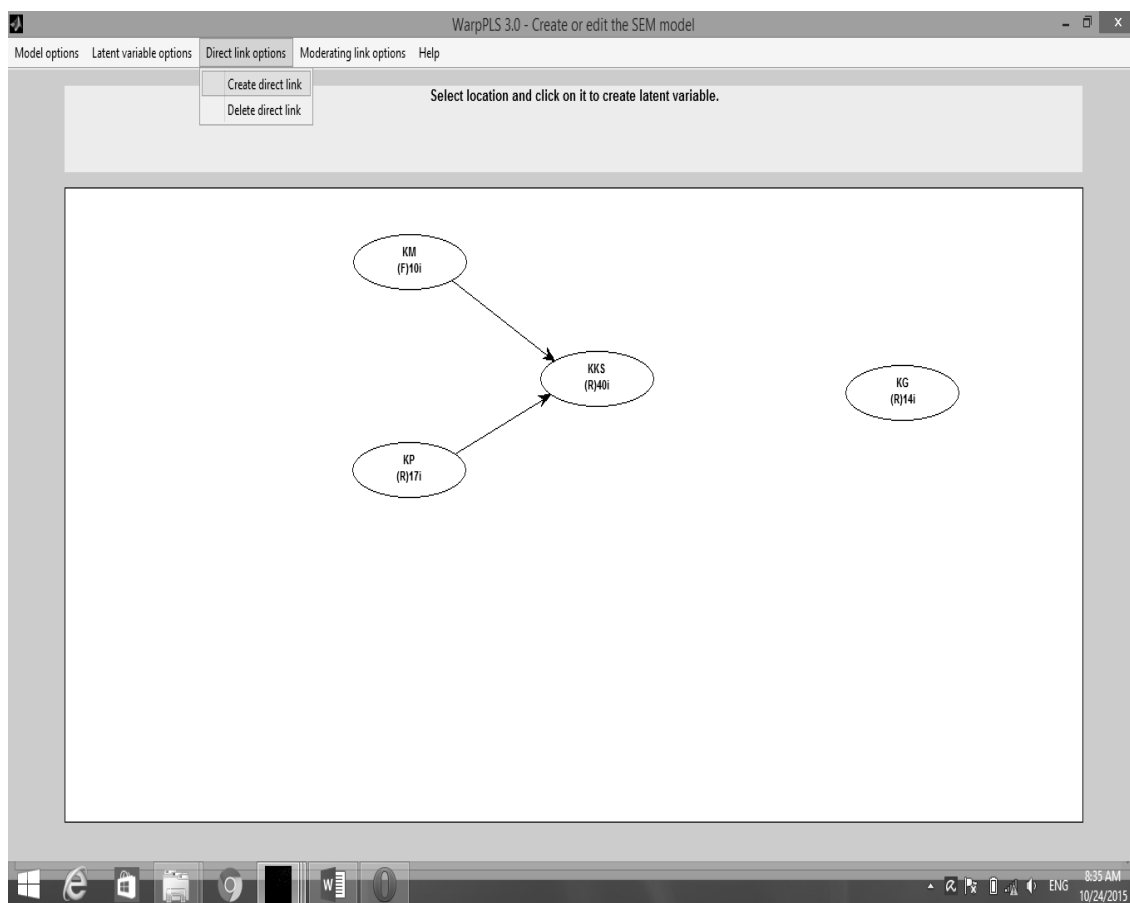
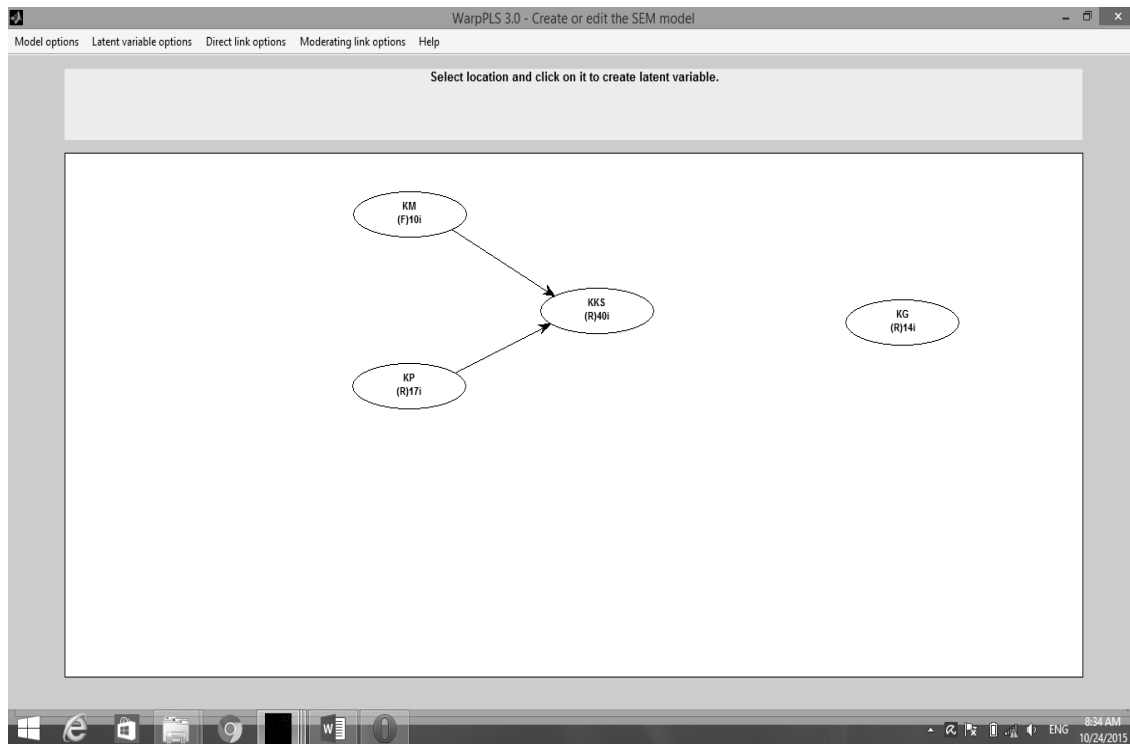
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



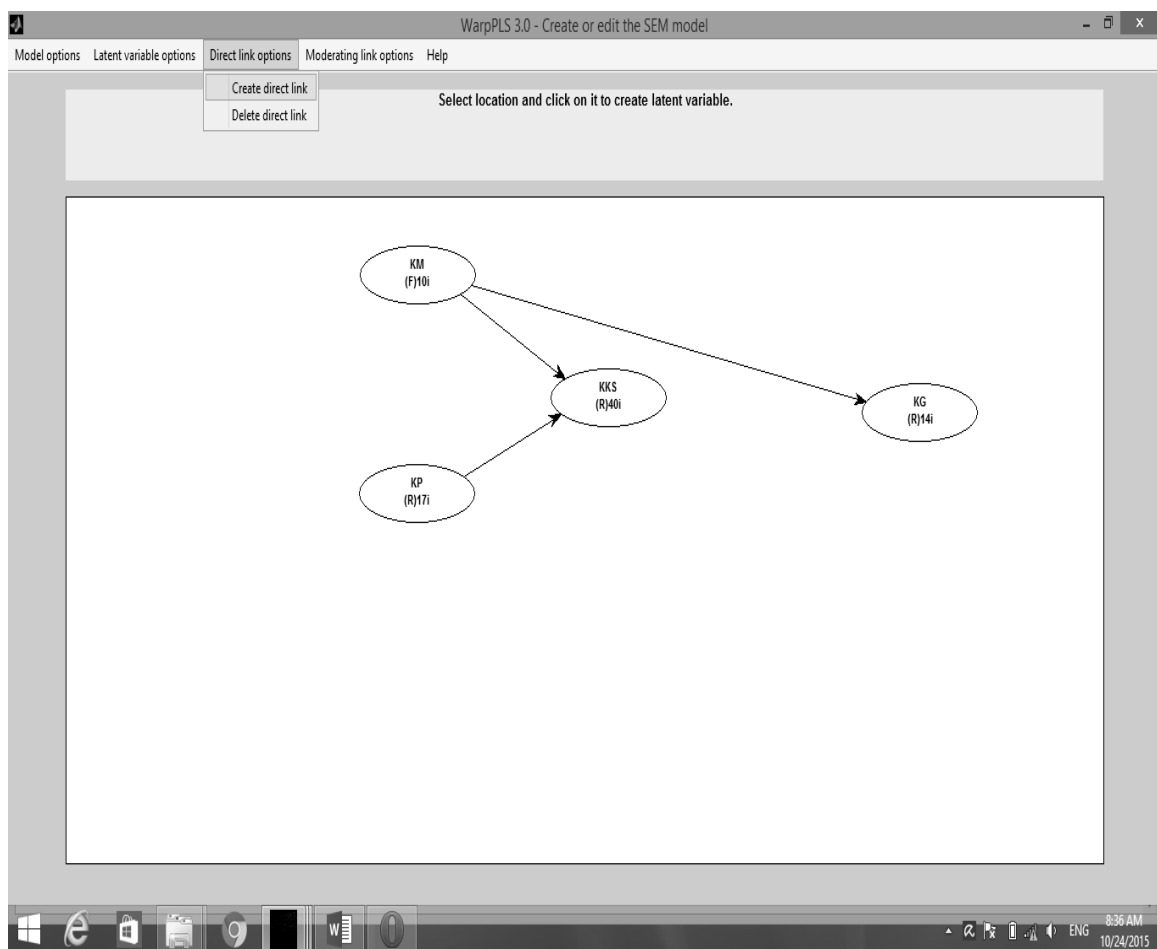
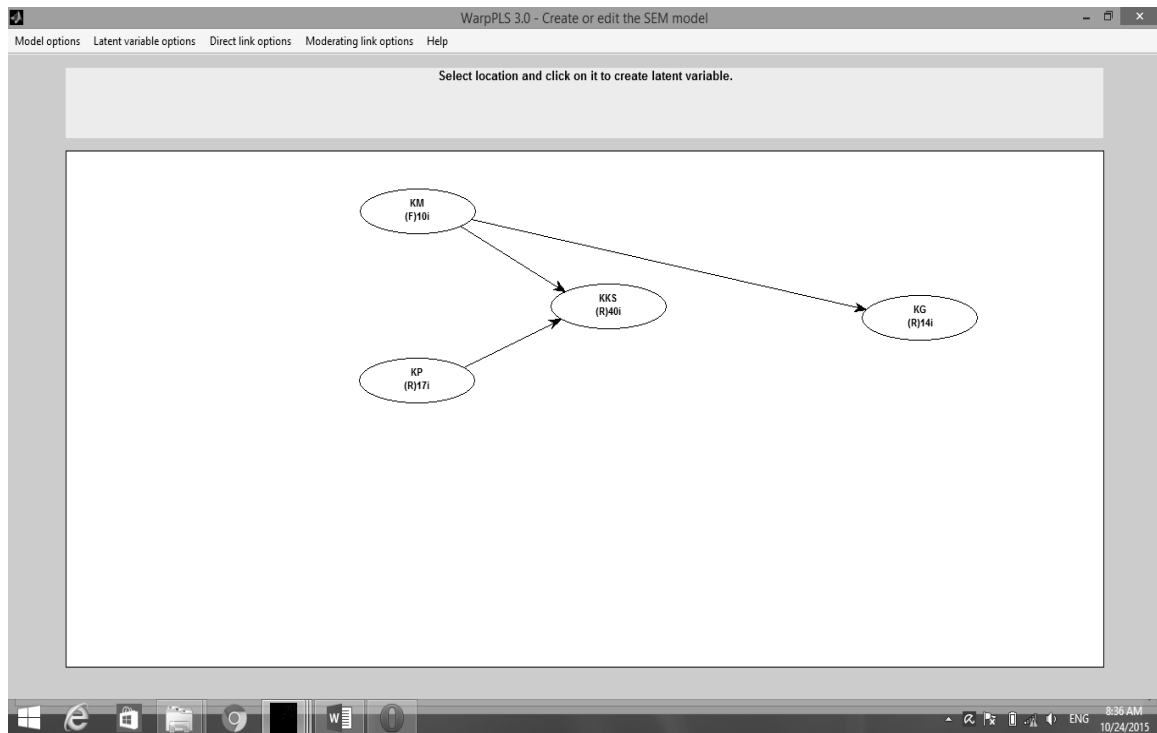
Buat jalur dengan mengklik pada variable laten predictor/eksogen (yaitu KM) lalu ke variable laten kriteria /endogen KKS sehingga diperoleh tampilan seperti berikut:



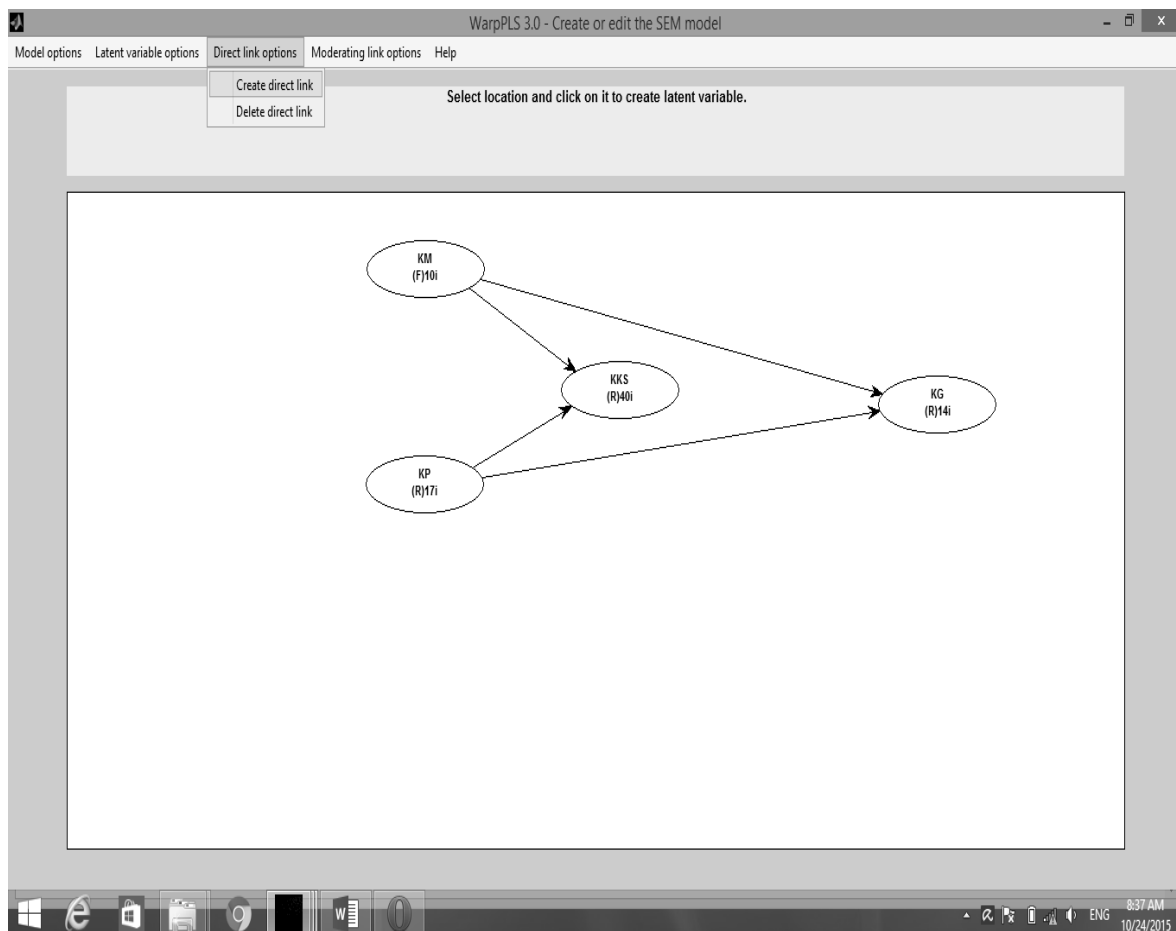
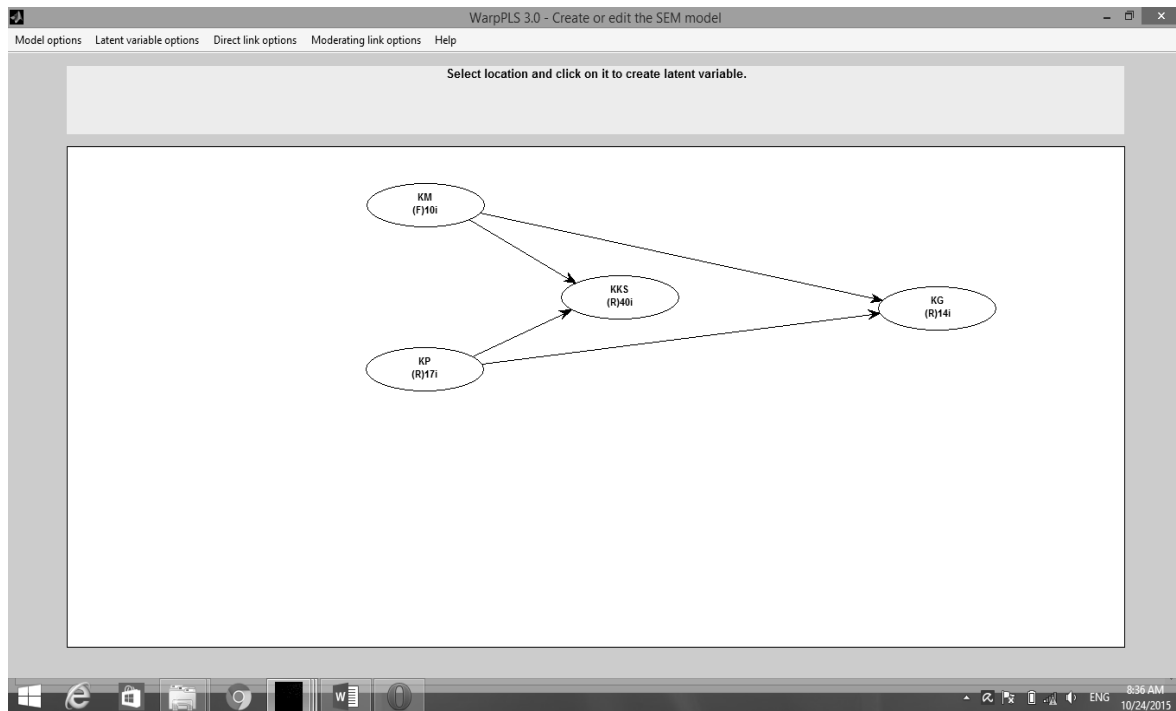
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



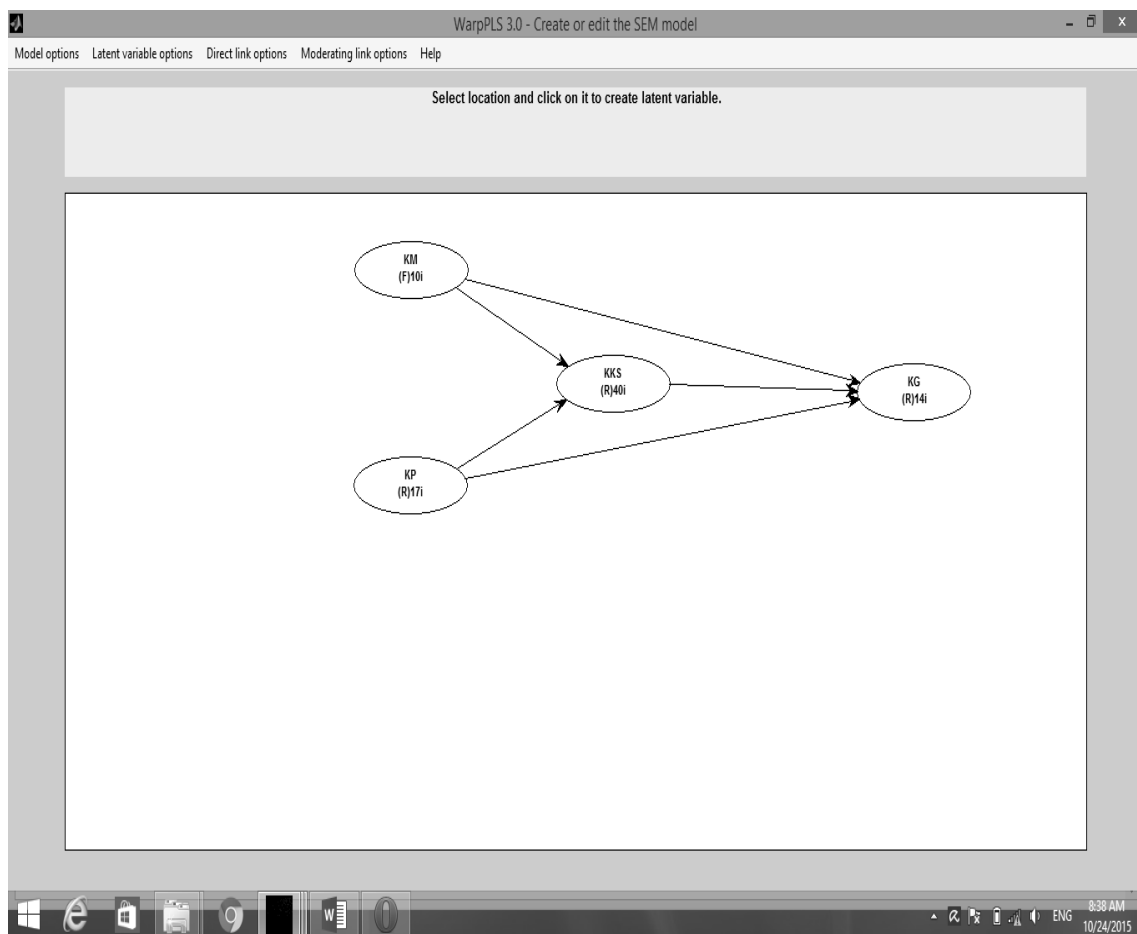
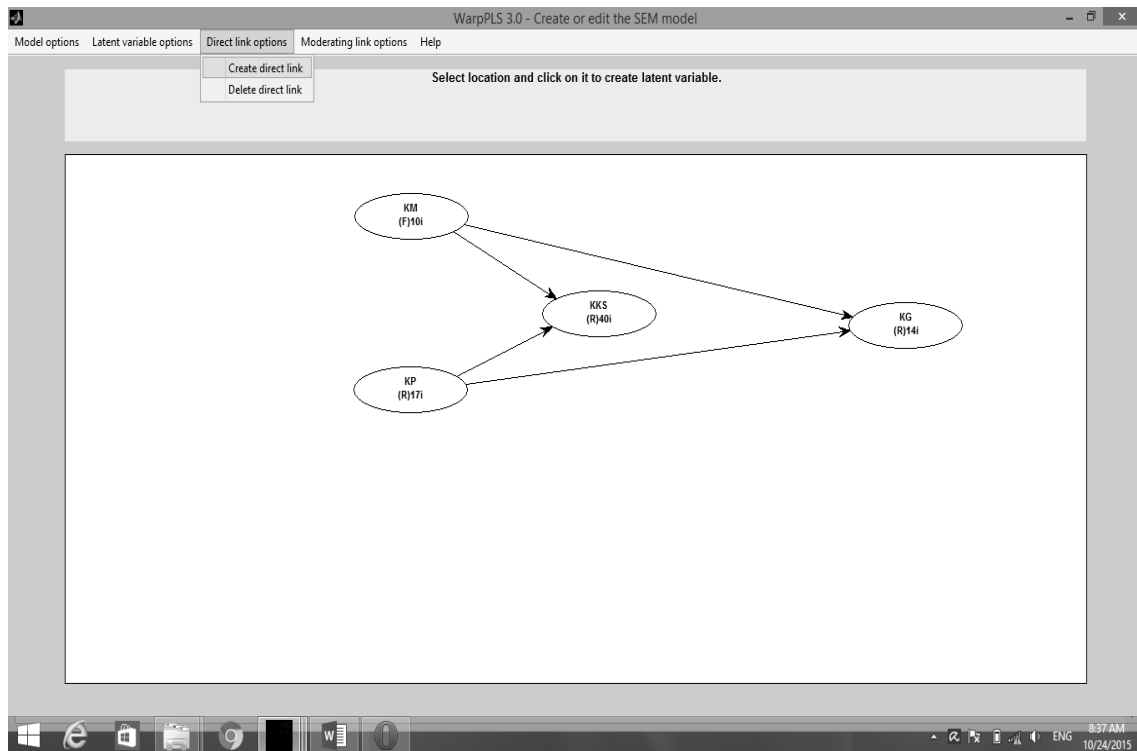
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



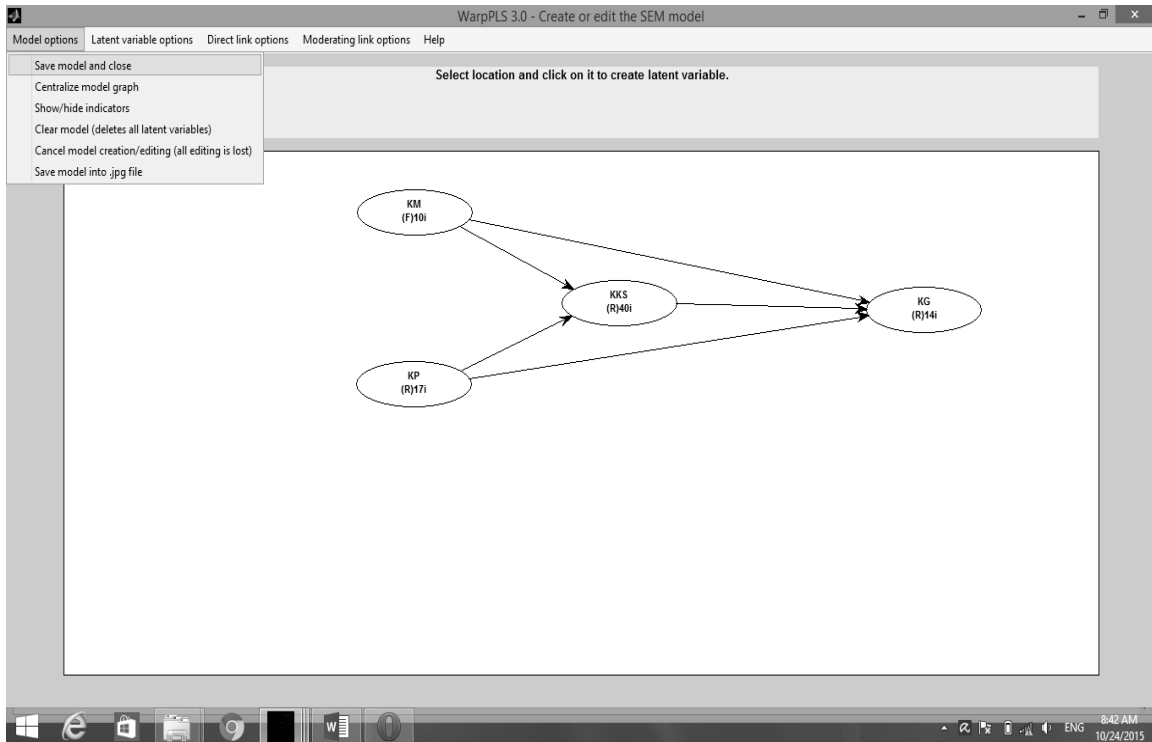
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

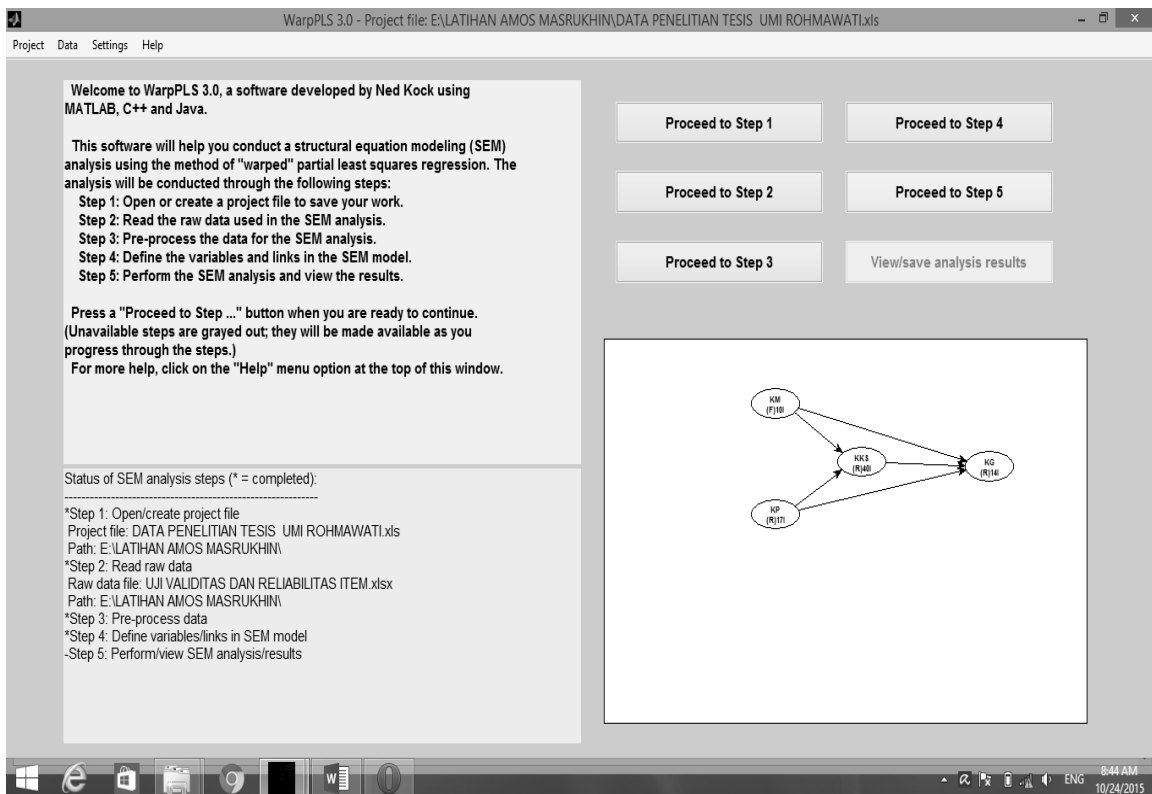


Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

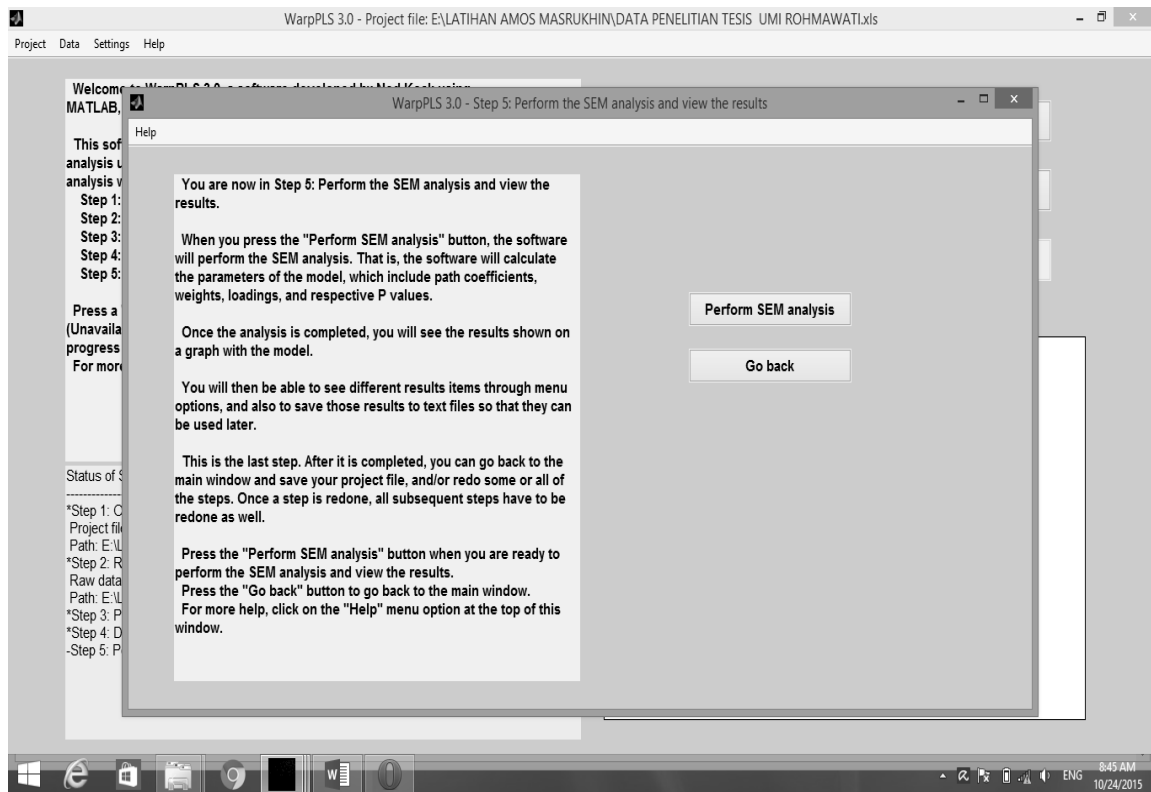


Langkah 5 : Menjalankan Analisi SEM dan Melihat Hasilnya (Output)

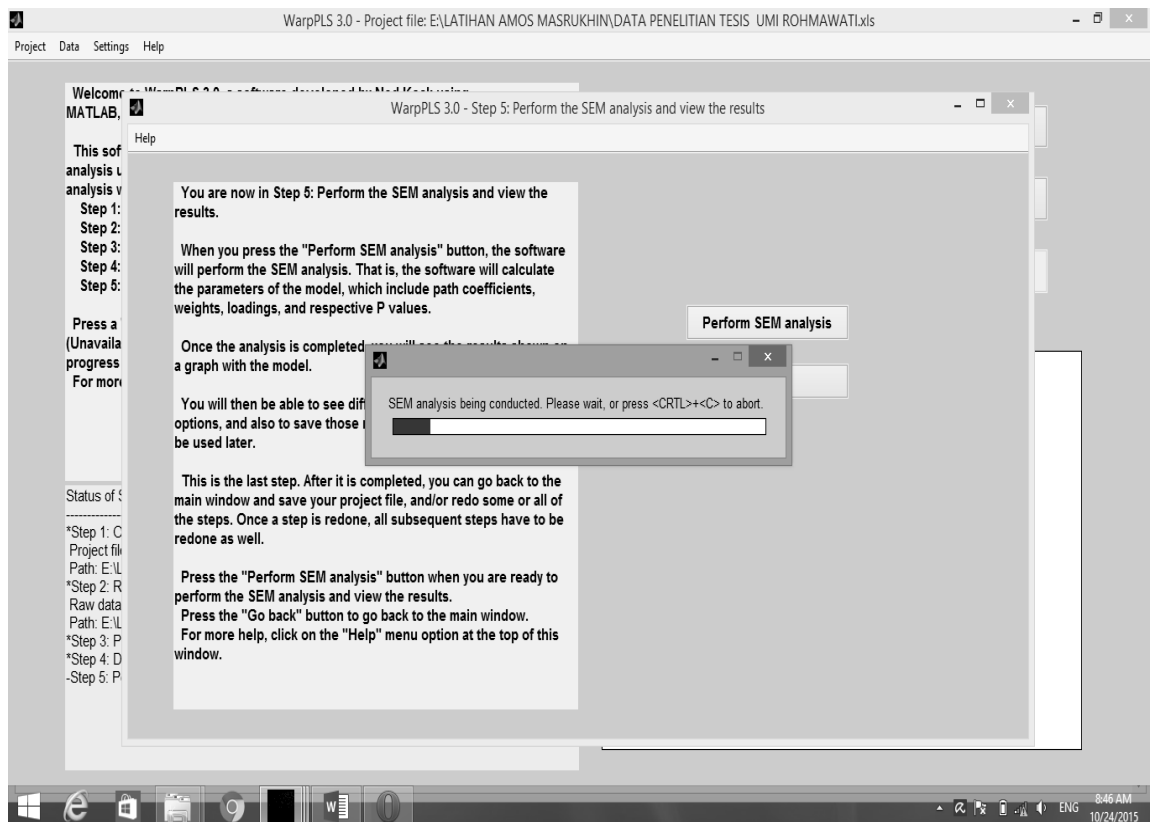
Pada langkah ini, kita akan menjalankan (*run*) analisis SEM untuk gambar model yang telah dibuat pada langkah 4. Kecepatan dari tahapan ini dalam memperoleh output umumnya sama dengan software PLS lain.



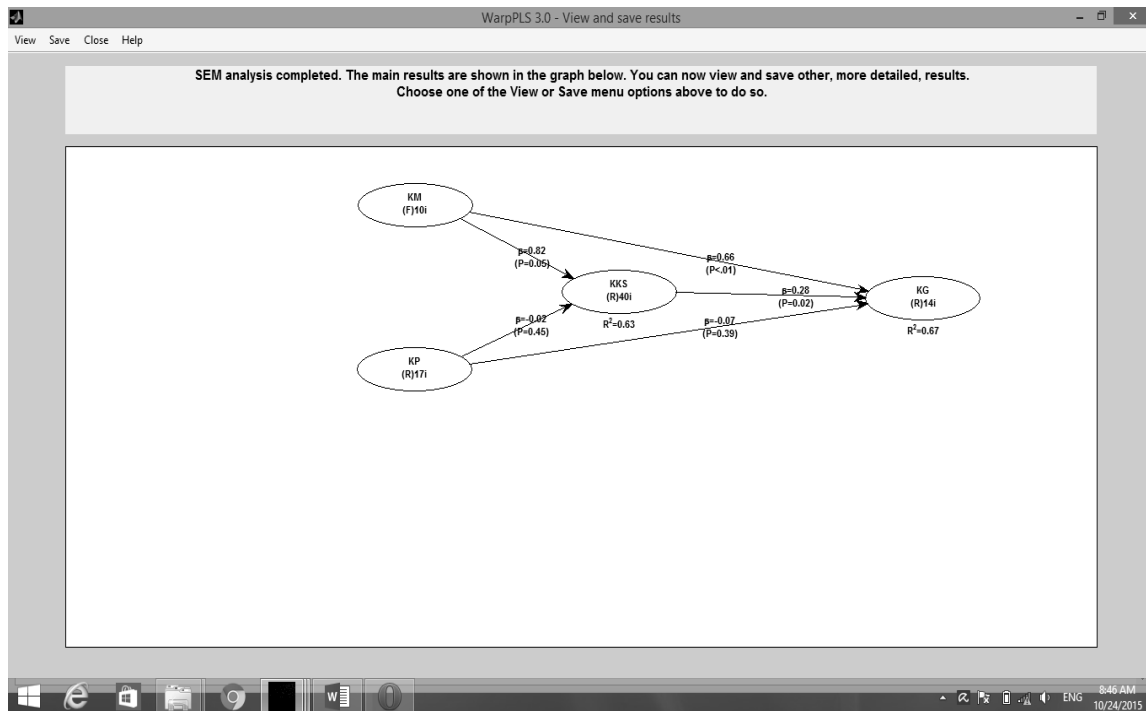
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



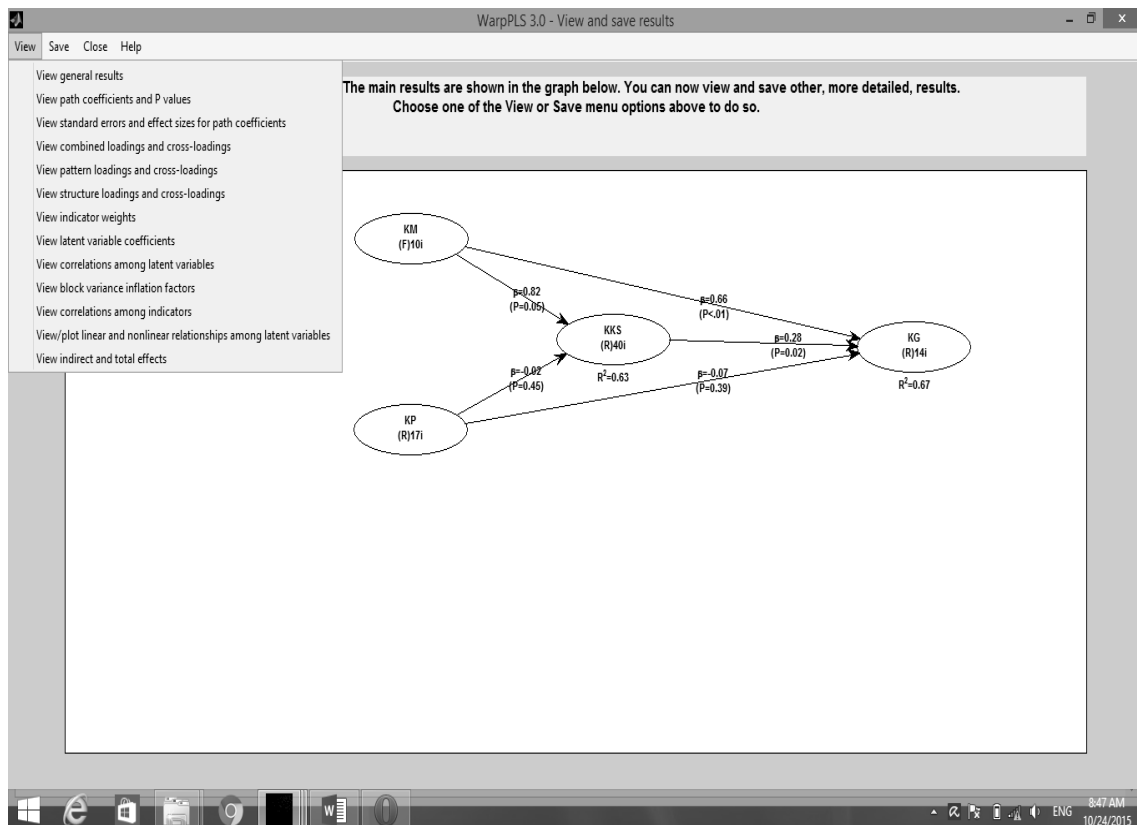
Klik *Perform SEM analysis* maka muncul jendela seperti berikut:



Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

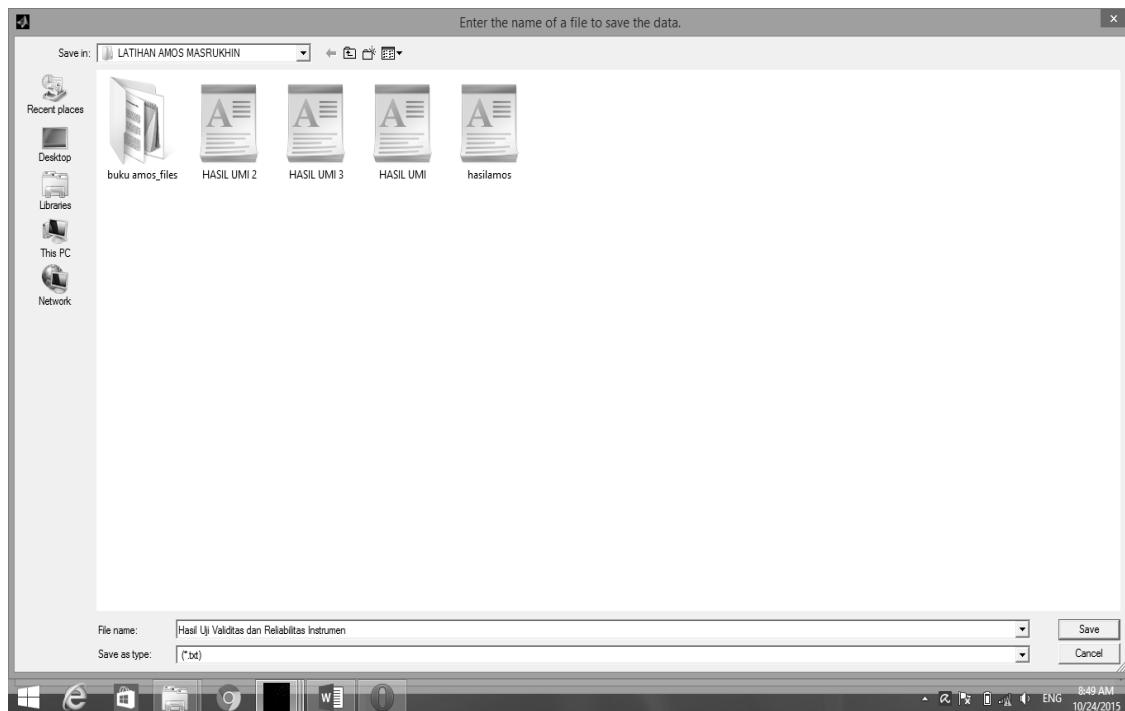
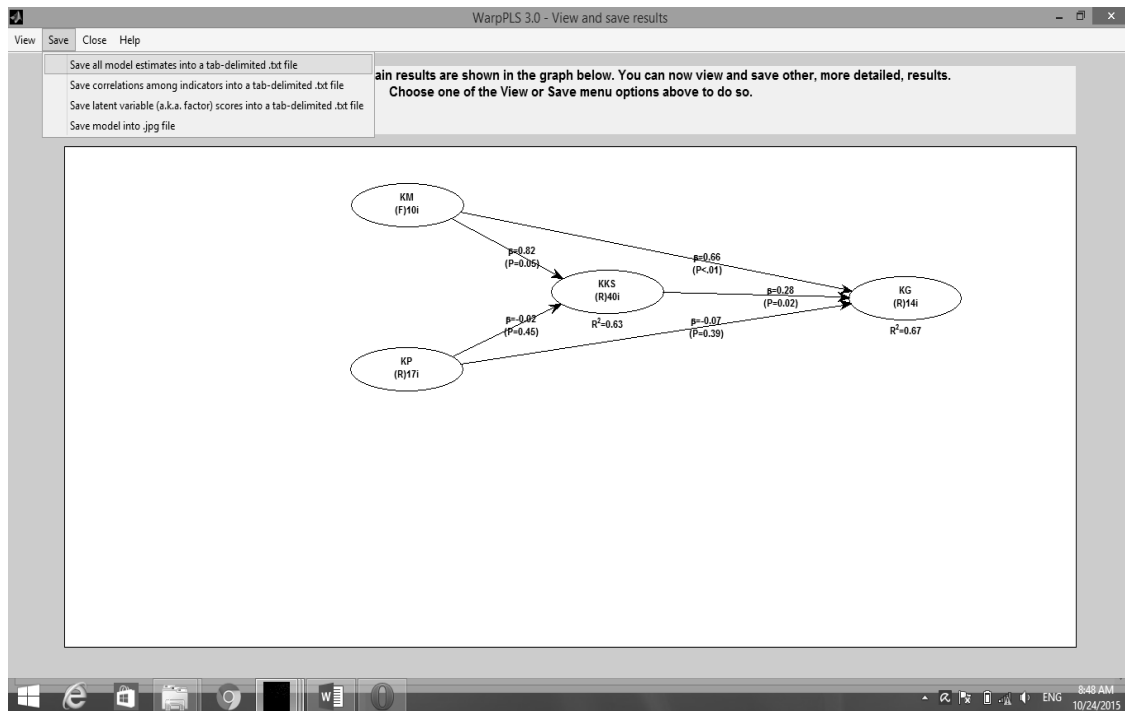


Untuk melihat output secara lengkap, klik *View* lalu pilih *output* yang digunakan.



Untuk menyimpan keseluruhan output yang ada, klik *save*, lalu pilih *Save all model estimates into a tab-delimited.txt file*.

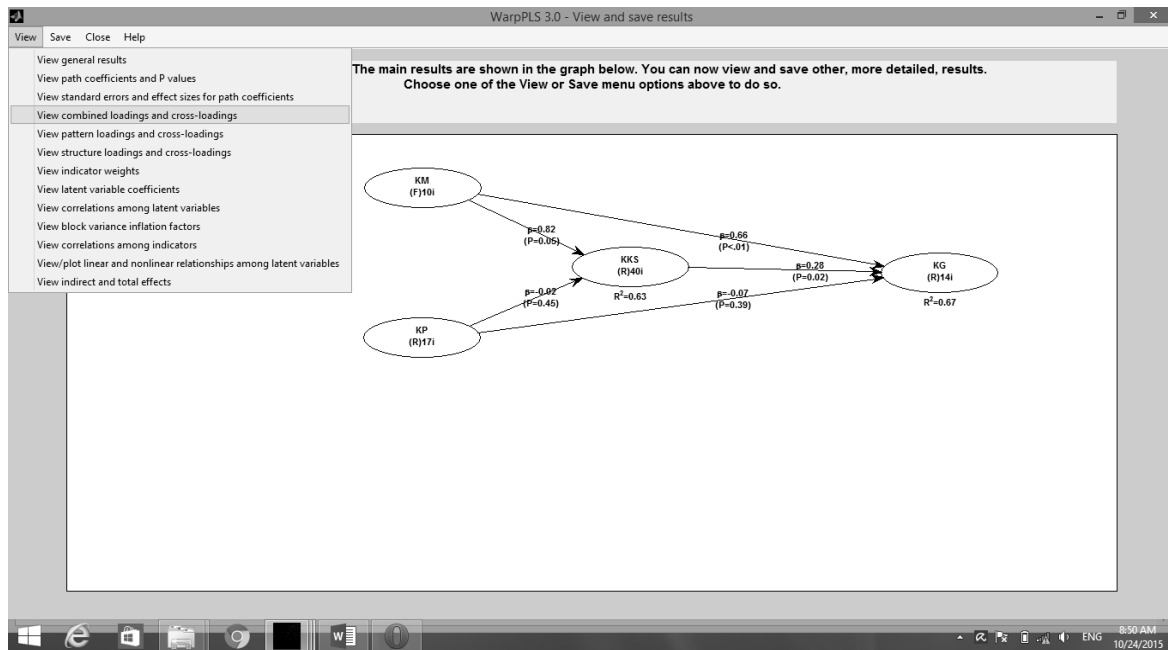
Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan



Khusus untuk uji Validitas

Klik *output combined loadings and cross-loadings*.

Output ini biasanya digunakan peneliti untuk melaporkan hasil pengujian validitas konvergen dari instrument pengukuran (kuesioner). Validitas konvergen merupakan bagian dari measurement model yang dalam SEM-PLS biasanya disebut sebagai *outer model* sedangkan dalam *covariance-based SEM* disebut *confirmatory factor analysis (CFA)*. Output ini menampilkan konstruk pada kolom dan indikator-indikatornya pada baris.



Sebagai contoh, konstruk MK1 diukur dengan indikator MK1-MK10. Loading setiap indikator pada konstruknya diberi tanda kurung serta terdapat nilai p-nya. Contoh *loading* dari *structure matrix* (*unrotated*) dan *cross-loading* dari *pattern matrix* (*rotated*). MK1 pada konstruk MK adalah sebesar 0,726 dan signifikansi dengan nilai p kurang dari 0,001 ($p < 0,001$).

Kriteria :

A. Terdapat dua kriteria untuk menilai apakah *outer model* memenuhi syarat validitas konvergen untuk konstruk reflektif yaitu :

1. Loading harus di atas 0,70
2. Nilai p signifikansi ($< 0,05$) (Hair dkk, 2013).

Syarat *loading* harus di atas 0,70 karena variable laten seharusnya minimal dapat menjelaskan variasi setiap indicator sebesar 50% ($0,70^2$ adalah mendekati 50%). Dalam kasus penelitian biasanya kasus tersebut sulit terpenuhi terlebih untuk kuesioner yang baru dikembangkan. Oleh karena itu, *loading* antara 0,40-0,70 harus tetap dipertimbangkan untuk dipertahankan. Indikator dengan *loading* dibawah 0,40 harus dihapus dari model. Namun untuk indikator dengan *loading* antara 0,40-0,70 sebaiknya kita analisis dampak dari keputusan menghapus indikator tersebut pada *average variance extracted* (AVE) dan *composite reliability*. Kita dapat menghapus indikator dengan *loading* antara 0,40-0,70 jika indikator tersebut dapat meningkatkan *average variance extracted* (AVE) dan *composite reliability* di atas batasannya (*threshold*).

Nilai batasan AVE adalah 0,50 dan *composite reliability* sebesar 0,70. Pertimbangan lain dalam menghapus indikator adalah dampaknya pada validitas isi (*content validity*) konstruk. Indikator dengan *loading* yang kecil kadang tetap dipertahankan karena punya kontribusi pada validitas isi konstruk.

Contoh :

MK1 pada konstruk MK adalah sebesar 0,726 dan signifikansi dengan nilai p kurang dari 0,001 ($p < 0,001$). Dengan syarat tersebut, pengukuran konstruk MK telah memenuhi syarat **validitas konvergen**.

B. Diharapkan *loading* ke konstruk lain (*cross-loadings*) bernilai lebih rendah daripada ke konstruk tersebut, maka hal ini terpenuhinya **validitas diskriminan**.

Contoh :

Konstruk KP = 0,059, KKS = 0,029 dan KG = 0,064 lebih rendah daripada Konstruk MK adalah sebesar 0,726. Hasil *cross-loadings* itu bisa menjadi indikasi terpenuhinya kriteria terpenuhinya **validitas diskriminan**.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

WarpPLS 3.0 - Combined loadings and cross-loadings

Close Help

	KM	KP	KKS	KG	SE	P value
MK1	(0.726)	0.059	0.029	0.064	0.129	<0.001
MK2	(0.771)	-0.181	0.062	-0.014	0.125	<0.001
MK3	(0.664)	-0.219	0.071	-0.277	0.083	<0.001
MK4	(0.786)	-0.202	-0.019	-0.148	0.129	<0.001
MK5	(0.705)	-0.066	-0.193	0.312	0.144	<0.001
MK6	(0.751)	0.010	0.015	0.025	0.144	<0.001
MK7	(0.731)	-0.013	-0.003	0.040	0.151	<0.001
MK8	(0.749)	-0.178	0.109	-0.087	0.120	<0.001
MK9	(0.634)	0.623	0.043	-0.113	0.144	<0.001
MK10	(0.730)	0.257	-0.113	0.183	0.136	<0.001
KP1	0.076	(0.725)	-0.021	0.042	0.133	<0.001
KP2	0.231	(0.721)	-0.048	0.155	0.143	<0.001
KP3	0.075	(0.757)	-0.026	0.144	0.176	<0.001
KP4	0.045	(0.791)	0.007	-0.102	0.159	<0.001
KP5	-0.090	(0.809)	0.132	-0.107	0.144	<0.001
KP6	-0.041	(0.842)	0.013	-0.183	0.128	<0.001
KP7	0.211	(0.787)	0.069	-0.104	0.130	<0.001
KP8	-0.159	(0.818)	0.040	0.099	0.145	<0.001
KP9	-0.017	(0.844)	0.047	-0.063	0.136	<0.001
KP10	-0.065	(0.842)	-0.137	-0.032	0.124	<0.001
KP11	0.046	(0.830)	0.094	-0.093	0.142	<0.001
KP12	-0.005	(0.790)	-0.099	-0.029	0.142	<0.001
KP13	-0.027	(0.784)	-0.000	0.040	0.131	<0.001
KP14	-0.018	(0.836)	0.072	-0.152	0.135	<0.001
KP15	-0.018	(0.856)	-0.045	-0.056	0.117	<0.001
KP16	-0.078	(0.789)	-0.026	0.110	0.169	<0.001
KP17	-0.133	(0.698)	-0.093	0.444	0.114	<0.001
KKS1	-0.475	0.268	(0.765)	0.117	0.118	<0.001

Note: P values < 0.05 are desirable for reflective indicators.

WarpPLS 3.0 - Combined loadings and cross-loadings

Close Help

	KM	KP	KKS	KG	SE	P value
KKS1	-0.475	0.268	(0.765)	0.117	0.118	<0.001
KKS2	-0.431	0.262	(0.805)	0.208	0.101	<0.001
KKS3	-0.350	0.156	(0.762)	0.226	0.121	<0.001
KKS4	-0.254	0.089	(0.778)	0.241	0.110	<0.001
KKS5	-0.076	-0.165	(0.777)	0.324	0.127	<0.001
KKS6	-0.351	0.104	(0.832)	0.176	0.111	<0.001
KKS7	0.030	-0.204	(0.784)	0.191	0.123	<0.001
KKS8	-0.026	-0.186	(0.667)	0.316	0.160	<0.001
KKS9	0.260	-0.261	(0.704)	0.111	0.155	<0.001
KKS10	-0.111	0.113	(0.688)	0.225	0.152	<0.001
KKS11	0.205	-0.122	(0.838)	0.034	0.129	<0.001
KKS12	0.288	-0.223	(0.812)	-0.124	0.123	<0.001
KKS13	0.421	-0.313	(0.819)	-0.192	0.117	<0.001
KKS14	0.412	-0.267	(0.764)	-0.193	0.117	<0.001
KKS15	0.548	-0.347	(0.791)	-0.201	0.131	<0.001
KKS16	0.305	-0.202	(0.709)	-0.012	0.141	<0.001
KKS17	0.429	-0.373	(0.769)	-0.179	0.122	<0.001
KKS18	0.338	-0.392	(0.770)	0.064	0.139	<0.001
KKS19	0.535	-0.337	(0.723)	-0.525	0.156	<0.001
KKS20	0.325	-0.308	(0.706)	-0.168	0.181	<0.001
KKS21	0.203	-0.202	(0.704)	-0.075	0.181	<0.001
KKS22	0.362	-0.257	(0.747)	-0.378	0.130	<0.001
KKS23	-0.330	0.098	(0.636)	0.021	0.167	<0.001
KKS24	-0.337	0.161	(0.665)	0.212	0.093	<0.001
KKS25	-0.502	0.304	(0.732)	0.137	0.110	<0.001
KKS26	-0.272	0.343	(0.722)	-0.077	0.128	<0.001
KKS27	0.036	-0.086	(0.705)	-0.075	0.151	<0.001
KKS28	-0.227	0.274	(0.669)	-0.074	0.159	<0.001

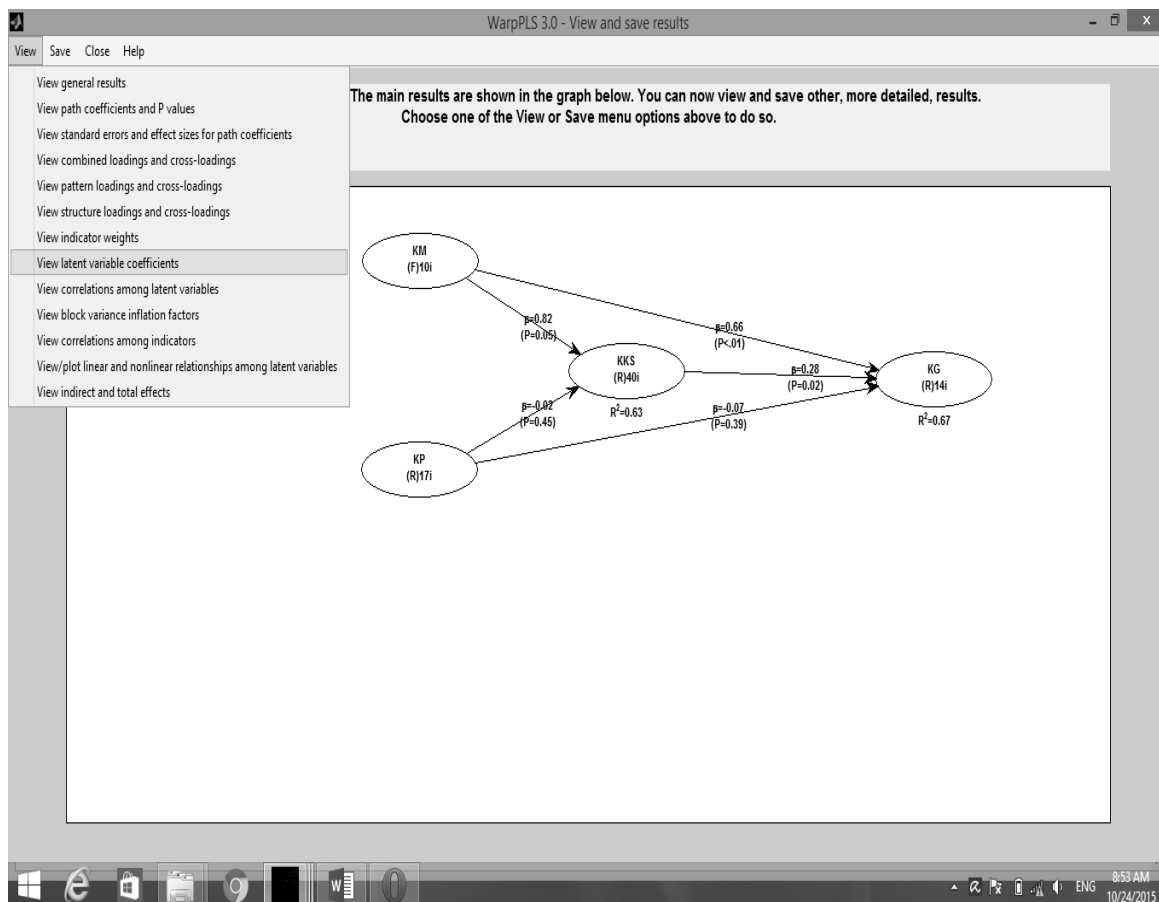
Note: P values < 0.05 are desirable for reflective indicators.

Pengembangan dan Pengujian Instrumen untuk Penelitian Pendidikan

WarpPLS 3.0 - Combined loadings and cross-loadings

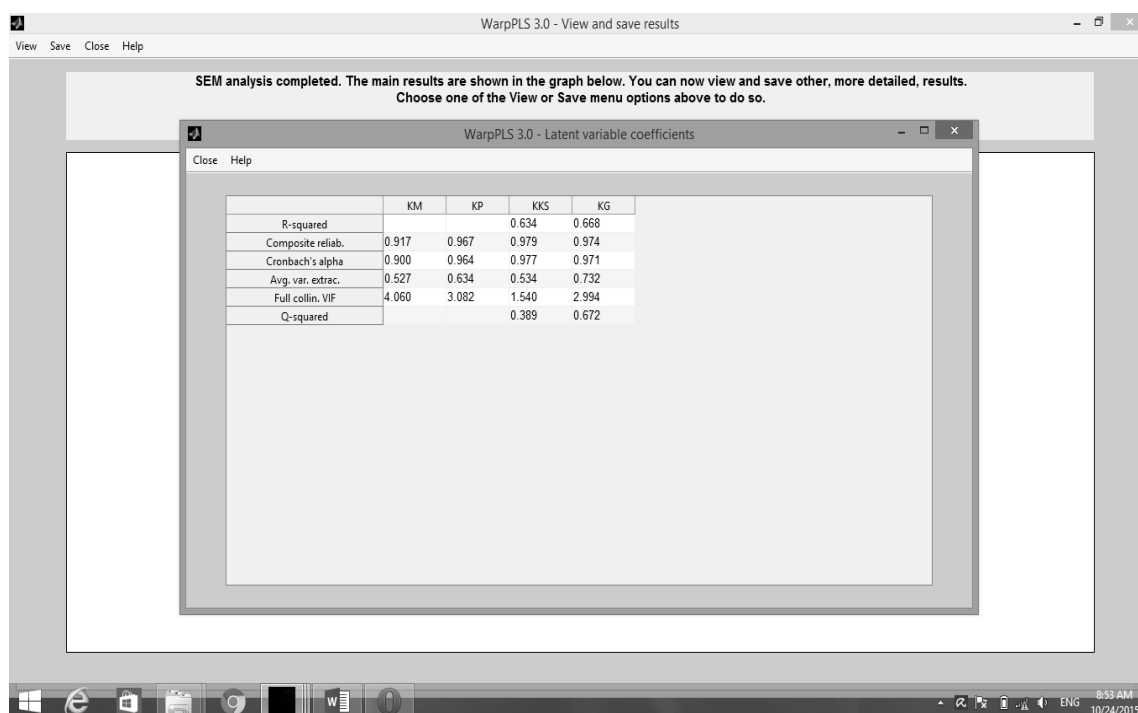
	KM	KP	KKS	KG	SE	P value
KKS27	0.036	-0.086	(0.705)	-0.075	0.151	<0.001
KKS28	-0.227	0.274	(0.669)	-0.074	0.159	<0.001
KKS29	-0.545	0.427	(0.609)	0.071	0.161	<0.001
KKS30	-0.078	0.219	(0.712)	-0.111	0.150	<0.001
KKS31	-0.434	0.419	(0.676)	0.313	0.183	<0.001
KKS32	-0.246	0.298	(0.676)	0.149	0.192	<0.001
KKS33	0.128	0.061	(0.735)	-0.150	0.178	<0.001
KKS34	0.085	0.157	(0.706)	-0.158	0.186	<0.001
KKS35	-0.003	0.178	(0.705)	-0.117	0.184	<0.001
KKS36	0.041	0.110	(0.679)	-0.086	0.197	<0.001
KKS37	-0.511	0.225	(0.642)	0.373	0.160	<0.001
KKS38	-0.164	0.358	(0.686)	-0.115	0.191	<0.001
KKS39	0.278	-0.069	(0.712)	-0.221	0.191	<0.001
KKS40	0.146	0.018	(0.759)	-0.205	0.184	<0.001
KG1	-0.216	0.056	-0.132	(0.878)	0.174	<0.001
KG2	-0.326	0.206	0.028	(0.889)	0.180	<0.001
KG3	-0.305	0.114	-0.029	(0.842)	0.181	<0.001
KG4	-0.172	0.079	0.037	(0.891)	0.185	<0.001
KG5	-0.091	0.021	-0.017	(0.892)	0.173	<0.001
KG6	0.119	-0.044	-0.001	(0.840)	0.192	<0.001
KG7	0.169	-0.029	-0.042	(0.869)	0.156	<0.001
KG8	0.085	-0.032	0.036	(0.842)	0.168	<0.001
KG9	0.192	-0.090	0.036	(0.858)	0.170	<0.001
KG10	-0.336	0.275	0.029	(0.845)	0.171	<0.001
KG11	0.400	-0.327	0.004	(0.909)	0.150	<0.001
KG12	-0.154	0.170	-0.022	(0.854)	0.175	<0.001
KG13	0.096	-0.081	0.060	(0.882)	0.167	<0.001
KG14	0.684	-0.392	0.019	(0.664)	0.238	0.003

Note: P values < 0.05 are desirable for reflective indicators.



Uji Reliabilitas

Klik View pilih *Output Latent Variable Coefficients*



Output seperti gambar diatas merupakan output yang harus dilaporkan dalam tulisan ilmiah. Beberapa hasil yang ditampilkan adalah koefisien determinasi, reliabilitas instrument, validitas diskriminasi, *full collinearity test*, dan validitas prediktif (*predictive validity*).

Koefisien determinasi menggunakan R-squared yang menunjukkan berapa persentase variansi konstruk endogen/criterion dapat dijelaskan oleh konstruk yang dihipotesiskan mempengaruhinya (eksogen/prediktor). Semakin tinggi R-Squared menunjukkan model yang baik. R-Squared hanya ada untuk konstruk endogen. Sebagai contoh, R-Squared konstruk KKS = 0,634 menunjukkan bahwa variansi KKS dapat dijelaskan sebesar 63,4% oleh variansi KM dan KP. KG sebesar 0,668 menunjukkan bahwa variansi KG dapat dijelaskan sebesar 66,8% oleh variansi KKS.

Q-Squared (biasanya juga disebut Stoner-Geisser Coefficient) merupakan ukuran nonparametric yang diperoleh melalui algoritma blindfolding. Q-Squared digunakan untuk penilaian validitas prediktif atau revelansi dari sekumpulan variable laten predictor pada variable kriterion. Q-Squared analog dengan R-Squared namun hanya dapat diperoleh melalui *resampling*. Q-Squared dapat bernilai negative sedangkan nilai R-squared selalu positif. **Model dengan validitas prediktif harus mempunyai nilai Q-squared lebih besar dari nol.** Hasil estimasi model kita menunjukkan validitas prediktif yang baik (yaitu 0,389 dan 0,672) karena diatas nol.

Output juga menyajikan dua ukuran reliabilitas instrument penelitian yaitu composite reliability dan cronbach's alpha. Keduanya harus bernilai di atas 0,70 sebagai syarat reliabilitas (Fornell dan Lacker, 1981, Nunnaly, 1978). Output diatas menunjukkan reliabilitas instrument telah terpenuhi karena diatas 0,70. Average variance extracted (AVE) juga digunakan untuk evaluasi validitas konvergen. Karena harus di atas 0,50 (Fornell dan Lacker, 1981). Output menunjukkan kriteria tersebut telah terpenuhi.

Full collinearity VIF merupakan hasil pengujian kolinieritas penuh yang meliputi multikolinieritas vertical dan lateral. Hasil pengujian full collinearity disajikan untuk setiap variable laten. Kolinieritas yang kita kenal selama ini disebut vertical atau klasik yaitu kolinieritas antarvariabel predictor dalam blok yang sama.

Terdapat konsep baru, yaitu kolinearitas lateral yaitu kolinearitas antara variabel laten predictor dengan kriteria. Kolinearitas lateral sering diabaikan padahal dapat menyebabkan hasil penelitian menjadi bias. Kolinearitas lateral juga digunakan untuk menguji *common method bias*. Kriteria untuk *full collinearity test* adalah nilainya harus lebih rendah dari 3,3 (Kock, 2013). Output diatas menunjukkan nilai *full collinearity VIF* kurang dari 3,3 yaitu KP = 3,082, KKS = 1,540, dan KG= 2,994 sehingga model bebas dari masalah kolinearitas vertical, lateral, dan *common method bias*. Kecuali MK = 4,060 diatas menunjukkan nilai *full collinearity VIF* lebih besar dari 3,3 sehingga model bermasalah kolinearitas vertical, lateral, dan *common method bias*.

**BAB
IX**

DAFTAR PUSTAKA

- Amudi Pasaribu, *Pengantar Statistik*, Medan : Imballo, 1965.
- Anas Sudijono, *Pengantar Evaluasi Pendidik*, PT. Raja Grafindo Persada, Jakarta, 1990.
- Anas Sudijono, *Pengantar Statistik Pendidik*, PT. Raja Grafindo Persada, Jakarta, 1980
- Anonymous. 2009. “Aspek Penilaian dalam KTSP Bag 1 (Aspek Kognitif)”. (Online)<http://massofa.wordpress.com/feed/>. Diakses Tanggal 10 Oktober 2009
- Anonymous. 2009. “Pengembangan Perangkat Penilaian Afektif”. (Online) <http://akhmadsudrajat.wordpress.com/2008/08/15/pengertian-fungsi-dan-mekanisme-penetapan-kriteria-ketuntasan-minimal-kkm/>. Diakses Tanggal 10 Oktober 2009
- Anonymous. 2009. “Pengembnagan Perangkat Penilaian Psikomotor dan Prosedur Penilaian”. (Online) <http://nurmanspd.wordpress.com/2009/09/17/pengembangan-perangkat-penilaian-psikomotor/>. Diakses Tanggal 10 Oktober 2009
- Anonymous. 2009. “Pengukuran Ranah Kognitif, Afektif, dan Psikomotor”. (Online) <http://hadirukiyah.blogspot.com/2009/08/pengukuran-ranah-kognitif-afektif-dan.html>. Diakses Tanggal 10 Oktober 2009
- Anonymous. 2009. “Penilaian Ranah Psikomotorik Siswa”. (Online) <http://delapanratus.blogspot.com/2009/04/penilaian-ranah-psikomotorik-siswa.html>. Diakses Tanggal 10 Oktober 2009
- Anonymous. 2009. “Sistem Penilaian”. (Online) <http://smak.yski.info/>. Diakses Tanggal 10 Oktober 2009
- Anton Dajan, *Pengantar Metode Statistik Deskriptif*, Jakarta LP3ES, 1973.
- Bambang Sumintono dan Wahyu Widhiarso, *Aplikasi Model Rasch untuk Penelitian Ilmu-Ilmu Sosial*, Trim Komunikata Publishing House, Cimahi, 2003.
- Bloom, B. S. ed. et al. *Taxonomy of Educational Objectives: Handbook 1, Cognitive Domain*. New York: David McKay, 1956.

- Bloom, B.S., (Ed). *Taxonomy of Educational Objectives. The Classification of Educational Goals. Handbook I : Cognitive domain*. Longman, New York, 1956.
- Cronbach, Lee, J. *Essentials of Psychological Testing*, Harper & Row, New York, 1960.
- Crow, Lester D. Ph.D. and Alice Crow Ph.,D. *General Psychology*. Littlefield, Adam & Co. 1963.
- Daryanto, Drs. *Evaluasi Pendidikan*, Rineka Cipta, 2000.
- Djarwanto, Ps. Drs, *Statistika Non Parametrik*, Universitas Sebelas Maret, Surakarta, 1999.
- Ghozali, Imam, 2002. *Aplikasi Analisis Multivariate dengan Program SPSS*, Semarang: UPT. Undip.
- Hamzah B. Uno, Herminanto Sofyan, dan I Made Candiasa, *Pengembangan Instrumen Untuk Penelitian*, Penerbit Delima Press, Jakarta, 2001.
- Ismet Basuki dan Hariyanto, *Asesmen Pembelajaran*, Penerbit Rosda Karya, Bandung, 2014.
- Krathwohl c.s., *Taksonomi of Educational Objectives: Book 2*, Longman Group Ltd., London, 1972.
- Krathwohl, (E.d). dalam Robert M. Gagne, Leslie J. Briggs. *Principles of Instructional design, Fourth Edition*, Harcourt Brace Jovanovich College Publeshers, New York, 1992.
- Krathwohl, D. R. ed. et al., *Taxonomy of Educational Objectives: Handbook II, Affective Domain*. New York: David McKay, 1964.
- M. Chabib Thoba. *Teknik Evaluasi Pendidikan*, Rosyda Karya, Bandung, 2000.
- Masrukhin, *Buku Latihan Aplikasi SPSS Statistik Deskriptif dan Inferensial*, Penerbit Media Ilmu Press, Kudus, 2010.
- Masrukhin, *Evaluasi Pendidikan*, Penerbit Media Ilmu Press, Kudus, 2008.
- Masrukhin, M.Pd. *Statistik Deskriptif Berbasis Komputer*, Media Ilmu Kudus, 2007
- Masrukhin, M.Pd. *Statistik Inferensial Aplikasi Program SPSS*, Media Ilmu Kudus, 2007
- Masrukhin, *Metodologi Penelitian Kuantitatif*, Penerbit Media Ilmu Press, Kudus, 2010.
- Masrukhin, *Statistik Deskriptif dan Inferensial Aplkasi Program SPSS*, Penerbit Media Ilmu Press, Kudus, 2014
- Moh. Nazir, *Metode Penelitian*, Ghalia, Jakarta, 1999.
- Muchtar Buchari, *Technik-Technik Evaluasi dalam Pendidikan*, Bapemsi, Bandung, 1967.
- Mulyasa, *Kurikulum Tingkat Satuan Pendidikan*. Rosyda Karya, Bandung, 2006.
- Nana Sudjana, *Dasar-Dasar Proses Belajar Mengajar*. Sinar Baru Algensindo Bandung, 2000.

- Nana Syaodih Sukmadinata. *Landasan Psikologi Proses Pendidikan*. Bandung: P.T. Remaja Rosdakarya. W. Gulo. *Strategi Belajar Mengajar*. Jakarta : Grasind, 2005.
- Ngalim Purwanto, MP. *Prinsip-Prinsip Teknik Evaluasi Pengajaran*, Rosyda Karya, Bandung, 2002.
- Norman E. gronlund & Robert L. Linn, *Measurement and Evaluation in Teaching*. Macmillan Publishing Company, New York, 1990.
- Remmers H.H. and Gade N.L. *A Practical Introduction to Measurement and Evaluation*, Harper & Row, New York, 1960.
- Retno Sriningsih Satmoko, *Statistika Inferensial*, IKIP Semarang Press, 1993
- Retno Sriningsih Satmoko, *Statistika Inferensial*, IKIP Semarang Press, 1993
- Robert B. Ewen, *Introductory Statistics For The Behavioral Seciences*, New York and London University, 1971.
- Simpson, E.J. *The Classifiacation of Educational Objectives : Psychomotor domain*, University of Illinois Press, Urbana, 1996.
- Singgih Santoso, *SPSS Statistik Multivariat*, Elex Media Komputindo, 2002
- Slameto, *Evaluasi Pendidikan*. Bumi Aksara, Jakarta, 2000.
- Soemadi Soeryabrata. *Psycho Diagnostik*. Rakepress Yogyakarta 1973
- Sri Wardani. 2004. *Penilaian Pembelajaran Matematika Berbasis Kompetensi*. Yogyakarta: Departemen Pendidikan Nasional
- Stanley Elan, *Performance Based Teacher Education*. Washington DC : American Association of Colleges for Teacher Education, One Dupont Circle, 1971.
- Stanley, Julien C., *Measurement in Today's Schools*. Printice-Hall, Inc., 1964.
- Sudjana, *Metoda Statistika*, Tarsito Bandung, 1996.
- Sudjana, Nana. 1989. *Penilaian Hasil Proses Belajar*. Bandung: PT. Remaja Rosdakarya Offset
- Sudjono, Anas. 2008. *Pengantar Evaluasi Pendidikan*. Jakarta: PT.RajaGrafindo Persada.
- Sugiyono, dan Eri Wibowo, S.Pd. *Statistika Penelitian dan Aplikasinya Dengan SPSS 10.00 For Windows*, CV. Alfabeta Bandung, 2001.
- Sugiyono, dan Eri Wibowo, S.Pd. *Statistika Penelitian dan Aplikasinya Dengan SPSS 10.00 For Windows*, CV. Alfabeta Bandung, 2001.
- Sugiyono, *Statistika Untuk Penelitian*, CV. Alfa Beta Bandung, 2002.
- Sugiyono, *Statistika Untuk Penelitian*, CV. Alfa Beta Bandung, 2002.
- Suharsimi Arikunto, *Dasar-Dasar Evaluasi PendididkanteK*, Pen. Renika Cipta, Jakarta, 1998.

- Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, PT. Bina Aksara, Jakarta, 1999.
- Suharsimi Arikunto, *Dasar-Dasar Evaluasi Pendidikan*, Jakarta : PT. Bina Aksara, 1999
- Suharsimi Arikunto, *Manajemen Penelitian*, Jakarta : PT. Renika Cipta, 1983.
- Suharsimi Arikunto, *Prosedur Penelitian : Suatu Pendekatan Praktis*, Jakarta : PT. Bina Aksara, 1983.
- Suharsimi Arikunto, *Prosedur Penelitian suatu pendekatan praktek*, Penerbit Renika Cipta, Jakarta, 1998.
- Sutrisno Hadi, *Statistik, Jilid I*, Andi Offset : Yogyakarta, 2000.
- Sholihin Mahfudh dan Retno Dwiyanto. *Analisis WarpPLS 3.0*. Penerbit Andi Yogyakarta, 2013.
- Syaifuddin Azwar, *Dasar-Dasar Psikometri*, Pustaka Pelajar, Yogyakarta, 2001
- Syaifuddin Azwar, *Reliabilitas dan Validitas*, Pustaka Pelajar, Yogyakarta, 2000.
- Thorndike, Robert L, and Elisabeth Hogen, *Measurement and Evaluation in Psychology and Education*. John Willey & Sons, Inc. New York, 1961.
- Vardiansyah, Dani. *Filsafat Ilmu Komunikasi: Suatu Pengantar*, Indeks, Jakarta 2008.
- Wandt, Edwin and Gerald W. Brown. *Essential of Educational Evaluation*. Holt Rinehart and Winston, New York, 1957.
- Wijaya IR, *Statistika Non Parametrik (Aplikasi Program SPSS)*, Alfa Beta Bandung, 2000.
- Witherington, *Psychologi Pendidikan*, (Terjemahan). Mochtar Buchori, M.Ed. Bapemsi, Bandung, 1967.
- Woodwart, Robert S. And Donald G. Marquis, *Psychology*, Holt Rinehart and Winston, Inc, New York, 1961.
- Zaenal Arifin, *Evaluasi Pembelajaran Program Peningkatan Kualifikasi Guru Madrasah dan Guru PAI pada Sekolah*, Dirjen Pendidikan Islam Depag RI, 2009.
- Zakiah Daradjat, *Metodik Khusus Pengajaran Agama Islam*. Bumi Aksara Jakarta, 1995.

**LAMPIRAN
DAFTAR TABEL**

- 1. Table t-test**
- 2. Tabel r Product Moment**
- 3. Tabel Chi Square (χ^2)**
- 4. Tabel Rho (Spearman)**

Tabel 1.
TABLE T-TEST

dk	Harga Kritik t-test Pada Taraf Signifikansi				
	5%	1%	dk	5%	1%
1	12.71	63.66	24	2.06	2.80
2	4.30	9.92	25	2.06	2.79
3	3.18	5.84	26	2.06	2.78
4	2.78	4.60	27	2.05	2.77
5	2.57	4.03	28	2.05	2.76
6	2.45	3.71	29	2.04	2.76
7	2.36	3.50	30	2.04	2.75
8	2.31	3.36	35	2.03	2.72
9	2.26	3.25	40	2.02	2.71
10	2.23	3.17	45	2.02	2.69
11	2.20	3.11	50	2.01	2.68
12	2.18	3.06	60	2.00	2.65
13	2.16	3.01	70	2.00	2.65
14	2.14	2.98	80	1.99	2.64
15	2.13	2.95	90	1.99	2.63
16	2.12	2.92	100	1.98	2.63
17	2.11	2.90	125	1.98	2.62
18	2.10	2.88	150	1.98	2.61
19	2.09	2.86	200	1.97	2.60
20	2.09	2.84	300	1.97	2.59
21	2.08	2.83	400	1.97	2.59
22	2.07	2.82	500	1.96	2.59
23	2.07	2.81	1000	1.96	2.58

Tabel 2
Tabel r Product Moment

dk=N	5%	1%	dk =N	5%	1%
3	0.997	0.999	38	0.32	0.413
4	0.95	0.99	39	0.316	0.408
5	0.878	0.959	40	0.312	0.403
6	0.811	0.917	41	0.308	0.398
7	0.754	0.874	42	0.304	0.393
8	0.707	0.834	43	0.301	0.389
9	0.666	0.798	44	0.297	0.384
10	0.632	0.765	45	0.294	0.38
11	0.602	0.735	46	0.291	0.376
12	0.576	0.708	47	0.288	0.372
13	0.553	0.684	48	0.284	0.368
14	0.532	0.661	49	0.281	0.364
15	0.514	0.641	50	0.279	0.361
16	0.497	0.623	55	0.266	0.345
17	0.482	0.606	60	0.254	0.33
18	0.468	0.59	65	0.244	0.317
19	0.456	0.575	70	0.235	0.306
20	0.444	0.561	75	0.227	0.296
21	0.433	0.549	80	0.22	0.286
22	0.423	0.537	85	0.213	0.278
23	0.413	0.526	90	0.207	0.27
24	0.404	0.515	95	0.202	0.263
25	0.396	0.505	100	0.195	0.256
26	0.388	0.496	125	0.176	0.23
27	0.381	0.487	150	0.159	0.21
28	0.374	0.478	175	0.148	0.194
29	0.367	0.47	200	0.138	0.181
30	0.361	0.463	300	0.113	0.148
31	0.355	0.456	400	0.098	0.128
32	0.349	0.449	500	0.088	0.115
33	0.344	0.442	600	0.08	0.105
34	0.339	0.436	700	0.074	0.097
35	0.334	0.43	800	0.07	0.091
36	0.329	0.424	900	0.065	0.086
37	0.325	0.418	1000	0.062	0.081

Tabel 3. Tabel Chi Square (χ^2)

Dk	Taraf Signifikansi	
	5%	1%
1	3.841	6.635
2	5.991	9.21
3	7.815	11.345
4	9.488	13.227
5	11.07	15.086
6	12.592	16.812
7	14.067	18.475
8	15.507	20.09
9	16.919	21.666
10	18.307	23.209
11	19.675	24.275
12	21.026	26.217
13	22.362	27.688
14	23.685	29.141
15	24.996	30.578
16	26.296	32
17	27.587	33.409
18	28.869	34.805
19	30.144	36.191
20	31.41	37.566
21	32.617	38.932
22	33.924	40.289
23	35.172	41.638
24	36.145	42.98
25	37.652	44.314
26	38.885	45.642
27	40.113	46.963
28	41.337	48.278
29	42.557	49.588
30	43.778	50.892

Tabel 4. Tabel Rho (Spearman)

Dk	TarafSignifikansi	
	5%	1%
5	1	
6	0.886	1
7	0.786	0.929
8	0.738	0.881
9	0.683	0.833
10	0.648	0.794
12	0.91	0.777
14	0.544	0.715
16	0.506	0.665
18	0.475	0.625
20	0.45	0.591
22	0.428	0.562
24	0.409	0.537
26	0.392	0.515
28	0.377	0.496
30	0.364	0.478