

RELIABILITAS DAN VALIDITAS AITEM

Saifuddin Azwar

Pada Buletin Psikologi yang lalu telah diketengahkan artikel mengenai salah-satu prosedur analisis aitem dan penggunaannya dalam seleksi aitem untuk penyusunan suatu skala psikologi (Azwar, 1994).

Parameter aitem yang dipakai sebagai kriteria pemilahan aitem, yang dijelaskan dalam artikel tersebut, adalah koefisien korelasi antara skor setiap aitem dengan skor total tes atau skala - biasa disebut **korelasi aitem-total** - baik yang belum dikoreksi (r_{ix}) maupun yang telah dikoreksi dari efek *spurious overlap* ($r_{i(x-i)}$).

Koefisien korelasi aitem-total ini sering disalahartikan sebagai koefisien validitas aitem, padahal validitas aitem memiliki pengertian yang sama sekali berbeda. Koefisien korelasi aitem-total tidak memberikan informasi mengenai validitas aitem tetapi memberikan informasi mengenai **konsistensi** antara apa yang diukur oleh aitem dengan apa yang diukur oleh tes. Aitem-aitem yang berkorelasi tinggi dengan skor total berarti memiliki kesesuaian tujuan ukur dengan tes. Kesesuaian tujuan ukur antara aitem dengan tes ini tidak menjamin apakah tes tersebut memang mengukur apa yang ingin diukur (valid) dikarenakan validitas menghendaki adanya suatu kriteria lain. Oleh karena itu, sekalipun korelasi aitem-total pada kesemua aitem sudah baik, masih diperlukan pengujian validitas tes.

Tulisan ini diharapkan berguna sebagai suatu klarifikasi mengenai pengertian reliabilitas dan validitas aitem.

Indeks Reliabilitas Aitem

Indeks reliabilitas aitem dirumuskan sebagai:

$$ira = s_i \cdot r_{ix}$$

ira = Indeks reliabilitas aitem

s_i = Deviasi standar skor aitem i

r_{ix} = Koefisien korelasi aitem i dan skor total tes X

Apabila aitem-aitem dalam tes diberi skor dikotomi (hanya berisi angka 0 dan 1) maka indeks reliabilitas aitem dapat dituliskan sebagai:

$$r_{ix} = (p_i q_i) r_{ix}$$

p_i = Proporsi subjek yang mendapat skor 1 pada aitem i

q_i = $1 - p_i$

r_{ix} = Koefisien korelasi point-biserial antara skor aitem i dengan skor total tes X

Untuk memperjelas uraian, perhatikan data fiktif hasil tes yang disajikan pada Tabel 1.

Tabel 1. Skor Aitem Dikotomi, Skor Tes X, dan Skor Kriteria Y

Skor aitem (i)				Skor Tes (X)	Skor Kriteria (Y)
10	20	30	40		
1111111111	1111111111	1111111111	1111111111	40	49
1100000000	0000000000	0000000000	0000000000	2	17
1000000000	0001110000	0000000000	0000000000	4	10
1111011111	1111011111	1111111111	1111111110	37	34
1111111111	1111100000	0000000000	0000000000	15	27
1110111110	0000000000	1000000000	0001000000	10	22
1111111111	0111111111	1111111111	1111111100	37	27
1111101111	1111000000	0000000000	0000000000	13	37
1110000000	0000000000	0000000000	0000000000	3	17
1111111111	0111111110	1111111111	1000000000	29	30
1111111111	1111111111	1111111111	1111000000	34	33
1110000000	0000000000	0000111000	0000000000	6	12
1111111111	1011111111	1100000010	0000000000	22	32
1110111111	0111111111	1000000000	0010000000	20	30
1011111110	1111111111	1111111100	0000000000	26	27
1111110111	1110000000	0000000000	0000000000	12	29
1111111111	1000000000	0000000000	0000000000	11	34
1111111111	1111111111	1111100000	0000000000	25	25
1111111111	1111111110	0000000000	0000000000	19	22
0111111111	1111111111	1111111111	1111110000	35	45
1111111111	1111111000	0000000000	0000000000	17	22
1111011111	1111110000	0000001000	0000000000	16	21
1111111000	0000000000	0000000000	0000000000	7	20
1100000000	0000000000	0000000000	0000000000	2	10
1111100100	0000000000	0010000000	0000000000	7	18
1111111111	1000000000	0000000000	0000000000	11	33

Skor aitem (i)				Skor Tes (X)	Skor Kri- teria (Y)
10	20	30	40		
1111100000	0000000000	0100000000	0000000000	6	27
1111111111	1011111111	1111111111	0000000000	29	37
1111111000	0000010000	0000110000	0000000000	10	23
1111111110	0000000000	0000000000	0000000000	9	34
1111110111	1111110000	0000000000	0100000000	16	36
1111111111	1101111100	0000100000	0100000000	19	36
1111111111	1111111111	1000000000	0000000000	21	39
1111111111	1111111111	1111000000	0000000000	24	40
1111000000	0000000000	0001000000	0000001000	6	17
1111111011	1111101111	1111110000	0000100000	25	43
1111110000	0000000000	0000000000	0000000000	6	24
1111111111	1111101111	1111111111	1110000000	32	44
1111000000	0000000000	0000000000	0000000000	4	17
1011111111	1111111110	1111111111	1111110000	34	43

Tabel 1. menyajikan data hasil tes yang terdiri dari 40 aitem yang diberi skor dikotomi. Skor tes tersebut, yang diberi simbol X, dan skor kriteria validasinya, yang diberi simbol Y, juga disajikan.

Komputasi reliabilitas tes X pada Tabel 1. menghasilkan koefisien reliabilitas $r_{xx'} = 0,9663$ dengan $s_x^2 = 128,1537$. Korelasi antara skor X dengan skor Y besarnya adalah $r_{xy} = 0,7350$ dan merupakan koefisien validitas tes X.

Tabel 2. menyajikan hasil komputasi deviasi standar aitem, korelasi aitem-total, dan indeks reliabilitas aitem dari data pada Tabel 1.

Karena indeks reliabilitas aitem merupakan hasil perkalian antara deviasi standar aitem dengan korelasi aitem-total, dan karena deviasi standar aitem selalu berharga positif, maka indeks reliabilitas aitem akan berharga negatif manakala korelasi aitem-totalnya negatif.

Lebih lanjut, dikarenakan deviasi standar aitem dikotomi (*biased estimation*) tidak lebih besar daripada 0,50 dan korelasi aitem-total tidak dapat lebih besar daripada 1,00 maka indeks reliabilitaspun tidak akan dapat lebih besar daripada 0,50. Keseluruhan indeks reliabilitas aitem ini apabila dijumlahkan maka besarnya akan sama dengan deviasi standar skor tes yang bersangkutan (dalam contoh ini besarnya adalah 11,32).

Tabel 2. Indeks Reliabilitas Aitem untuk Tes X

Aitem	s	r_{ix}	ira	Aitem	s	r_{ix}	ira
11	.1581	-.2503	-.04	121	.5006	.8237	.41
12	.2667	-.0970	-.03	122	.4903	.7952	.39
13	.2667	.3785	.10	123	.4830	.7908	.38
14	.3848	.4454	.17	124	.4830	.7861	.38
15	.4229	.4163	.18	125	.4903	.7166	.35
16	.4385	.6314	.28	126	.4743	.6980	.33
17	.4641	.6652	.31	127	.4641	.7062	.33
18	.4641	.6310	.29	128	.4385	.8148	.36
19	.4641	.7189	.33	129	.4385	.7941	.35
110	.4903	.7155	.35	130	.4229	.7995	.34
111	.5006	.5878	.29	131	.4051	.7705	.31
112	.5038	.7213	.36	132	.4229	.6763	.29
113	.5006	.7914	.40	133	.4051	.7202	.29
114	.4961	.7688	.38	134	.3848	.6141	.24
115	.5038	.6898	.35	135	.3616	.6442	.23
116	.5057	.6538	.33	136	.3349	.6450	.22
117	.5057	.8600	.43	137	.3038	.3720	.11
118	.5038	.8656	.44	138	.2667	.5216	.14
119	.5006	.8645	.43	139	.2207	.4305	.10
120	.4830	.7580	.37	140	.1581	.3220	.05

Indeks Validitas Aitem

Validitas tes diperlihatkan oleh korelasi antara skor tes dan skor suatu kriteria (*criterion-related validity*). Kriteria validasi dapat berupa tes lain yang mengukur atribut yang sama, dapat berupa variabel yang hendak diprediksikan oleh tes, dan dapat pula berupa ukuran-ukuran lain yang relevan seperti *rating* atau hasil observasi yang reliabel.

Koefisien validitas aitem juga diperlihatkan oleh korelasi antara skor aitem dengan skor suatu kriteria yang relevan, yaitu kriteria validasi tes yang bersangkutan, sedangkan indeks validitasnya diperoleh dari perkalian koefisien korelasi tersebut dengan deviasi standar aitem yang bersangkutan.

Indeks validitas aitem dirumuskan sebagai:

$$iva = s_i \cdot r_{iy}$$

s_i = Deviasi standar skor aitem i

r_{iy} = Koefisien korelasi aitem i dengan skor suatu kriteria eksternal Y
(koefisien validitas aitem)

Bila aitem-aitem dalam tes diberi skor dikotomi, maka indeks validitas aitem dapat dituliskan sebagai:

$$iva = (p_i q_i) r_{iy}$$

Tabel 3. memuat deviasi standar setiap aitem, koefisien korelasi aitem-kriteria, dan indeks validitas aitem dari data pada Tabel 1.

Tabel 3. Indeks Validitas Aitem untuk Tes X

Aitem	s	r_{iy}	iva	Aitem	s	r_{iy}	iva
I1	.1581	-.2656	-.04	I21	.5006	.5832	.29
I2	.2667	.0549	.01	I22	.4903	.5598	.27
I3	.2667	.4669	.12	I23	.4830	.5025	.24
I4	.3848	.5449	.21	I24	.4830	.4972	.24
I5	.4229	.6176	.26	I25	.4903	.4139	.20
I6	.4385	.6048	.27	I26	.4743	.4070	.19
I7	.4641	.5998	.28	I27	.4641	.3255	.15
I8	.4641	.5833	.27	I28	.4385	.4853	.21
I9	.4641	.7210	.33	I29	.4385	.5144	.23
I10	.4903	.6967	.34	I30	.4229	.5127	.22
I11	.5006	.6678	.33	I31	.4051	.4821	.20
I12	.5038	.5599	.28	I32	.4229	.5430	.23
I13	.5006	.5861	.29	I33	.4051	.4821	.20
I14	.4961	.5317	.26	I34	.3848	.3519	.14
I15	.5038	.4533	.23	I35	.3616	.4917	.18
I16	.5057	.3080	.16	I36	.3349	.4207	.14
I17	.5057	.5817	.29	I37	.3038	.1069	.03
I18	.5038	.6172	.31	I38	.2667	.2326	.06
I19	.5006	.5832	.29	I39	.2207	.2994	.07
I20	.4830	.5554	.27	I40	.1581	.3302	.05

Dari Tabel 2. dan Tabel 3. tampak bahwa aitem yang memiliki koefisien korelasi aitem-total rendah cenderung akan memiliki korelasi aitem-kriteria yang rendah pula. Walaupun secara keseluruhan validitas tes X sudah memuaskan akan tetapi beberapa diantara aitemnya sendiri berkorelasi rendah, bahkan negatif, dengan kriteria validasi tes (Y).

Indeks reliabilitas dan indeks validitas aitem menjadi penting dikarenakan rumusnya yang melibatkan varians skor aitem. Varians skor aitem memiliki kontribusi dalam reliabilitas dan validitas tes.

Fungsi indeks reliabilitas dan indeks validitas aitem dalam penyusunan tes psikologi adalah pada seleksi aitem yang bertujuan peningkatan reliabilitas dan validitas tes secara keseluruhan.

Penggunaan indeks reliabilitas dan indeks validitas aitem menjadi penting dalam pemilihan aitem apabila reliabilitas dan validitas tes secara keseluruhan belum diketahui atau sudah diketahui dan hasilnya belum memuaskan. Apabila reliabilitas tes secara keseluruhan sudah memuaskan, pemilihan aitem dapat didasarkan pada koefisien korelasi aitem-total saja.

Indeks reliabilitas aitem berkaitan dengan tingginya reliabilitas tes. Aitem yang variansnya lebih besar memiliki kontribusi yang lebih besar pula terhadap reliabilitas tes (Magnusson, 1966; Allen & Yen, 1979; Crocker & Algina, 1986). Karena itu, bila harus memilih satu diantara dua aitem yang koefisien korelasi aitem-totalnya sama hendaklah dipilih yang memiliki varians yang lebih besar. Sebaliknya, apabila tes berisi aitem yang variansnya homogen (deviasi standar aitem kurang lebih setara satu sama lain) maka pemilihan aitem dapat disandarkan pada besarnya koefisien korelasi aitem-total saja.

Memilih Aitem Berdasarkan *ira* dan *iva*

Koefisien validitas tes dinyatakan secara langsung oleh perbandingan antara jumlah indeks validitas aitem dengan jumlah indeks reliabilitas aitem. Validitas tes dengan kriteria akan maksimal apabila tes berisi aitem-aitem yang memiliki indeks validitas setara indeks reliabilitasnya (Allen & Yen, 1979). Oleh karena itu, guna meningkatkan validitas tes perlu dipilih aitem-aitem yang memiliki indeks reliabilitas dan indeks validitas kurang lebih setara.

Bila pertimbangan seleksi aitem adalah peningkatan reliabilitas konsistensi internal tes se-maksimal mungkin maka perlu dipilih aitem-aitem yang memiliki koefisien korelasi aitem-total yang tinggi. Pada Tabel 4. dimuat aitem-aitem yang dipilih berdasar tujuan dan kriteria yang berbeda.

Tabel 4. Indeks Reliabilitas dan Indeks Validitas Aitem sebagai kriteria Pemilihan Aitem

Aitem	s	riy	iva	rix	ira
I1	.1581	-.2656	-.04	-.2503	-.04
I2	.2667	.0549	.01	-.0970	-.03
I3	.2667	.4669	.12	.3785	.10 *
I4	.3848	.5449	.21	.4454**	.17 *
I5	.4229	.6176	.26	.4163**	.18 *
I6	.4385	.6048	.27	.6314**	.28 *
I7	.4641	.5998	.28	.6652**	.31 *
I8	.4641	.5833	.27	.6310**	.29 *
I9	.4641	.7210	.33	.7189**	.33 *
I10	.4903	.6967	.34	.7155**	.35 *
I11	.5006	.6678	.33	.5878**	.29 *
I12	.5038	.5599	.28	.7213**	.36 *
I13	.5006	.5861	.29	.7914**	.40
I14	.4961	.5317	.26	.7688**	.38
I15	.5038	.4533	.23	.6898**	.35
I16	.5057	.3080	.16	.6538**	.33
I17	.5057	.5817	.29	.8600**	.43
I18	.5038	.6172	.31	.8656**	.44
I19	.5006	.5832	.29	.8645**	.43
I20	.4830	.5554	.27	.7580**	.37 *
I21	.5006	.5832	.29	.8237**	.41
I22	.4903	.5598	.27	.7952**	.39
I23	.4830	.5025	.24	.7908**	.38
I24	.4830	.4972	.24	.7861**	.38
I25	.4903	.4139	.20	.7166**	.35
I26	.4743	.4070	.19	.6980**	.33
I27	.4641	.3255	.15	.7062**	.33
I28	.4385	.4853	.21	.8148**	.36
I29	.4385	.5144	.23	.7941**	.35
I30	.4229	.5127	.22	.7995**	.34
I31	.4051	.4821	.20	.7705**	.31
I32	.4229	.5430	.23	.6763**	.29 *
I33	.4051	.4821	.20	.7202**	.29 *
I34	.3848	.3519	.14	.6141**	.24 *
I35	.3616	.4917	.18	.6442**	.23 *
I36	.3349	.4207	.14	.6450**	.22 *
I37	.3038	.1069	.03	.3720	.11 *
I38	.2667	.2326	.06	.5216**	.14 *
I39	.2207	.2994	.07	.4305**	.10 *
I40	.1581	.3302	.05	.3220	.05 *

* = Aitem terpilih berdasarkan kesetaraan ira dan iva

** = Aitem terpilih berdasarkan tingginya rix

Bila dipilih 36 aitem berdasarkan tingginya r_{xx} sebagaimana dalam Tabel 4., diperoleh koefisien reliabilitas (konsistensi internal) sebesar $r_{xx}' = 0,9698$ dan koefisien validitas tes sebesar $r_{xy} = 0,7307$. Sebagaimana dikatakan di muka, pemilihan aitem berdasarkan korelasi aitem-total ini akan meningkatkan reliabilitas tes tetapi sedikit mengurbankan validitas. Hal ini tampak pada peningkatan koefisien reliabilitas dari 0,9663 menjadi 0,9698. Kecilnya peningkatan ini dikarenakan reliabilitas tes semula memang sudah sangat tinggi. Tampak pula terjadinya penurunan koefisien validitas dari 0,7350.

Bila aitem-aitem dipilih berdasarkan kesetaraan indeks reliabilitas dan indeks validitas aitem guna peningkatan validitas tes, pada Tabel 4. dari 20 aitem diperoleh validitas tes $r_{xy} = 0,8041$. Validitas ini merupakan peningkatan yang substansial dari koefisien semula yang hanya 0,7350. Bersamaan dengan peningkatan validitas ini, terjadi penurunan koefisien reliabilitas menjadi $r_{xx}' = 0,9213$.

Tentu saja yang paling diinginkan adalah meningkatkan reliabilitas dan validitas tes sekaligus. Sayangnya hal itu tidak dapat dicapai oleh suatu komposisi aitem yang sama. Suatu kumpulan aitem yang meningkatkan reliabilitas akan mengurangi validitas dan sebaliknya suatu komposisi aitem yang meningkatkan validitas akan mengurangi reliabilitas. Oleh karena itu prioritasnya dikembalikan kepada penyusun tes sendiri, apakah lebih mementingkan peningkatan validitas atau lebih mementingkan peningkatan reliabilitas.

Demikianlah sekilas uraian mengenai indeks reliabilitas dan indeks validitas aitem serta salah-satu fungsinya dalam seleksi aitem. Dalam pengembangan tes kognitif, fungsi indeks reliabilitas aitem merupakan pertimbangan penting pula dalam penetapan varians/penyebaran skor tes.

Referensi

- Allen, M.J. & Yen, W.M. 1979, *Introduction to Measurement Theory* Monterey: Brooks/Cole Publishing Company.
- Azwar, S., Seleksi Aitem dalam Penyusunan Skala Psikologi, *Buletin Psikologi*, Thn. II No. 2 Desember 1994 h.26-33.
- Crocker, L. & Algina, J. 1986 *Introduction to Classical and Modern Test Theory*. Forth Worth: Holt, Rinehart, and Winston, INC.
- Magnusson, D. 1967, *Test Theory*, Reading: Addison-Wesley Publishing Company.