

## Uji Validitas Alat Ukur Ketidakjujuran Akademik

Fayna Faradiena

Institut Asesmen Indonesia, Indonesia

faradienaf@gmail.com

### Abstract

*It cannot be denied that academic dishonesty is one of the problems that happened amongst students all over the world, this construct includes behavior that is intentionally carried out to fulfill assignments or academic assessments. This study aims to examine the validity of the construct of academic dishonesty which was adapted from the Academic Dishonesty Scale (McCabe & Trevino, 1993) and Academic Dishonesty Instrument (Iyer & Eastman, 2008) using the Confirmatory Factor Analysis (CFA) method. In addition, this study produced three standardized forms of academic dishonesty scale, namely: long-form (20-item), ideal-form (11-item), and short-form (7-item). Regarding the research sample, this finding included 355 students from all faculties at the State Islamic University of Syarif Hidayatullah Jakarta using snowball non-probability sampling method.*

**Keywords:** *academic dishonesty; measurement; confirmatory factor analysis*

### Abstrak

Ketidakjujuran akademik adalah salah satu masalah yang banyak terjadi di kalangan mahasiswa, hal ini mencakup perilaku yang dengan sengaja dilakukan untuk memenuhi penugasan maupun penilaian akademik. Adapun tujuan dari penelitian ini adalah untuk menguji validitas konstruk ketidakjujuran akademik yang diadaptasi dari alat ukur *Academic Dishonesty Scale* (McCabe & Trevino, 1993) dan *Academic Dishonesty Instrument* (Iyer & Eastman, 2008) melalui metode Confirmatory Factor Analysis (CFA) menggunakan *software* Mplus 8. Selain dilakukan uji validitas, penelitian ini menghasilkan tiga bentuk alat ukur ketidakjujuran akademik yang dibakukan, yaitu bentuk yang valid namun kurang ideal (20-item), bentuk yang ideal (11-item), dan bentuk ringkas (7-item). Sampel penelitian ini adalah 355 mahasiswa aktif dari berbagai Fakultas di Universitas Islam Negeri Syarif Hidayatullah Jakarta menggunakan teknik *snowball non-probability sampling*.

**Kata kunci:** ketidakjujuran akademik; pengukuran; confirmatory factor analysis

## Pendahuluan

Ketidajjuran akademik merupakan perilaku tidak jujur yang disengaja dalam rangka memenuhi atau mengerjakan tugas akademik (Gitanjali, 2004). Konsep ketidajjuran akademik lainnya juga dikemukakan oleh Iyer dan Eastman (2008) bahwa ketidajjuran akademik meliputi empat hal, yaitu: perilaku menyontek, bantuan dari luar, plagiarisme, dan menyontek menggunakan *gadget* pada saat tes (*electronic cheating*). Sedangkan ketidajjuran akademik menurut McCabe dan Trevino (1993) hanya terdiri dari perilaku menyontek dan plagiarisme. Dalam penelitian ini, konsep ketidajjuran akademik yang digunakan adalah berdasarkan hasil penelitian dari Iyer dan Eastman. Hal ini dikarenakan perilaku ketidajjuran akademik yang didefinisikan cukup menggambarkan ketidajjuran akademik yang terjadi di kalangan mahasiswa di Indonesia. Adanya perilaku ketidajjuran akademik yang terjadi di kalangan mahasiswa tentu sangat mengkhawatirkan, hal ini didasarkan penelitian Karassavidou dan Glaveli (dalam Iyer & Eastman, 2008) bahwa mahasiswa yang melakukan ketidajjuran akademik cenderung untuk melakukan perilaku-perilaku ketidajjuran lainnya (*unethical behavior*) di masa mendatang. Maka dari itu, untuk mendeteksi apakah mahasiswa melakukan ketidajjuran akademik atau tidak sebagai dasar evaluasi pendidikan adalah penting.

Untuk mendeteksi ketidajjuran akademik secara ilmiah, tentu diperlukan alat ukur yang mengukur tingkat tinggi rendahnya mahasiswa dalam melakukan ketidajjuran akademik. Selanjutnya, alat ukur yang digunakan haruslah bersifat unidimensional (hanya mengukur satu faktor saja) artinya alat ukur tersebut hanya mengukur konstruk ketidajjuran akademik. Dalam penelitian ini, alat ukur yang diujikan validitasnya merupakan modifikasi dan kombinasi antara *Academic Dishonesty Scale* (McCabe & Trevino, 1993) dan *Academic Dishonesty Instrument* (Iyer & Eastman, 2008) serta beberapa item yang disesuaikan kondisinya di Indonesia agar tidak terjadi *bias* budaya. Setelah dilakukan modifikasi, alat ukur ketidajjuran akademik ini terdiri dari 20 item dengan model 4 skala *Likert* (4 = pernah > 2 kali; 3 = pernah 2 kali; 2 = pernah 1 kali; 1 = tidak pernah) dan bersifat unidimensional.

Selanjutnya, hipotesis mayor penelitian pengukuran ketidajjuran akademik berbunyi “*seluruh butir soal (item) yang dibuat dalam skala pengukuran ketidajjuran akademik yang dibuat memang mengukur satu faktor saja*” dan hipotesis minor penelitian pengukuran ketidajjuran akademik yang berbunyi “*setiap butir soal (item) signifikan dalam mengukur ketidajjuran akademik*”.

## Uji Validitas Alat Ukur Ketidajjuran Akademik

Dalam melakukan uji validitas alat ukur ketidajjuran akademik, peneliti menggunakan metode *Confirmatory Factor Analysis* (CFA) menggunakan *software* Mplus 8.0 (Muthen&Muthen, 2017). Adapun logika dalam CFA menurut Umar (2015) adalah sebagai berikut :

1. Terdapat sebuah konsep atau *trait* berupa kemampuan yang didefinisikan secara operasional, sehingga dapat di susun pertanyaan atau pernyataan untuk mengukur kemampuan tersebut, kemampuan ini disebut faktor. Untuk melakukan pengukuran pada setiap faktor harus dilakukan analisis terlebih dahulu terhadap respon atas item-itemnya.
2. Dalam hal ini diteorikan bahwa semua item mengukur hanya satu faktor yang sama (unidimensional).
3. Selain itu, juga diteorikan bahwa masing-masing item hanya mengukur satu faktor saja. Ini berarti bahwa setiap item bersifat unidimensional pada dirinya sendiri (tidak mengandung bias).
4. Jika data sudah tersedia, maka data tersebut digunakan untuk mengestimasi parameter dari model unidimensional yang diteorikan. Dalam hal ini parameternya adalah koefisien muatan faktor ( $\lambda$ ) dan varians kesalahan pengukuran ( $\Theta$ ). Karena model bersifat unidimensional, parameter yang berupa varians dari konstruk yang diukur (faktor) tidak perlu diestimasi, karena ditetapkan ( $fixed = 1$ ). Hal

ini diperlukan untuk menentukan skala ukuran bagi faktor yang hendak diukur, yang disini berupa skala baku.

5. Nilai hasil estimasi parameter digunakan untuk memprediksi korelasi yang seharusnya diperoleh menurut teori, korelasi ini disebut *sigma* ( $\sigma$ ).
6. Matriks nilai korelasi hasil prediksi ( $\Sigma$ ) kemudian dibandingkan dengan matriks korelasi yang diperoleh dari data (S). Jika model teoritis yang diuji (unidimensional) sesuai dengan model teori tersebut memang benar adanya karena didukung oleh data. Dalam hal ini, yang dilakukan secara statistik " $H_0 = S - \Sigma = 0$ ". Artinya, tidak ada perbedaan antara matriks korelasi yang diharapkan oleh teori dengan matriks korelasi yang diperoleh dari data.
7. Ada terdapat banyak indeks statistik yang bisa digunakan untuk menguji hipotesis nihil ( $H_0 = S - \Sigma = 0$ ) diantaranya yang paling utama adalah *chi-square* dan RMSEA. Untuk nilai *chi-square*, model dikatakan fit jika memiliki *probability* yang lebih besar dari 5% ( $p > 0.05$ ). Sedangkan untuk RMSEA, biasanya *software* yang digunakan (Mplus) menyajikan tiga kriteria, model dikatakan fit jika dua dari tiga hasil RMSEA, yaitu: a.) RMSEA *estimate*  $< 0.05$ , b.) *Confident interval* (CI)  $< 0.05$ , dan c.) *probability* bahwa RMSEA  $< 0.05$  adalah  $> 0.05$  ( $P \text{ RMSEA} < 0.05 \text{ adalah } > 0.05$ ). Jika nilai *chi-square* dan/atau RMSEA tidak signifikan, maka disimpulkan bahwa hipotesis nihil tidak ditolak. Artinya model teori bahwa seluruh item memang hanya menguji satu dimensi (unidimensional) adalah benar dan dapat diterima. Hal ini disebut bahwa model fit dengan data dan kegiatannya dinamakan "uji model fit".
8. Jika model unidimensional dinyatakan fit dengan data, maka dapat disimpulkan bahwa instrumen (alat ukur) yang terdiri dari himpunan item tersebut adalah valid untuk mengukur konstruk yang didefinisikan (validitas konstruk terbukti).
9. Jika model unidimensional tidak fit, maka biasanya diperlukan modifikasi model dengan cara menambah parameternya yang dalam hal ini terutama adalah korelasi antar kesalahan pengukuran (residual).
10. Apabila model unidimensional fit dengan data, tetapi disertai dengan korelasi antar residual, maka dapat ditafsirkan bahwa item-item tersebut bersifat multidimensional pada dirinya masing-masing. Misalnya, dua buah item yang selain mengukur kemampuan matematika ternyata juga mengukur kemampuan berbahasa (karena kedua item berbentuk soal cerita). Item yang bersifat multidimensional seperti ini disebut mengandung bias.
11. Jika model fit telah diperoleh, langkah selanjutnya adalah melihat apakah setiap item secara signifikan mengukur konstruk yang hendak diukur. Dalam hal ini terdapat dua kriteria utama dalam menentukan validitas item, yaitu: a.) Apakah koefisien muatan faktor (*factor loading*) bernilai positif, dan b.) Apakah koefisien muatan faktor signifikan secara statistik ( $T\text{-value} > 1,96$  atau  $P\text{-value} < 0.05$ ). Jika koefisien muatan faktor bernilai negative, maka item harus di-*drop* (tidak dapat digunakan) meskipun secara statistik signifikan. Jika koefisien muatan faktor bernilai positif dan signifikan, berarti item valid dan dapat digunakan. Apabila item bernilai positif tetapi tidak signifikan, dapat diserahkan kepada peneliti apakah akan digunakan atau tidak. Namun demikian, untuk pengembangan tes psikologi yang baku sebaiknya tidak digunakan.

## Metode

### Subjek Penelitian

Pada penelitian ini, peneliti menggunakan data mentah (*raw score*) dari skripsi dengan jumlah subyek 355 mahasiswa aktif UIN Syarif Hidayatullah Jakarta. Adapun sampel yang diteliti merupakan mahasiswa dari 12 fakultas, yaitu: Fakultas Ilmu Tarbiyah dan Keguruan, Fakultas Dirasat Islamiyah, Fakultas Ekonomi dan Bisnis, Fakultas Sains dan Teknologi, Fakultas Ushuluddin, Fakultas Syariah dan Hukum, Fakultas Adab dan Humaniora, Fakultas Ilmu Dakwah dan Komunikasi, Fakultas Ilmu Sosial dan Ilmu Politik, Fakultas Psikologi, Fakultas Ilmu Kesehatan, dan Fakultas Kedokteran yang berusia 18 hingga 27 tahun ( $M = 20.57$ ,  $SD=1.310$ ). Metode pengambilan sampel dalam penelitian ini menggunakan *non-probability snowball-sampling*.

### Alat Ukur

Variabel ketidakjujuran akademik (KA) dalam penelitian ini diukur menggunakan modifikasi dari dua alat ukur, yaitu *Academic Dishonesty Scale* (McCabe & Trevino, 1993) dan *Academic Dishonesty Instrument* (Iyer & Eastman, 2008). McCabe dan Trevino (1993) mengembangkan alat ukur KA yang terdiri dari 12 item menggunakan skala Likert dengan empat pilihan jawaban, yaitu 1 = tidak pernah sampai dengan 4 = sangat sering. Sedangkan, instrumen yang dikembangkan oleh Iyer dan Eastman (2008) terdiri dari 17 item dengan lima alternatif jawaban menggunakan skala *Likert*, yaitu tidak pernah, jarang, kadang-kadang, sering, dan selalu.

Selanjutnya, peneliti melakukan modifikasi dari kedua alat ukur tersebut dengan menghapus item yang memiliki kesamaan. Selanjutnya, peneliti menambahkan item berdasarkan hasil pengamatan tentang perilaku ketidakjujuran akademik oleh mahasiswa di Indonesia yang tidak tercantum pada dua alat ukur sebelumnya, yaitu *menitipkan absen pada teman yang masuk*. Sehingga didapatkan *blue print* alat ukur ketidakjujuran akademik yang tercantum pada tabel 1 dan 20 item alat ukur ketidakjujuran akademik pada tabel 2, sedangkan skor pengukuran ketidakjujuran akademik menggunakan skala frekuensi yang tertera pada tabel 3.

Tabel 1 *Blueprint* skala ketidakjujuran akademik

No	Dimensi	Indikator	No Item ( <i>Favorable</i> )	Jumlah
I	Ketidakjujuran akademik	Menyontek secara manual ataupun bantuan <i>gadget</i>	1, 2, 6, 18, 19	5
		Bekerja sama dalam berbuat ketidakjujuran	3, 5, 7, 10, 11, 14, 15	7
		Melakukan pemalsuan (daftar pustaka, data, dan informasi)	9, 13, 16, 20	4
		Menyalin pekerjaan dan/atau menggunakan pendapat orang lain tanpa sepengetahuannya	4, 8, 12, 17	4

Tabel 2. Instrumen Ketidakjujuran Akademik 20-item

Nomor Item	Pernyataan
1	Ketika ujian saya menggunakan contekan berupa catatan kecil
2	Saat ujian saya mencoba untuk menyontek pekerjaan teman yang duduknya berdekatan
3	Mencoba mencari bocoran soal sebelum ujian dilaksanakan
4	Menyalin tulisan mahasiswa lain tanpa sepengetahuan mereka
5	Mencoba menyalin pekerjaan teman dengan seizinnya
6	Mencari contekan jawaban apapun caranya
7	Membantu teman dalam menjawab ujian
8	Menyalin tulisan orang lain dan menyerahkannya seakan itu buatan saya sendiri
9	Memasukkan daftar bacaan pada daftar pustaka meskipun tidak membacanya
10	Meminta orang lain untuk menyelesaikan tugas-tugas perkuliahan
11	Bekerja sama dalam menyelesaikan tugas yang seharusnya individual
12	Menyalin beberapa kalimat dari buku/jurnal tanpa mencantumkan sumbernya
13	Mencoba mendekati dosen agar merubah nilai yang diberikan lebih baik
14	Meminta bantuan orang lain yang lebih pandai untuk memeriksa tugas/tulisan saya
15	Mencari informasi tentang soal ujian kepada mahasiswa lain yang sudah menempuhnya
16	Meminta dicantumkan nama pada tugas kelompok, padahal tidak ikut mengerjakan
17	Meminta pendapat orang lain, namun digunakan seakan pendapat sendiri
18	Mencari jawaban ujian melalui bantuan <i>smartphone</i>
19	Mengambil gambar (memotret) soal ujian secara sembunyi-sembunyi
20	Menitipkan absen pada teman yang masuk

Tabel 3. Skor Pengukuran Skala Frekuensi

Alternatif Pilihan Jawaban Skala Frekuensi	Pernyataan	
	<i>Favorable</i>	<i>Unfavorable</i>
Pernah lebih dari 2x	4	1
Pernah 2x	3	2
Pernah 1x	2	3
Tidak Pernah	1	4

## Hasil Penelitian

### Hasil Uji Validitas Konstruk

Dalam penelitian ini, dilakukan analisis faktor konfirmatori dari 20 item yang mengukur ketidakjujuran akademik. Skala pengukuran ketidakjujuran akademik dalam penelitian ini fit setelah dilakukan iterasi sebanyak 30, 000 kali, dan menghasilkan RMSEA *estimate* = 0.058 ( $p < 0.05$ ), 90 percent C.I = 0.050 sampai 0.066 ( $p < 0.05$ ), dan RMSEA *probability*  $< 0.05 = 0.053$  ( $p > 0.05$ ). Adapun koefisien muatan factor (*factor loading*) untuk setiap item pengukuran ketidakjujuran akademik disajikan dalam tabel 4.

Tabel 4. Koefisien Muatan Faktor Item Ketidakjujuran Akademik

Nomor Item	Koefisien	S.E.	T-value	P-value	Signifikan
1	0.730	0.032	22.788	0.000	✓
2	0.769	0.030	25.770	0.000	✓
3	0.650	0.037	17.336	0.000	✓
4	0.559	0.067	8.392	0.000	✓
5	0.698	0.036	19.236	0.000	✓
6	0.742	0.036	20.348	0.000	✓
7	0.724	0.039	18.583	0.000	✓
8	0.477	0.061	7.763	0.000	✓
9	0.440	0.053	8.339	0.000	✓
10	0.398	0.071	5.605	0.000	✓
11	0.461	0.049	9.316	0.000	✓
12	0.375	0.055	6.796	0.000	✓
13	0.376	0.079	4.788	0.000	✓
14	0.432	0.052	8.317	0.000	✓
15	0.613	0.041	14.895	0.000	✓
16	0.229	0.087	2.636	0.008	✓
17	0.415	0.070	5.975	0.000	✓
18	0.762	0.033	22.769	0.000	✓
19	0.618	0.052	11.950	0.000	✓
20	0.557	0.046	12.200	0.000	✓

#### Keterangan:

S.E : Standard error dari *factor loading*

T-Value : Nilai *t-test*

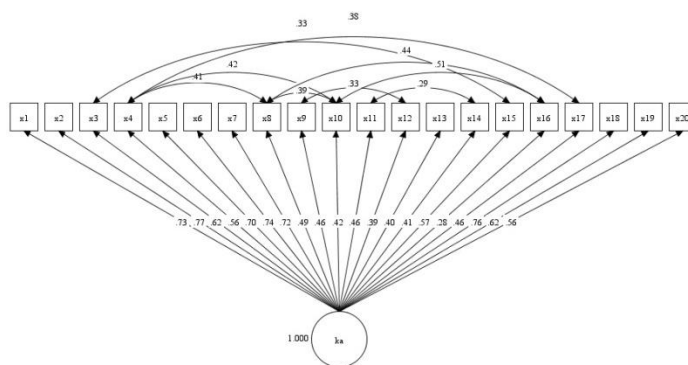
P-Value : Nilai *probability*

Berdasarkan tabel 4 dapat diketahui bahwa tidak ada item yang didrop pada skala ketidakjujuran akademik, karena seluruh item memiliki T-value > 1.96 (signifikan) dan bermuatan positif. Semua item pada variabel ini

telah memenuhi kriteria model fit, sehingga dapat diikutsertakan pada analisis berikutnya. Hasil analisis konstruk ketidakjujuran akademik (KA) menunjukkan model yang tidak fit ( $\chi^2 = 508.945$ ,  $df = 170$ ,  $p < 0.000$ ,  $RMSEA_{estimate} = 0.075$ ,  $90\% C.I = 0.068$  sampai  $0.082$ ,  $probability RMSEA < 0.05 = 0.000$ ,  $CFI = 0.889$ , dan  $TLI = 0.876$ ). Karena baik *chi-square* ( $\chi^2$ ) maupun RMSEA adalah signifikan, hal ini menandakan bahwa model unidimensional (satu faktor) tidak fit dengan data. Ini berarti bahwa sekurangnya ada beberapa item yang mengukur konstruk lain selain ketidakjujuran akademik. Namun demikian, tidak fitnya model ini bisa juga berarti bahwa seluruh butir soal sebenarnya valid mengukur KA (satu dimensi yang hendak diukur, tetapi ada beberapa butir soal yang bersifat multidimensi).

Artinya, terdapat butir soal yang selain mengukur KA juga mengukur faktor lain yang dalam hal ini belum diketahui. Hal ini dikenal juga sebagai bias pada butir soal. Dalam kasus ini, mungkin terdapat beberapa butir soal yang “*error of measurement*”-nya saling berkorelasi karena butir-butir tersebut selain mengukur KA juga mengukur hal lain yang belum diketahui. Oleh sebab itu, langkah berikutnya yang peneliti lakukan adalah melakukan modifikasi terhadap model unidimensional yang diuji dengan cara membiarkan korelasi antar kesalahan pengukuran tersebut menjadi parameter bebas. Hal ini dilakukan satu per satu sampai diperoleh model fit. Adapun cara untuk mengetahui butir mana yang bersifat multidimensional adalah dengan melihat besarnya residual pada butir tersebut. Namun, melalui perangkat lunak Mplus terdapat metode yang lebih mudah dan efisien untuk mendeteksi butir soal yang bersifat multidimensional, yaitu dengan memberikan indeks yang disebut “*modification index*”. Korelasi antar residual pada butir soal yang memiliki koefisien *modification index* paling besar adalah petunjuk terhadap korelasi antar residual yang harus dibebaskan.

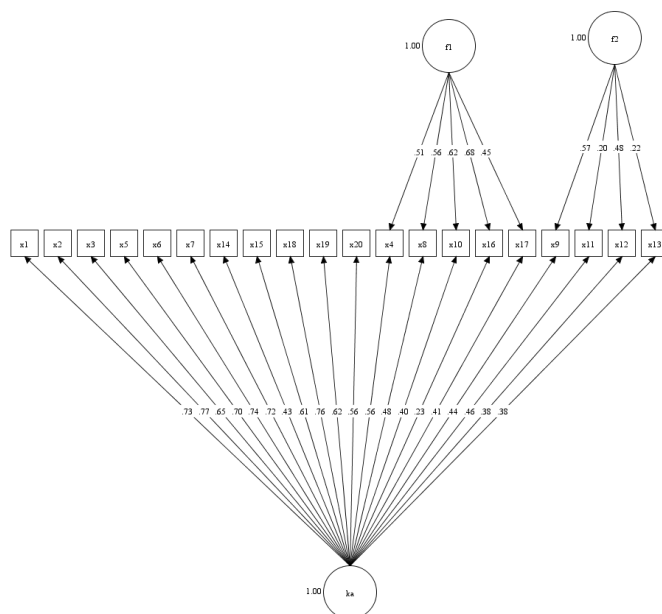
Sebagai hasil modifikasi dengan menggunakan CFA, diperoleh model unidimensional yang fit dengan data ( $\chi^2 = 344.981$ ,  $df = 161$ ,  $p < 0.000$ ,  $RMSEA_{estimate} = 0.057$ ,  $90\% C.I = 0.048$  sampai  $0.065$ ,  $probability RMSEA < 0.005 = 0.088$ ). Jika hanya melihat *chi-square*, model memang belum fit, tetapi jika melihat RMSEA dan koefisien lain seperti CFI dan TLI ( $p > 0.90$ ) maka model sudah fit dengan data (CFI = 0.940, TLI= 0.929). Seperti yang telah banyak dikemukakan, misalnya Jöreskog dan Sörbom (2004) disebutkan bahwa penggunaan *chi-square* untuk menguji model fit kurang tepat jika sampel cukup besar karena nilai *chi-square* adalah hasil perkalian langsung antara besarnya sampel dengan *log likelihood* yang dihasilkan. Dengan demikian, jika menggunakan ukuran *chi-square* sebagai kriteria, maka model yang sebetulnya sangat baik serta fit dengan data menjadi dianggap tidak fit hanya karena sampelnya besar. Namun, di sisi lain jika sampel tidak besar maka estimasi terhadap parameter dengan menggunakan metode *maximum likelihood* menjadi kurang bisa dipercaya. Oleh sebab itu, dalam penelitian ini peneliti lebih menggunakan RMSEA sebagai kriterianya. Adapun model yang fit dengan data disini adalah model unidimensional yang disertai dengan sembilan buah korelasi antar residual seperti pada diagram di bawah berikut:



Gambar I. Model Unidimensional yang Fit dengan Data namun disertai Sembilan Korelasi antar Residual

Untuk mengetahui faktor apa yang diukur oleh kesembilan item yang mengandung bias tersebut, peneliti melakukan analisis CFA dengan model *bi-factor*. Dalam hal ini, seluruh item diteorikan tetap mengukur

konstruk yang hendak diukur yaitu KA. Namun, ada sembilan item diantaranya yang pada saat bersamaan juga mengukur konstruk lain yang tidak berkorelasi dengan KA. Peneliti mencoba dengan model dimana konstruk yang lain itu hanya satu buah (sembilan item yang mengandung bias hanya mengukur satu faktor saja). Artinya, hanya ada satu jenis bias. Namun, hasilnya tetap menghasilkan model yang tidak fit dengan data. Kemudian, dengan mengelompokkan sembilan item yang mengandung bias tersebut menjadi dua kelompok, yang berarti bahwa ada dua konstruk lain yang diukur selain mengukur KA, maka akhirnya peneliti mendapatkan model *bi-factor* yang fit dengan data seperti terlihat pada gambar 2 berikut:



Gambar 2. Diagram Skala Pengukuran Ketidajujuran Akademik 20 Item Menggunakan Pendekatan *Bi-factor*

Berdasarkan gambar 2 di atas, model ini memiliki  $\chi^2 = 353.528$ ,  $df = 161$ ,  $p < 0.000$ ,  $RMSEA_{estimate} = 0.058$ ,  $90\% C.I = 0.50$  sampai  $0.066$ ,  $probability RMSEA < 0.005 = 0.053$ ,  $CFI = 0.937$ , dan  $TLI = 0.925$ . Dari gambar di atas juga terlihat bahwa ada dua faktor (konstruk) lain yang diukur oleh sembilan item yang mengandung bias di atas. Konstruk yang pertama diukur oleh item nomor 4, 8, 10, 16, dan 17. Setelah peneliti menganalisis isi dari kelima butir tersebut, ternyata semua pernyataan berkaitan dengan suatu konstruk yang menurut peneliti adalah keinginan untuk diakui (*need for recognition*). Sebagai contoh satu diantaranya berbunyi “*meminta dicantumkan nama pada tugas kelompok, padahal tidak ikut mengerjakan*”.

Sedangkan konstruk yang kedua diukur oleh item nomor 9, 11, 12, dan 13 yang setelah dianalisis ternyata isinya adalah berkenaan dengan kemampuan melakukan manipulasi. Artinya, mereka yang dapat skor tinggi pada butir-butir ini selain menunjukkan ketidajujuran akademik juga menunjukkan kemampuan yang tinggi dalam melakukan manipulasi. Sebagai ilustrasi, salah satu itemnya berbunyi “*memasukkan daftar bacaan pada daftar pustaka meskipun tidak membacanya*”. Untuk model *bi-factor* yang telah fit ini, dilakukan uji signifikan terhadap koefisien muatan faktor dari setiap butir yang ada. Tujuannya adalah untuk menguji hipotesis apakah setiap butir soal secara statistik signifikan dalam mengukur suatu konstruk yang diukur, yaitu ketidajujuran akademik. Ternyata, seluruh item signifikan ( $p < 0.01$ ) sehingga tidak ada item yang di-drop. Adapun koefisien muatan faktor untuk masing-masing item disertai dengan uji statistika dapat dilihat pada tabel 1.5, sedangkan kurva karakteristik soal (ICC) untuk setiap butir disajikan pada gambar 3.

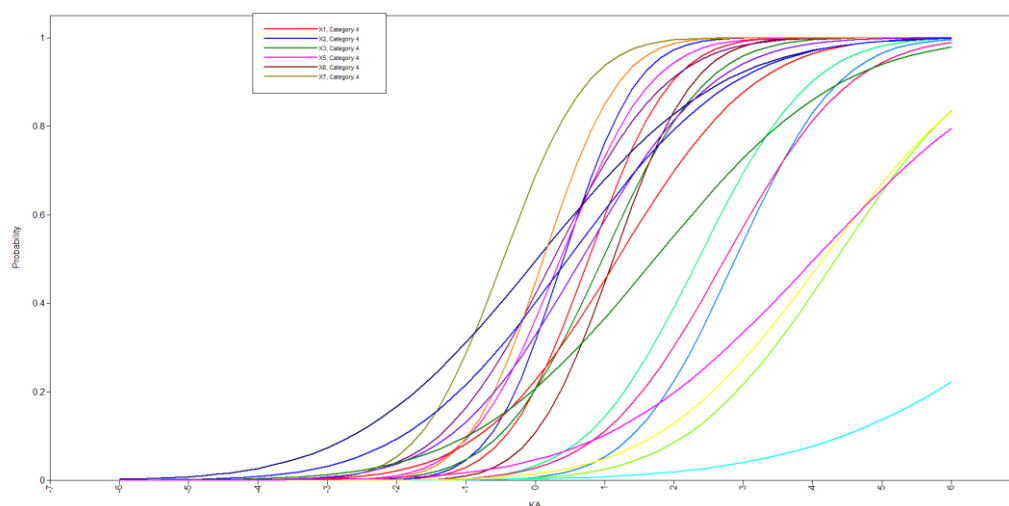


Tabel 5. Koefisien Muatan Faktor *Standardized* 20 Item Ketidakhujuran Akademik

Nomor Item	Koefisien Muatan	S.E.	T-value	P-value	Koefisien Bias
1	0.730	0.032	22.788	0.000*	0.000
2	0.769	0.030	25.770	0.000*	0.000
3	0.650	0.037	17.336	0.000*	0.000
4	0.559	0.067	8.392	0.000*	0.512
5	0.698	0.036	19.236	0.000*	0.000
6	0.742	0.036	20.348	0.000*	0.000
7	0.724	0.039	18.583	0.000*	0.000
8	0.477	0.061	7.763	0.000*	0.557
9	0.440	0.053	8.339	0.000*	0.575
10	0.398	0.071	5.605	0.000*	0.618
11	0.461	0.049	9.316	0.000*	0.205
12	0.375	0.055	6.796	0.000*	0.481
13	0.376	0.079	4.788	0.000*	0.219
14	0.432	0.052	8.317	0.000*	0.000
15	0.613	0.041	14.895	0.000*	0.000
16	0.229	0.087	2.636	0.008*	0.678
17	0.415	0.070	5.975	0.000*	0.448
18	0.762	0.033	22.769	0.000*	0.000
19	0.618	0.052	11.950	0.000*	0.000
20	0.557	0.046	12.200	0.000*	0.000

## Keterangan:

<i>S.E</i>	:	<i>Standard error factor loading</i>	<i>P-Value</i>	<i>Nilai probability</i>
<i>T-Value</i>	:	Nilai t-test	*)	Signifikan pada level 0.05

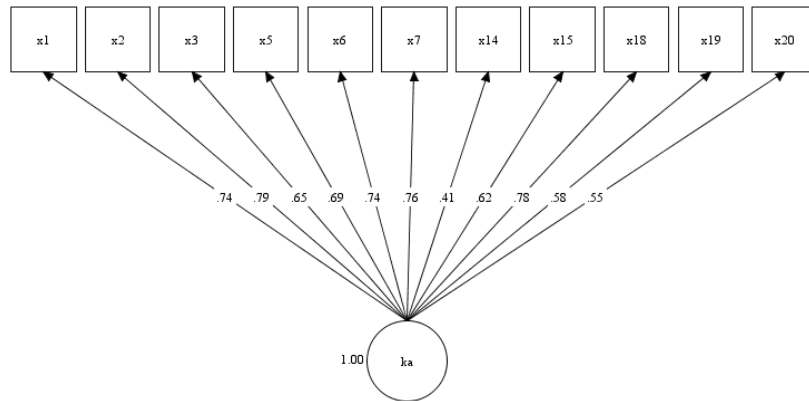


Gambar 3. Kurva Karakteristik dari Pengukura Ketidajujuran Akademik 20 item

Pada kolom paling kanan tabel 5 disampaikan pula koefisien muatan faktor yang spesifik untuk menunjukkan bias dimana selain mengukur ketidakjujuran akademik, butir tersebut mengukur konstruk lain yang lima diantaranya mengukur “*need for recognition*” dan empat diantaranya mengukur “kemampuan memanipulasi”. Jika bias ini tidak diperhitungkan di dalam skoring, maka akan terjadi *over* estimasi terhadap skor ketidakjujuran akademik yang dihasilkan. Misalnya, skor menjadi lebih tinggi hanya karena subjek memiliki *need for recognition* yang tinggi. Begitu pula halnya dengan faktor kemampuan memanipulasi. Sebagai ringkasan dari model *bi-factor* ini, dapat disimpulkan adanya temuan sebagai berikut:

1. Tidak ada item yang perlu di-*drop* karena semuanya signifikan mengukur ketidakjujuran akademik, asalakan cara skoringnya menggunakan perhitungan *true score* dengan model *bi-factor* seperti pada gambar 4.5 di atas.
2. Ditemukan dua jenis bias pada instrumen ini. Bias yang pertama adalah adanya lima butir soal yang selain mengukur ketidakjujuran akademik ternyata juga mengukur *need for recognition*, yaitu item nomor 4, 8, 10, 16, dan 17. Selanjutna, terdapat empat butir soal (item nomor 9, 11, 12, dan 3) mengukur kemampuan seseorang dalam melakukan manipulasi pada umumnya (tidak terbatas pada konteks ketidakjujuran akademik).

Meskipun tidak ada item yang dieliminasi, namun untuk mendapatkan *true score* pada model *bi-factor* seperti di atas, hanya dapat dilakukan dengan menggunakan perangkat lunak yang canggih seperti misalnya Mplus dan tidak dapat dilakukan dengan *software* di bidang *Item Response Theory* (IRT) dimana asumsi unidimensionalitas tanpa bias harus terpenuhi. Jika hanya *software* setingkat IRT (Bilog, Winstep, dsb) yang tersedia, maka sembilan item yang mengalami bias harus di-*drop* terlebih dahulu, barulah dapat diestimasi *true score* untuk setiap orang. Adapun model unidimensional tanpa bias yang dalam hal ini terdiri dari 11 item menghasilkan model fit ( $\chi^2 = 95.010$ ,  $df = 44$ ,  $p < 0.000$ ,  $RMSEA_{estimate} = 0.057$ ,  $90\% C.I = 0.041$  sampai  $0.073$ , *probability*  $RMSEA < 0.005 = 0.215$ ,  $CFI = 0.979$ , dan  $TLI = 0.973$ ). Model unidimensional yang lebih sempurna ini dapat disajikan dengan diagram pada gambar 4 sedangkan koefisien muatan faktor (daya pembeda) dari setiap item dapat disajikan pada tabel 6, sedangkan kurva karakteristik soal (ICC) untuk setiap butir disajikan pada gambar 5.



Gambar 4. Diagram Skala Pengukuran Ketidakjujuran Akademik II Item Unidimensional Menggunakan Pendekatan IRT

Tabel 6. Koefisien Muatan Faktor *Standardized* II Item Ketidakjujuran Akademik

Nomor Item	Koefisien	S.E.	T-value	P-value
1	0.743	0.032	22.997	0.000*
2	0.785	0.030	25.917	0.000*
3	0.649	0.039	16.776	0.000*
5	0.689	0.038	17.968	0.000*
6	0.741	0.036	20.359	0.000*
7	0.756	0.040	18.792	0.000*
14	0.405	0.055	7.343	0.000*
15	0.623	0.042	14.843	0.000*
18	0.781	0.033	23.395	0.000*
19	0.581	0.053	11.067	0.000*
20	0.553	0.047	11.752	0.000*

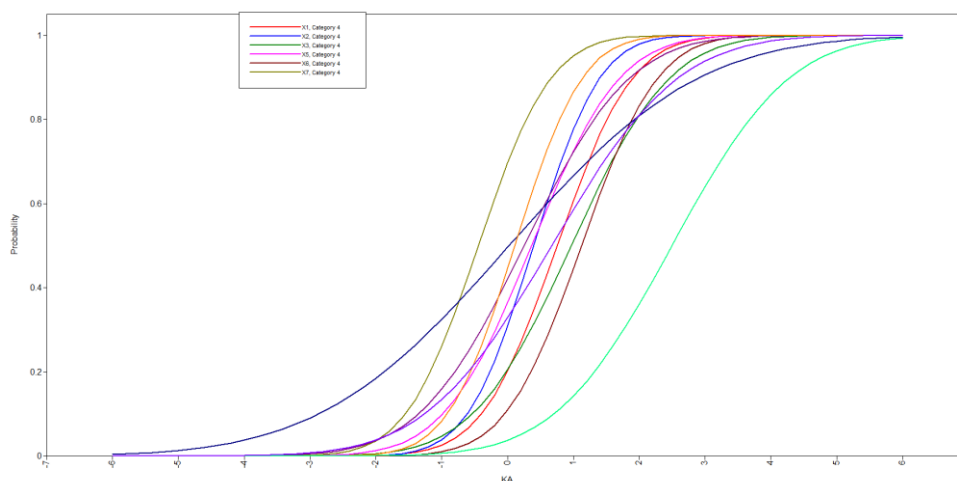
**Keterangan:**

S.E : Standard error dari factor loading

T- Value : Nilai t-test

P- Value : Nilai probability

\*) : Signifikan pada level 0.05



Gambar 5. Kurva Karakteristik dari Pengukuran Ketidakjujuran Akademik II item

Seperti terlihat pada gambar 4 tidak terdapat lagi korelasi antar kesalahan pengukuran, sehingga asumsi unidimensionalitas dan asumsi independensi lokal yang disyaratkan oleh IRT telah terpenuhi dengan baik. Selain itu, pada tabel 6 juga terlihat bahwa seluruh koefisien muatan faktor signifikan mengukur ketidakjujuran akademik dan tidak ada koefisien yang nilainya di bawah 0.4 (semuanya memiliki daya pembeda yang cukup tinggi). Dengan demikian, *true score* untuk instrumen yang terdiri dari II item ini dapat dengan mudah diperoleh jika menggunakan perangkat lunak IRT. Namun demikian, karena daya pembeda dari II item tersebut masih bervariasi, maka penggunaan *software* IRT yang berbasis model satu parameter (*rasch model*) seperti misalnya Winstep tidak dianjurkan untuk digunakan dalam skoring. Apalagi jika hanya menggunakan teori tes klasik, dimana skoring dilakukan hanya dengan menjumlahkan skor setiap item. Adapun kesebelas butir item tersebut sebagai berikut:

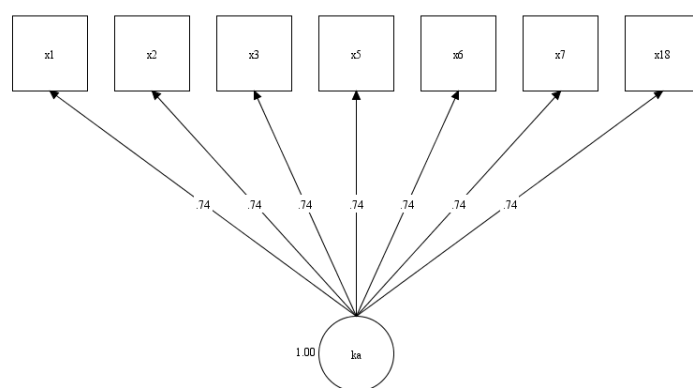
Tabel 7. Instrumen Ketidakjujuran Akademik II-item Unidimensional

Nomor Item	Pernyataan
1	Ketika ujian saya menggunakan contekan berupa catatan kecil
2	Saat ujian saya mencoba untuk menyontek pekerjaan teman yang duduknya berdekatan
3	Mencoba mencari bocoran soal sebelum ujian dilaksanakan
5	Mencoba menyalin pekerjaan teman dengan seizinnya
6	Mencari contekan jawaban apapun caranya
7	Membantu teman dalam menjawab ujian
14	Meminta bantuan orang lain yang lebih pandai untuk memeriksa tugas/tulisan saya
15	Mencari informasi tentang soal ujian kepada mahasiswa lain yang sudah menempuhnya
18	Mencari jawaban ujian melalui bantuan <i>smartphone</i>
19	Mengambil gambar (memotret) soal ujian secara sembunyi-sembunyi
20	Menitipkan absen pada teman yang masuk

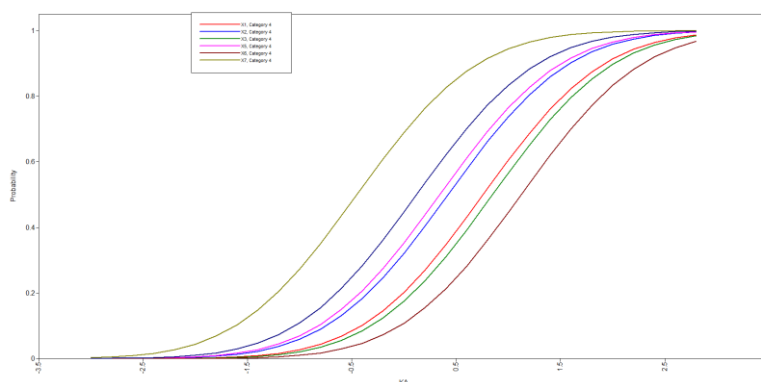
Jika pengguna instrumen ketidakjujuran akademik ini tidak memiliki akses terhadap perangkat lunak Mplus ataupun IRT. Sehingga tidak dapat diperoleh *true score* untuk setiap subjek yang diukur, maka diperlukan

model yang bersifat lebih restriktif. Yaitu, model dimana selain asumsi unidimensionalitas dan asumsi independensi lokal telah terpenuhi, juga harus ditambah dengan asumsi bahwa seluruh butir soalnya paralel. Adapun yang dimaksud dengan paralel adalah jika semua soal memiliki daya pembeda (koefisien muatan faktor) yang sama dan tingkat *threshold* (tingkat kesukaran) yang sama, serta varian kesalahan pengukuran yang sama pula. Jika seluruh asumsi ini terpenuhi, maka berlakulah teori tes klasik, dimana dapat diperoleh skor komposit (*raw score*) dengan cara menjumlahkan skor pada setiap butir soal.

Namun demikian, sebuah alat ukur dapat dianggap memiliki butir yang paralel (disebut *Tau equivalence*) dan pada instrumen seperti ini boleh dilakukan penggunaan *raw score* sebagai estimasi dari *true score* dengan catatan harus melampirkan/melaporkan tingkat reliabilitas dari *raw score* yang dihasilkan (misalnya koefisien alpha dari Cronbach). Jika situasi ini terpenuhi, maka skoring dapat dilakukan oleh siapa saja tanpa bantuan perangkat lunak apapun karena cukup dengan cara menjumlahkan skor pada butir-butir soalnya saja. Setelah peneliti melakukan uji paralelitas yang dalam hal ini adalah kondisi "*Tau equivalence*" (dimana seluruh item harus sama koefisien muatan faktornya dan paralel kurva karakteristik soalnya). Penelitian ini menghasilkan model fit ( $\chi^2 = 51.148$ ,  $df = 20$ ,  $p < 0.000$ ,  $RMSEA\ estimate = 0.066$ ,  $90\% C.I = 0.044$  sampai  $0.089$ ,  $probability\ RMSEA < 0.005 = 0.108$ ,  $CFI = 0.982$ , dan  $TLI = 0.981$ ). Artinya, model yang paralel dengan item terbukti fit dengan data dan boleh dipakai. Adapun diagram hasil CFA pada 7 butir soal tersebut disajikan pada gambar 6 dan kurva karakteristik soal yang paralel dapat dilihat pada gambar 7 sebagai berikut:



Gambar 6. Diagram Skala Pengukuran Ketidajujuran Akademik 7 item paralel menggunakan pendekatan Maximum Likelihood



Gambar 7. Kurva Karakteristik dari Pengukuran Ketidajujuran Akademik 7 item paralel

Sedangkan koefisien untuk 7 item adalah 0.740 yang semuanya signifikan. Alat ukur ketidajujuran akademik dengan 7 item yang paralel ini, dapat dikatakan sebagai bentuk *short-form* dari alat ukur yang semula terdiri 20 butir soal, dimana seluruh asumsi teori tes klasik telah terbukti terpenuhi. Oleh sebab itu,

penggunanya dapat melakukan skoring cukup dengan menjumlahkan skor pada setiap butir soalnya. Instrumen dalam bentuk *short-form* ini terdiri dari 7 butir yang pernyataannya disajikan pada tabel 7.

Tabel 7. Instrumen Ketidakhujuran Akademik 7 Item Paralel

Nomor Item	Pernyataan
1	Ketika ujian saya menggunakan contekan berupa catatan kecil
2	Saat ujian saya mencoba untuk menyontek pekerjaan teman yang duduknya berdekatan
3	Mencoba mencari bocoran soal sebelum ujian dilaksanakan
5	Mencoba menyalin pekerjaan teman dengan seizinnya
6	Mencari contekan jawaban apapun caranya
7	Membantu teman dalam menjawab ujian
18	Mencari jawaban ujian melalui bantuan <i>smartphone</i>

Untuk alat ukur yang terdiri dari 20 item (*bi-factor*) peneliti menyebutnya sebagai bentuk tingkat rendah dari instrumen pengukuran KA, karena memerlukan *software* yang canggih untuk skoringnya. Sedangkan yang terdiri dari 11 item, peneliti menyebutnya sebagai bentuk yang ideal namun tetap memerlukan *software* IRT untuk skoringnya. Adapun yang terdiri dari 7 butir dengan cara skoring yang sangat sederhana dan dapat digunakan tanpa *software* peneliti sebut sebagai *short-form*. Karena penggunaan alat ukur *short-form* menggunakan *raw score* dan mensyaratkan indeks reliabilitas, maka perlu dilaporkan disini bahwa perhitungan indeks reliabilitas terhadap *raw score* tersebut dengan cara yang lazim adalah dengan menghitung koefisien alpha. Adapun reliabilitas dari *raw score* pada versi *short-form* ini dapat dihitung dengan cara yang lebih akurat daripada misalnya Cronbach Alpha, seperti diketahui koefisien alpha dari Cronbach hanyalah merupakan *lower bound* (batas bawah) dari koefisien reliabilitas.

Untuk mendapatkan koefisien yang lebih akurat, dapat digunakan rumus berikut (Raykov, 1997; Umar, 2012) dengan syarat tersedia informasi mengenai koefisien muatan faktor dan varians residual dari setiap item :

$$R_{KA} = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \theta_{ii}}$$

**Keterangan:**

$R_{KA}$  = Reliabilitas skor ketidakhujuran akademik

$\lambda$  = Koefisien muatan faktor item "i" yang *standardized*

$\theta_{ii}$  = Varians residual untuk item i yang diperoleh dari rumus  $\theta_{ii} = (1 - \lambda_i^2)$

Berikut di bawah ini adalah perhitungan reliabilitas dari alat *short-form* alat ukur ketidakhujuran akademik yang terdiri dari 7 item paralel seperti yang disebutkan sebelumnya pada tabel 5.

$$R_{KA} = \frac{(7 \times 0.740)^2}{(7 \times 0.740)^2 + (1 - 0.740^2)}$$

$$R_{KA} = \frac{(7 \times 0.740)^2}{(7 \times 0.740)^2 + 7(1 - 0.740^2)}$$

$$R_{KA} = \frac{26.8324}{26.8324 + 7(0.4524)}$$

$$R_{KA} = \frac{26.8324}{26.8324 + 3.1668}$$

$$R_{KA} = \frac{26.8324}{29.9992}$$

$$R_{KA} = 0.894$$

$$R_{KA} \approx 0.89$$

Jadi, bagi pengguna awam yang tidak memiliki akses untuk menghitung *true score* maka dapat menggunakan *raw score* yang memiliki tingkat reliabilitas = 0.89. Alat ukur ini disebut juga sebagai alat ukur pada tingkat yang lebih praktis, karena skoring yang dilakukan hanya dengan menjumlah setiap skor item.

Selanjutnya, setelah mendapatkan reliabilitas maka *standard error of measurement* (SEM) dapat dihitung dengan rumus:

$$\mathbf{SEM = SD\sqrt{1 - \Omega}}$$

**Keterangan:**

SD = Standar deviasi

$\Omega$  = Reliabilitas

Berikut di bawah ini merupakan perhitungan SEM alat ukur ketidakjujuran akademik bentuk singkat (*short-form*):

$$\mathbf{SEM = 6.044\sqrt{1 - 0.89}}$$

$$\mathbf{SEM = 6.044\sqrt{0.11}}$$

$$\mathbf{SEM = 6.044 (0.331)}$$

$$\mathbf{SEM = 2.00}$$

Jika sudah didapat SEM, maka bisa didapatkan tingkat kepercayaan (*confident interval*) sebesar 95% pada setiap skor menggunakan rumus:

$$\mathbf{C. I. True Score = X \pm 1.96 \times SEM}$$

**Keterangan:**

C.I = Tingkat kepercayaan

X = Skor

SEM = *Standard error of measurement*

Pada alat ukur bentuk singkat ini, skor terendah adalah 4 dan skor tertinggi adalah 28. Dengan tingkat kepercayaan 95% pada individu yang memiliki skor 17, maka individu tersebut akan memiliki *true score* yang berkisar dari 13.08 sampai dengan 20.92. Adapun rentang skor tersebut didapat dari perhitungan berikut ini:

$$\mathbf{C. I. True Score = 17 \pm 1.96 \times SEM}$$

$$\mathbf{C. I. True Score = 17 - (1.96 \times 2.00) = 13.08}$$

$$\mathbf{C. I. True Score = 17 + (1.96 \times 2.00) = 20.92}$$

Sebagai pembandingan, peneliti melakukan uji validitas konstruk untuk mendapatkan alat ukur dengan himpunan item yang bersifat paralel menggunakan estimator MLR. Setelah melakukan uji model fit, diperoleh  $p\text{-value} = 1.000$ , nilai tersebut menunjukkan bahwa model 7 item paralel fit dengan data. Adapun 7 item ini memiliki koefisien muatan faktor ( $\lambda$ ) sebesar 0.662 dan signifikan secara statistik. Sehingga dapat menjadi instrumen alternatif jika dirasa pengukuran ketidakjujuran akademik *short-form* memiliki butir pernyataan yang lebih sedikit. Dalam hal skoring, alat ukur versi 7 item ini memiliki cara yang sama dengan *short-form*, yaitu dengan cara menjumlahkan masing-masing *raw score* dari ketujuh item. Reliabilitas alat ukur versi 7 item sebesar 0.88 yang diperoleh dari perhitungan sebagai berikut:

$$R_{KA} = \frac{(10 \times 0.662)^2}{(10 \times 0.662)^2 + 10(1 - 0.662^2)}$$

$$R_{KA} = \frac{(6.62)^2}{(6.62)^2 + 10(1 - 0.662^2)}$$

$$R_{KA} = \frac{(6.62)^2}{(6.62)^2 + 10(0.561)}$$

$$R_{KA} = \frac{43.824}{43.824 + 5.61}$$

$$R_{KA} = \frac{43.824}{49.441}$$

$$R_{KA} = 0.886$$

$$R_{KA} = \sim 0.88$$

## Penutup

Berdasarkan hasil uji validitas skala ketidakjujuran akademik yang telah peneliti jabarkan di atas. Hipotesis mayor penelitian pengukuran ketidakjujuran akademik yang berbunyi "*seluruh butir soal (item) yang valid dalam mengukur konstruk ketidakjujuran akademik*" dapat diterima, artinya 20 item yang diteorikan mengukur konstruk ketidakjujuran akademik. Meskipun ditemukan terdapat 9 item yang mengukur faktor lainnya (bias), yaitu item 4, 8, 9, 10, 11, 12, 13, 16, dan 17. Namun demikian, alat ukur model 20 item (*long-form*) ini tetap dapat digunakan meskipun kurang ideal. Selanjutnya, hipotesis minor penelitian pengukuran ketidakjujuran akademik yang berbunyi "*setiap butir soal (item) signifikan dalam mengukur ketidakjujuran akademik*" dapat diterima, hal ini dikarenakan tidak ada item yang memiliki koefisien yang negatif sehingga tidak ada item yang dieliminasi.



Dalam penelitian ini, selain jenis *long-form* peneliti juga menemukan dua jenis alat ukur lainnya berdasarkan jumlah item dan skoring yang dinamakan *short-form* dan *ideal-form*. Untuk jenis *short-form* yang terdiri dari 7 butir soal mengenai ketidakjujuran akademik, dapat digunakan dengan skoring yang lebih mudah yaitu menjumlahkan *raw score* ketujuh item yang sudah paralel tersebut. Sedangkan untuk versi *ideal-form* (11 item) dan *long-form* (20 item) diperlukan *software* untuk membantu dalam hal skoring. Adapun ketiga jenis alat ukur yang sudah dibakukan tersebut dapat digunakan sesuai dengan kemampuan dan kebutuhan peneliti di kemudian hari.

## Daftar Pustaka

- Gitanjali, B. (2004). Academic dishonesty in Indian medical colleges. *Journal of Postgraduate Medicine*, 50(4), 281–4.
- Iyer, R., & Eastman, J. K. (2008). The Impact of Unethical Reasoning on Academic Dishonesty: Exploring The Moderating Effect of Social Desirability on Academic Dishonesty. *Marketing Education Review*, 18(2), 21–33.
- Lord, F.M. (1980). *Applications of item response theory to practical testing problems*. Hillsdale, NJ: Lawrence Erlbaum Associates.
- McCabe, D. L., & Trevino, L. K. (1993). Academic Dishonesty: Honor Codes and Other Contextual Influences. *The Journal of Higher Education*, 64(5), 522–538.
- Raykov, T. (1997). Estimation of composite reliability for congeneric measures. *Applied Psychological Measurement*, 21(2), 173–184.
- Umar, J. (2012). Mengenal Lebih Dekat Konsep Reliabilitas Skor Tes. *Jurnal Pengukuran Psikologi Dan Pendidikan Indonesia*, 2, 126–140.
- Umar, J. (2015). *Bahan Ajar Statistika 3*. Fakultas Psikologi Universitas Islam Negeri Syarif Hidayatullah : Jakarta.