

PENGARUH KUALITAS ITEM TERHADAP RELIABILITAS ALPHA CRONBACH PADA TES BAKAT NUMERIKAL DAN TES POTENSI AKADEMIK

Jumailiyah

Dosen Jurusan Psikologi Pendidikan dan Bimbingan /BK FIP IKIP Mataram

E-mail: jumailiyah@gmail.com

ABSTRACT: The quality of the items in this study is limited to the level of difficulty and item discrimination in classical model and reliability test is Alpha Cronbach. Object of research is numerical standardized test consists of 18 groups and 12 groups of tests that have not been standardized each of 20 items,. Each group calculated Alpha Cronbach coefficient, the mean level of difficulty of the questions and the mean discrimination items. Testing the hypothesis is calculated by multiple regression analysis with SPSS ver PSAW. 18. The results obtained indicate, simultaneously the value of $F = 67.625$ ($\text{sig} = 0.000$), the effect of item difficulty level (Mean P) to the reliability of the value $t = 2,949$ ($\text{sig} = 0.007$) and the average impact biserial item (Mean bis) obtained $t = 3,653$ ($\text{sig} = 0.001$). In conclusion there is the influence of the quality of the items on Cronbach's alpha coefficient.

Keyword: Reliabilitas Alpha Cronbach, Mean P, Mean bis.

PENDAHULUAN

Proses penentuan tes yang memenuhi standar kualitas membutuhkan usaha – usaha profesional di bidang pengukuran psikologi dan pendidikan. Cohen berpendapat pengembangan tes melalui tahapan ”merencanakan tes, menyusun tes, ujicoba tes, analisis butir, dan revisi tes” (Cohen, 2005 : 190–240). Robert Gregory secara lebih rinci pengembangan tes ”menentukan batasan suatu tes, memilih metode pengukuran, menulis butir tes, mengujikan butir tes, merevisi tes dan menerbitkan tes” (Gregory, 2000 : 117). Thorndike mengemukakan 10 langkah dalam pengembangan tes yaitu: menentukan definisi dan sasaran yang hendak diukur tes, menentukan tujuan tes, keterbatasan penggunaan tes, menyusun kisi-kisi tes, spesifikasi format butir tes termasuk bentuk tes dan cara menjawab tes, ujicoba tes dan analisis ujicoba butir tes, perhitungan statistik karakteristik butir, formulasi metode standarisasi dan menyusun norma, pengumpulan data untuk mengevaluasi validitas atau bias, mengembangkan manual tes (Thorndike, 1982: 10 -11)., Langkah pengembangan tes bakat (potensi intelektual) yaitu: (1) pengembangan spesifikasi tes, (2) penulisan soal, (3) penelaahan soal, (4) perakitan soal untuk keperluan uji-coba, (5) uji-coba tes, (6) analisis butir soal, (7) seleksi dan perakitan soal, (8) pencetakan tes, (9) administrasi tes bentuk akhir, (10) penyusunan skala dan norma (Suryabarata, 2005 : 169).

Pengembangan atau standarisasi tes pada tes seleksi seperti tes yang terkenal SAT

mengalami perkembangan dan penyempurnaan, sejak tahun 1926, dilakukan revisi setiap tahun dengan sampel yang terus meningkat sampai 10 000 siswa. (Hanford, 1996, Kaplan, 2012: 315). Dari prosedur pengembangan tes akan tampak tahapan yang menunjukkan bahwa tes yang terstandar diharapkan menunjukkan item-item tes yang berkualitas, tes yang teridentifikasi kualitas item dengan standar-standar tertentu. Tahap uji-coba bertujuan utama sebagai usaha meningkatkan tes dengan memperbaiki atau menghapus item infektif. Fungsi penting lain dari analisis item, terutama analisis item pada tes pencapaian, adalah memberikan informasi diagnostic mengenai apa yang diketahui dan tidak diketahui oleh peserta tes (Aiken, 2006: 80).

Reliabilitas dapat diartikan ”keterpercayaan, keterandalan, kejelasan, kestabilan, konsistensi, dan sebagainya, namun ide pokok yang terkandung dalam konsep reliabilitas adalah sejauhmana hasil suatu pengukuran dapat dipercaya” (Azwar, 2003 : 4). Reliabilitas alat ukur erat berkaitan erat dengan masalah error pengukuran (error of measurement). Error pengukuran sendiri dapat menunjukkan sejauhmana inkonsistensi hasil pengukuran terjadi apabila pengukuran diulang pada kelompok subjek yang sama.

Dua teori pengukuran psikologi dan pendidikan yaitu teori pengukuran klasik dan pengukuran modern atau teori respons butir. Teori respons butir dikembangkan untuk mengatasi keterbatasan pada teori pengukuran klasik, terutama kelemahan yang bersifat

dependen antara kelompok peserta tes dan butir. Bersifat dependen diartikan bahwa hasil pengukuran tergantung pada kelompok peserta yang mengerjakan tes. Jika tes diujikan pada kelompok peserta dengan berkemampuan tinggi maka tingkat kesukaran butir soal akan rendah. Sebaliknya jika tes diujikan kepada kelompok peserta berkemampuan rendah maka tingkat kesulitan butir tersebut menjadi tinggi (Hambleton, H. 1991: 3). Ciri item lainnya yaitu kemampuan item untuk membedakan peserta tes berdasarkan kemampuannya disebut dayabeda item. Dua ciri item ini menjadi perhatian kedua model pengukuran tersebut dan mempengaruhi konsistensi pengukuran.

Pada model logistic dayabeda butir mempengaruhi keakuratan hasil pengukuran, sedangkan tingkat kesukaran item kurang mempengaruhinya (Embretson, 2000: 169), hal ini menimbulkan pertanyaan yang membutuhkan pengkajian lebih mendalam, apakah gejala pada model logistic akan sama dengan model klasik? Dengan demikian rumusan masalah dalam kajian ini apakah ada pengaruh tingkat kesukaran dan dayabeda item terhadap reliabilitas Alpha Cronbach pada tes kemampuan. .

METODE PENELITIAN

Sumber Data

Tes kemampuan penelitian tes kemampuan berpikir numerikal dari tiga tes tes masing-masing 40 item. Tes tersebut dikelompokkan item tes ganjil, item tes genap dan dikelompokkan lagi berdasarkan banyak responden 250, 500, dan 1000 sehingga diperoleh 18 kelompok data tes kemampuan berpikir numerikal. Selain pengelompokan data diperoleh dari tes Tes Potensi Akademik (TPA) dari tes masuk IKIP Mataram tahun 2014 dan 2015 dikelompokkan ke dalam subtes verbal, subtes matematika dan subtes spatial, selanjutnya pengelompokan berdasarkan jumlah responden 250, 500, 1000 dan 2000 sehingga diperoleh 12 kelompok dari TPA dan total menjadi 30 kelompok.

Data tes kemampuan berpikir numerikal dikumpulkan dari 28 Sekolah Menengah Atas di Pulau Lombok, dan telah dilakukan proses standarisasi tes (Jumailiyah, 2015, Prosiding Seminar Internasional Malindo 4 di Denpasar Bali). Sedangkan TPA belum distandarisasi sebagaimana pengembangan tes potensi yang biasanya digunakan dalam praktek Bimbingan Konseling sekolah. Pada kedua macam tes tersebut dikelompokkan berdasarkan randomisasi responden, sedangkan item yang

diteliti adalah semua item yang ada pada setiap subtes yaitu masing-masing 20 item.

Instrumen Penelitian

Pembahasan di atas mengemukakan bahwa tes terdiri dari dua macam yang merupakan tes pilihan ganda dengan lima option. Instrumen tes merupakan tes pilihan ganda dengan lima option Tes kemampuan berpikir numerikal merupakan tes yang bertujuan untuk mengetahui kemampuan atau potensi siswa dalam menentukan penjurusan di SMA, suatu instrumen tes berisi Definisi operasional berpikir numerikal yaitu kemampuan dalam melakukan operasi dan penalaran aritmatika, dan dasar-dasar aljabar dan penerapan dalam masalah kehidupan. Operasi aritmetika: penjumlahan, pengurangan, perkalian dan pembagian pada bilangan bulat, pecahan biasa, pecahan desimal, bilangan berpangkat dan bilangan akar. Penalaran aritmatika: penerapan perkalian, pembagian, prosentase, dan nilai rata-rata dalam masalah kehidupan. Dasar-dasar matematika, aljabar dan penerapan dalam masalah kehidupan: melakukan operasi bilangan bulat dan pecahan dengan tanda plus dan minus, persamaan model matematika pada satu dan dua variabel, pertidaksamaan, penerapan diagram venn dalam masalah kehidupan. TPA sebagai tes seleksi masuk IKIP Mataram tahun 2014 dan 2015 mengungkapkan kemampuan Verbal terdiri dari pertanyaan-pertanyaan: persamaan kata, lawan kata dan padanan kata, dan subtes spatial berisi soal-soal yang mengungkapkan analogi masing-masing subtes terdiri dari 20 item tes, Subtes spatial merupakan tes yang mengungkapkan kemampuan menganalisis suatu persoalan yang dipresentasikan dalam bentuk gambar-gambar dua dimensi dan tiga dimensi, dan kemudian menentukan secara tepat pola perubahan sederet gambar, bagian-bagian gambar, kekhasan pola gambar, dan hubungan pola antar gambar, yang disajikan dalam suatu wacana soal, dan soal berupa gambar-gambar dua dimensi dan tiga dimensi. Kemudian peserta tes memilih satu di antara pilihan yang merupakan kesimpulan paling tepat sebagai pola perubahan sederet gambar, bagian-bagian gambar, kekhasan pola gambar, dan hubungan pola antar gambar (Atmoko, 2016: 7), subtes matematika berisi soal matematika dasar.

Prosedur dan Analisis Data Penelitian:

Analisis data awal mengacu pada penelitian deskriptif, suatu metode yang memaparkan atau mendeskripsikan data yang diperoleh dari nilai ciri item. Penskoran

dilakukan dengan penskoran dikotomi yaitu skor dengan pemberian angka satu (1) untuk jawaban yang benar dan skor nol (0) untuk jawaban yang salah pada setiap item, sehingga masing-masing responden akan mendapat skor berdasarkan jumlah item yang benar.

Reliabilitas Alpha Cronbach dihasilkan pada setiap kelompok tes berdasarkan output program ITEMAN versi 3.5, MICROCAT (tm) System menampilkan: (1) Rata-rata tingkat kesukaran item (2) Rata-rata dayabeda item, dan (3) Reliabilitas Alpha Cronbach, dengan rumus:

$$r = \alpha = \left(\frac{N}{N-1} \right) \left(\frac{S^2 - \sum S_i^2}{S^2} \right) \text{ (Kaplan, 2012: 115)}$$

Tahap analisis untuk menguji signifikansi pengaruh dua variabel bebas terhadap variabel tergantung. Variabel bebas penelitian yaitu tingkat kesukaran dan dayabeda item sedangkan variabel tergantung reliabilitas Alpha Cronbach dianalisis dengan teknik analisis regresi ganda. Regresi ganda merupakan teknik analisis untuk mengetahui pengaruh dua atau lebih variabel bebas terhadap variabel tergantung. Dua variabel bebas tingkat kesukaran item (X₁) dan dayabeda item (X₂), variabel tergantung reliabilitas Alpha Cronbach dianalisis dengan program SPSS PSW 18.0.

data, dari tes kemampuan berpikir numerikal merupakan tes yang telah melalui proses ujicoba dan telah dilakukan revisi pembuangan item yang tidak memenuhi standar dan ditambah dengan item tes yang baru. Pada analisis ITEMAN dilakukan dengan beberapa putaran sehingga item yang terdapat pada setiap tes mencerminkan item tes yang baik. Berbeda dengan data tes seleksi masuk IKIP Mataram tes tahun 2014 dan 2015, tes yang belum dilakukan pembakuan dan uji coba, tidak melalui proses menyempurnaan kualitas tes. Pengelompokan berdasarkan banyak responden walaupun bukan bermaksud untuk mengetahui perbedaan reliabilitas darisegi jumlah responden akan tetapi analisis dengan N pasangan angka yang lebih besar menunjukkan hasil menyakinkan.

HASIL DAN PEMBAHASAN

Analisis Deskriptif

Pada bagian sumber data dikemukakan dua tes yang digunakan sebagai pengumpul

Tabel 1. Pengelompokan Tes Alpha Cronbach, Kesulitan dan Dayabeda

SET TES	RESPONDEN	N	ALPHA	KESULITAN	DAYABEDA
		ITEM	CRONBACH		
A GANJIL	250	20	0.772	0.521	0.574
A GANJIL	500	20	0.776	0.525	0.578
A GANJIL	1000	20	0.775	0.527	0.579
A GENAP	250	20	0.812	0.558	0.619
A GENAP	500	20	0.812	0.547	0.616
A GENAP	1000	20	0.812	0.549	0.618
B GANJIL	250	20	0.805	0.525	0.601
B GANJIL	500	20	0.811	0.515	0.603
B GANJIL	1000	20	0.808	0.524	0.603
B GENAP	250	20	0.744	0.573	0.542
B GENAP	500	20	0.696	0.549	0.505
B GENAP	1000	20	0.812	0.549	0.618
C GANJIL	250	20	0.783	0.454	0.579
C GANJIL	500	20	0.807	0.469	0.458
C GANJIL	1000	20	0.799	0.465	0.595
C GENAP	250	20	0.794	0.481	0.447
C GENAP	500	20	0.811	0.49	0.46
C GENAP	1000	20	0.802	0.489	0.594
TPA MAT	250	20	0.55	0.263	0.434
TPA MAT	500	20	0.557	0.273	0.43
TPA MAT	1000	20	0.554	0.269	0.434
TPA MAT	2000	20	0.563	0.272	0.437
TPA					
VERBAL	250	20	0.37	0.29	0.37
TPA					
VERBAL	500	20	0.371	0.284	0.37

TPA VERBAL	1000	20	0.371	0.284	0.378
TPA VERBAL	2000	20	0.345	0.284	0.369
TPA SPATIAL	250	20	0.535	0.377	0.405
TPA SPATIAL	500	20	0.533	0.361	0.394
TPA SPATIAL	1000	20	0.463	0.364	0.375
TPA SPATIAL	2000	20	0.394	0.375	0.381

Dari tabel 1 di atas dikemukakan pada kolom pertama Set A, B, dan set C menunjukkan tes kemampuan berpikir numerikal sedangkan tes TPA sebagai tes seleksi masuk IKIP Mataram. Banyak item tes setiap kelompok sama yaitu sebanyak 20 item. Perhatikan kolom 4 menunjukkan koefisien Alpha Cronbach, koefisien disekitar 0,8 diperoleh dari subtes kemampuan berpikir numerikal yang telah diterapkan melalui standarisasi sedangkan TPA menunjukkan koefisien Alpha Cronbach yang rata-rata lebih rendah atau lebih kecil dari koefisien di bawah 0,6

Hasil analisis data deskriptif ditunjukkan rata-rata dan standar deviasi untuk tingkat kesukaran item, dengan kriteria penarikan kesimpulan tingkat kesukaran soal. Item baik adalah item yang tidak terlalu sukar dan tidak juga terlalu mudah, yaitu item dengan taraf kesukaran 0.5 atau disekitar 0.5 atau $0.33 \leq P \leq 0.67$ (Ebel, 1991: 232), Nitko berpendapat kriteria tingkat kesukaran butir yaitu, item sangat sukar bernilai $p > 0.25$ dan

item tingkat kesukaran mudah dengan nilai $p > 0.80$ (Nitko, 2003:323). Kualitas dayabeda item ditentukan berdasarkan pendapat $> 0.40 =$ Sangat bagus, $0.30 - 0.39 =$ Cukup bagus tidak perlu perbaikan, $0.20 - 0.29 =$ Kurang bagus dan perlu perbaikan, dan < 0.19 berarti jelek (Ebel, 1991: 232). Seberapa tinggi suatu koefisien reliabilitas harus dicapai hingga dikatakan “cukup tinggi” jawabannya akan tergantung pada penggunaan tes. Perkiraan reliabilitas dalam cakupan nilai antara 0,70 dan 0,80 telah dianggap cukup baik untuk sebagian besar tujuan penelitian dasar. (Kaplan, 2012 :123).

Dari table 2 dan kriteria kualitas item di atas maka tingkat kesukaran dan dayabeda item menunjukkan item yang baik sebagaimana ditunjukkan berdasarkan nilai rata-rata atau mean, demikian pula rata-rata reliabilitas Alpha Cronbach sebesar 0,661 suatu koefisien sedikit di bawah kriteria reliabilitas “cukup tinggi” sehingga dapat diartikan bahwa rata-rata reliabilitas tes menunjukkan kualitas yang baik.

Tabel 2. Mean dan Varian Tingkat Kesukaran, Dayabeda Item dan Reliabilitas

	N	Mini	Maxi	Mean	Std.	Variance	Kurtosis	
	Statis	Range	Mum	mum	Deviation			Std.
	Tik	Statis	Statistic	Statistic	Statistic	Statistic	Statistic	Error
ALPHACRON	30	,467	,345	,812	,66123	,171976	,030	,833
MEANP	30	,310	,263	,573	,43353	,111318	,012	,833
MEANBIS	30	,250	,369	,619	,49887	,096523	,009	,833
N	30							

Pada pembahasan di atas ditunjukkan oleh dinalai rata-rata dan selanjutnya akan diperhatikan variabilitas dari ketiga statistic tersebut, dengan memperhatikan rentang angka maksimum dan minimum, standar deviasi akan memungkinkan variasi angka –angka yang dapat dianalisis lebih lanjut dalam rangka mengetahui variable yang mempengaruhi reliabilitas..

Analisis Pengujian Hipotesis

Analisis pengujian hipotesis pengaruh dua variabel bebas terhadap variabel tergantung, variabel bebas tingkat kesukran item dan tingkat dayabeda item sedangkan variabel tergantung adalah reliabilitas Alpha Cronbach. Teknik regresi ganda diawali dengan perhitungan korelasi antar variabel yang diteliti. Hasil analisis data pada variabel penelitan ditampilkan pada Tabel 2 berikut:

Tabel 3. Koefisien Korelasi Mean P dan Mean Biserial dengan Reliabilitas

		ALPHACRON	MEANP	MEANBIS
Pearson Correlation	ALPHACRON	1.000	.867	.883
	MEANP	.867	1.000	.839
	MEANBIS	.883	.839	1.000
Sig. (1-tailed)	ALPHACRON	.	.000	.000
	MEANP	.000	.	.000
	MEANBIS	.000	.000	.

Tabel 3 di atas menunjukkan bahwa variabel bebas rata-rata tingkat kesulitan berkorelasi dengan rata-rata reliabilitas sebesar $r = 0,867$ dan dayabeda item dengan reliabilitas Alpha Cronbach diperoleh $r = 0,883$, kedua koefisien korelasi tersebut menunjukkan korelasi yang tinggi dan tingkat signifikansi 0.00 atau lebih kecil dari 0.05. Dengan demikian secara meyakinkan bahwa kedua variabel mempengaruhi reliabilitas.

Analisis regresi memerlukan uji asumsi klasik (a) Multikolenieritas, (b) Heteroskdistisitas, (c) normalitas, (d) Autokorelasi. Multikolenieritas dimana akan dikur tingkat asosiasi (keeratn) hubungan antar variabel bebas lebih besar dari 0,60 (pendapat lain 0,50 dan 0,90). Dikatakan tidak terjadi multikolenieritas jika koefisien korelasi antar variabel bebas lebih kecil atau sama dengan 0,60 ($r < 0,60$) (Danang Sunyoto, 2009: 79). Hasil analisis yang dikemukakan pada tabel 02 tingkat kesulitan item dengan dayabeda item berkorelasi 0,839 dengan tingkat signifikansi 0.00. Menunjukkan kedua variabel bebas berkorelasi tinggi lebih besar dari 0,60. Kedua variabel bebas yaitu tingkat kesulitan dan dayabeda item sebenarnya hanya satu variabel bebas sehingga analisis regresi yang digunakan

tidak perlu menggunakan teknik analisis regresi ganda akan tetapi cukup teknik regresi sederhana. Regresi sederhana merupakan teknik regresi yang digunakan untuk menganalisis data satu variabel bebas dan satu variabel tergantung.

Hasil analisis ini berbeda dengan analisis teori modern atau teori respons butir keakuratan atau reliabilitas ditunjukkan pada fungsi informasi, fungsi informasi akan maksimal dipengaruhi oleh dayabeda item sedangkan tingkat kesulitan item memberikan pengaruh yang kecil (Embretson, 2000: 169 - 170). Dayabeda item yang tinggi akan mengakibatkan fungsi informasi maksimal atau dapat dikatakan tingginya dayabeda akan lebih mempengaruhi reliabilitas tes dibandingkan dengan tingkat kesulitan item yang tinggi. Perbedaan makna pengaruh variabel bebas dalam teori klasik dengan teori modern menunjukkan sensitifitas tinggi pengaruh kualitas item terhadap reliabilitas tes pada teori modern, berarti teori modern lebih sensitif terhadap reliabilitas tes. Akan tetapi teori modern ini merupakan teori yang berkembang mutakhir dan dianggap teori yang kaku "rigid" memiliki persyaratan tinggi dalam penerapannya.

Tabel 4. Perhitungan ANOVA Tingkat Kesulitan, Dayabeda Item Terhadap Reliabilitas (Uji Serempak)

Model	Sum of Squares	Df	Mean Square	F	Sig.
Regression	.715	2	.357	67.625	.000 ^a
Residual	.143	27	.005		
Total	.858	29			

a. Predictors: (Constant), MEANBIS, MEANP

b. Dependent Variable: ALPHACRON

Tabel 4 di atas mengacu pada uji serempak pengaruh dua variabel bebas terhadap variabel tergantung diperoleh nilai F hitung sebesar 67,625 dengan tingkat signifikansi 0,000 lebih kecil dari signifikansi 0,05 sehingga dapat dikatakan bahwa H_0 statistik ditolak dan H_a penelitian diterima. Hipotesis statistik: tidak ada pengaruh dua variabel bebas terhadap variabel tergantung sedangkan hipotesis penelitian menyatakan ada pengaruh dua variabel bebas terhadap variabel tergantung. Karena H_a diterima maka dapat diartikan

bahwa ada pengaruh dua variabel bebas terhadap variabel tergantung atau ada pengaruh tingkat kesulitan item dan dayabeda terhadap reliabilitas Alpha Cronbach.

Selanjutnya dianalisis bagaimana pengaruh masing-masing variabel bebas terhadap variabel tergantung secara parsial, pengaruh tingkat kesulitan item (MeanP) dan dayabeda item (MeanBis) terhadap reliabilitas Alpha Cronbach, hal ini akan dikemukakan pada tabel berikut besarnya nilai T dan masing-masing signifikansinya.

Tabel 5. Perhitungan ANOVA Tingkat Kesulitan, Dayabeda Item Terhadap Reliabilitas (Uji Parsial)

Model	Standardized											
	Unstandardized					95.0% Confidence				Collinearity		
	Coefficients		Statistics			Interval for B		Correlations		Statistics		
	B	Std. Error	Beta	t	Sig.	Lower Bound	Upper Bound	Zero-order	Partial	Part	ce	VIF
(Constant)	-.092	.072		-1.274	.213	-.239	.056					
MEANP	.657	.223	.425	2.949	.007	.200	1.114	.867	.494	.232	.297	3.371
MEANBIS	.938	.257	.527	3.653	.001	.411	1.465	.883	.575	.287	.297	3.371

Dependent Variable: ALPHACRON

Dari Tabel 05 di atas dapat diketahui bahwa analisis parsial menunjukkan pengaruh tingkat kesulitan item (MeanP) diperoleh nilai T sebesar 2,949 dengan signifikansi 0,007 dimana tingkat signifikansi yang diperoleh lebih dari 0,05 sehingga dapat diartikan bahwa tingkat kesulitan item mempengaruhi reliabilitas tes Alpha Cronbach. Pengaruh tingkat dayabeda item (MeanBis) diperoleh nilai T sebesar 3,653 dengan signifikansi 0,001 dimana tingkat signifikansi yang diperoleh lebih dari 0,05 sehingga dapat diartikan bahwa dayabeda item mempengaruhi reliabilitas tes Alpha Cronbach.

Pembahasan yang kaitkan dengan rumus dalam perhitungan, konsep konsistensi internal item mengandung makna dan tafsiran yang berbeda-beda menurut para pakar di bidang psikometrika. Cronbach (1951) menggunakan istilah “konsistensi internal” dan “homogenitas” secara bergantian dan menyatakan bahwa internal konsistensi tes atau homogenitas harus independen dengan panjang tes (Wei Tang, Ying Cui, Oksana Babenko, 2014 : 205). Konsistensi internal mengacu pada makna reliabilitas, mengartikan istilah konsistensi internal dengan validitas item. Konsistensi internal dalam penelitian ini adalah koefisien Alpha Cronbach yang diperoleh dari rumus dikeukakan di atas. Berbeda dengan koefisien konsistensi internal item melalui proses perhitungan dimana setiap item dikorelasikan dengan skor total dimana item tes berasal dari tes skor dikotomi dan menghasilkan konsistensi internal untuk setiap item diperoleh dengan rumus KR20, sedangkan koefisien Alpha Cronbach diterapkan secara lebih luas yaitu dapat pada skoring dikotomi dan politomi (0, 1, 3 atau lebih) sehingga dapat diterapkan pada data penskalaan. Sebagai angka akhir dari koefisien Alpha Cronbach akan diperoleh pada setiap subtes bukan pada setiap item tes, dengan demikian untuk memperoleh angka yang lebih meyakinkan membutuhkan data yang jauh lebih banyak dibandingkan KR20.

Penelitian ini menunjukkan bahwa korelasi antara dua variabel bebas antara tingkat kesulitan dan dayabeda item berkorelasi $r = 0,839$ lebih besar dari batas toleransi 0,6. Sehingga dapat disimpulkan bahwa tingkat kesulitan item dan dayabeda item terjadi multikolinieritas, dua variabel yang mempunyai makna yang sama sehingga dalam penelitian ini dapat dianalisis salah satu dari keduanya sehingga analisis berubah menjadi regresi sederhana dimana satu variabel bebas dan satu variabel tergantug.

Pada bagian penjelasan instrumen pengumpulan data menggunakan tes kemampuan berpikir numerikal yang tujuan sebagai alat untuk menentukan potensi dalam pemilihan jurusan di SMA, berbeda dengan TPA untuk seleksi masuk IKIP Mataram suatu tes yang bertujuan untuk mengerahui kemampuan yang dapat diprediksi dapat menyelesaikan studi di IKIP Mataram dengan kata lain suatu tes yang mempunyai validitas prediktif yang tinggi. Akan tetapi TPA verbal tidak mengukur kemampuan berpikir kritis calon mahasiswa. Subtes spatial yang biasanya dijumpai pada tes inteligensi seperti Progressive Matrics (PM), Culture Fair Intelligence Test (CFIT) mengukur analogi dan kini soal yang sejenis tersebut merupakan tes spatial. Item tes yang mengungkapkan kemampuan kognitif peserta melalui gambar mental yang dapat diangkat dalam item subtes spasial responden diminta mengamati pasangan-pasangan gambar dua dimensi (2 D) yang menunjukkan bentuk-bentuk geometris tiga dimensi (3 D). Saat gambar ditampilkan bentuknya dirotasi dari 0 sampai 180 derajat. (Sternberg, 2008: 244 - 245, Marios Pittalis, 2010:191). Rangkaian tes ini merupakan tes persepsi ruang, yang dikenakan sebagai factor penting pada pekerjaan yang melibatkan tugas mekanis. Seperti yang tersirat dari namanya, persepsi ruang adalah kemampuan memvisualisasikan benda pada tiga dimensi dan mengendalikannya untuk

menghasilkan konfigurasi tertentu. (Aiken, 2006: 257). Perhatikan kembali data yang tertuang pada Tabel 01 koefisien Alpha Cronbach yang diperoleh umumnya lebih tinggi dibandingkan dengan subtes TPA lainnya walaupun tidak setinggi tes kemampuan berpikir numerikal yang telah melalui proses ujicoba dan revisi.

SIMPULAN

Dari analisis data dapat disimpulkan bahwa tingkat kesulitan item dan dayabeda item mempengaruhi koefisien Alpha Cronbach (secara serempak) dengan nilai $F = 67,625$ dengan tingkat signifikansi = 0.000. Secara parsial pengaruh tingkat kesukaran item (Mean P) terhadap reliabilitas dengan nilai $t = 2.949$ (sig = 0.007) dan pengaruh rata-rata biserial item (Mean bis) diperoleh $t = 3.653$ (sig = 0.001).

Karena dalam penelitian ini terjadi gejala multikolenieritas dua variabel bebas yang berkorelasi tinggi sesamanya sehingga dua variabel bebas itu dapat diambil salah satunya dan analisisnya berubah dari regresi ganda menjadi regresi sederhana.

SARAN

Dengan demikian maka disarankan untuk selalu mengembangkan tes yang sesuai dengan kebutuhan seleksi mahasiswa baru baik dari segi isi tes dan dari segi prosedur analisis dan ketajaman analisis.

DAFTAR RUJUKAN

- Aiken Lewis R., Gary Groth-Marnat, (2006), *Psychological Testing and Assessment*, Pearson Education Group, Boston.
- Anne, Anastasi, (1998), Susana Urbina, *Psychological Testing*, jilid 1 dan 2, alih bahasa Robertus Hariono S.Iman, Jakarta: Prehalindo,
- Azwar, Saifuddin (2003), *Reliabilitas dan Validitas*, Pustaka Pelajar, Yogyakarta
- Cohen, Ronald Jay and Mark E. Swerdlik, (2005).; Mc Graw Hill, Boston.
- Crocker, Linda, James Algina, (1986), *Introduction to Classical and Modern Test Theory*, Tokyo: Harcourt Brace Jovanovich College Publishers.
- Ebel, Robert L. dan David A Frisbie, (1991), *Essentials of Educational Measurement*, New Jersey: Prentice Hall.
- Gregory, Robert J, (2013) *.Psychological Testing, History Pinciples and Applications*, Alih Bahasa Prof Dr. Amitya Kumara (Universitas Gajah Mada), Penerbit Erlangga, Yogyakarta.
- Gronlund, Norman E, and Robert Linn, (1990), *Measurement and Evaluation in Teaching*, Mcmillan, Publishing Company, New York.
- Hambleton, Ronald K, H. Swaminathan, H.Jane Rogers (1991), *Fundamentals of Item Response Theory*, Sage Publication, International Professional Publishers, (London:
- Hayat, Bahrul, Sumarna S Prananto, Suprananto, (1997), *Manual Item and Test Analysis (ITEMAN)*, Pusat Penelitian dan Pengembangan Sistem Pengujian Badan Penelitian dan Pengembangan Pendidikan dan Kebudayaan, Departemen Pendidikan dan Kebudayaan, Jakarta.
- Jumailiyah. (2015, 22-23 May 2015). *Pengembangan Tes Bakat Berpikir Numerikal Model Logistik Dua Parameter*. Makalah disajikan pada Seminar dan Workshop Internasional Konseling Malindo ke-4, **Denpasar, Bali**.
- Kaplan, Robert M, Dennis P.Saccuzzo, (2012), *Psychological Testing Principles Applications and Issues*, Penerjemah Eko Prasetyo Widodo, Penerbit Salemba Humanika. Jakarta.
- Nitko, Anthony J, (2001), *Educational Assessment of Students* (New Jersey: Merrill Prentice Hall.
- Nunnally, Jum C, Ira H.Bernstein, (1994), *Psychometric Theory*. Third Edition, McGraw-Hill.Inc, New York.