

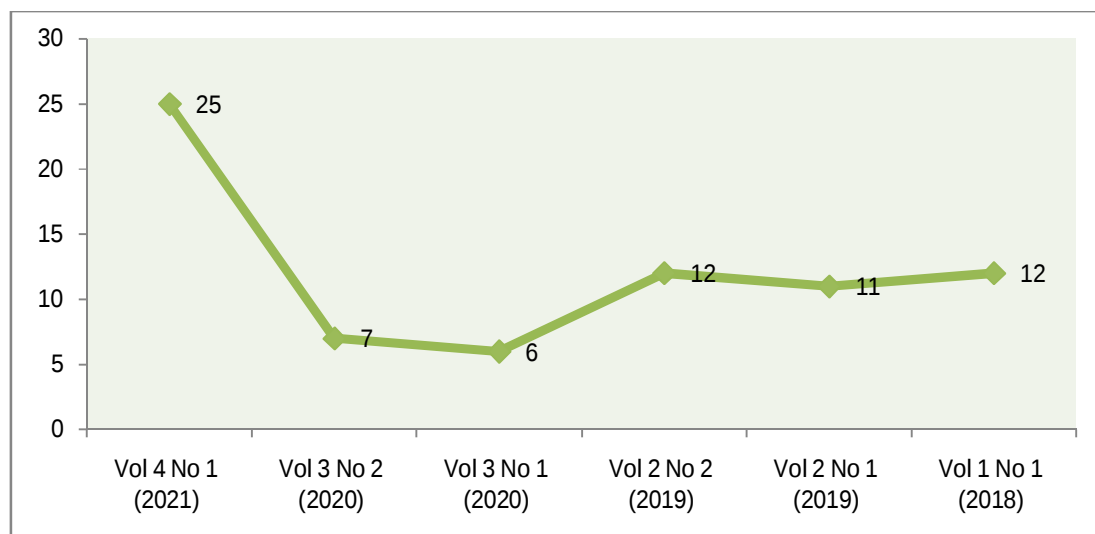
Editorial Note: Uji Validitas Dengan Korelasi Item-Total ?

Hendryadi
Email: hendry.basrah@gmail.com

Hendryadi is a lecturer at Sekolah Tinggi Ilmu Ekonomi Indonesia Jakarta, Jakarta, Indonesia. He is a part of the reviewers in some scientific journals and is the author and co-author of research papers. He is Head of the Research Center at the Imperium Research Institute, Jakarta. His current research interests include Islamic work ethics, leadership, workplace incivility, and quantitative research methods. Email: hendry.basrah@gmail.com. ORCID: <http://orcid.org/0000-0002-1579-8487>

PENDAHULUAN

Dewan redaksi JMSAB telah mempublikasikan artikel mulai edisi pertama sampai dengan Vol 4(1) adalah sebanyak 73 artikel. Tren peningkatan minat penulis untuk mempublikasikan artikelnya di JMSAB juga meningkat seiring dengan ditetapkannya JMSAB sebagai jurnal terakreditasi SINTA 3 pada awal tahun 2020 dan berlaku mulai edisi pertama 2018 sampai dengan Vol 5, No 2, Desember 2022.



Gambar 1.
Tren Jumlah Artikel JMSAB Periode 2018 – 2021

Saya ingin meninjau situasi JMSAB, selaku Managing Editor jurnal ini. Poin-poin penting yang ingin disampaikan dalam editorial note ini khusus mengenai penggunaan sampel uji validitas.

Uji Validitas

Validitas mengacu pada aspek ketepatan dan kecermatan hasil pengukuran. Dalam konteks Teori Skor-murni Klasik, Azwar (2012) menjelaskan bahwa makna validitas dapat dinyatakan sebagai sejauhmana besaran skor-tampak X mampu mendekati besaran skor-murni T . Semakin skor-tampak mendekati skor-murni berarti semakin tinggi validitas dan sebaliknya, semakin rendah validitas hasil pengukuran berarti semakin besar perbedaan skor-tampak dari skor-murni. Validitas adalah sejauh alat ukur (tes) benar-benar menggambarkan apa yang hendak diukur sehingga kesalahan dalam pengujian validitas akan berdampak fatal pada hasil penelitian secara keseluruhan. Dengan demikian, kita dapat menggambarkan pengujian validitas sebagai upaya peneliti untuk memastikan bahwa alat ukur yang digunakan memiliki ketepatan dari sisi konten, sehingga dapat menggambarkan keseluruhan dari variabel yang hendak diukur. Ada tiga jenis validitas yaitu *content validity* (validitas isi), *criterion validity* (validitas kriteria), dan *construct validity* (validitas konsep). Saya mempelajari lebih dari 70 persen artikel yang masuk ke JMSAB selama 3 tahun terakhir ini memiliki kekeliruan yang sama, yaitu dalam penggunaan teknik analisis dan minimnya sampel uji.

Kesalahan teknik analisis

Rata-rata artikel yang masuk menggunakan teknik Korelasi Item-Total (*item-total correlation*) atau kita singkat "ITC" untuk mengevaluasi validitas item. Sebenarnya jika merujuk pada beberapa referensi psikometrik (lihat artikel Widhiarso, 2012) yang berjudul "(Sekali Lagi) Korelasi Aitem Total Bukanlah Validitas Item" yang menjelaskan secara singkat mengenai kekeliruan yang banyak terjadi dalam pengujian validitas menggunakan teknik ini.

Lalu apa kegunaan korelasi item-total ?

Dasar kerja yang digunakan dalam analisis aitem dalam hal ini adalah memilih item-item yang fungsi ukurnya sesuai dengan fungsi ukur test seperti dikehendaki penyusunnya. Dengan kata lain adalah memilih item yang mengukur hal yang sama dengan apa yang diukur oleh tes secara keseluruhan. Pengujian keselarasan fungsi item dengan fungsi ukur tes dilakukan dengan menghitung koefisien korelasi antara distribusi skor pada setiap item dengan distribusi skor total tes itu sendiri. Prosedur ini akan menghasilkan koefisien korelasi item total (r_{it}) yang juga dikenal dengan sebutan parameter daya beda item. (*item discrimination*). Menurut Azwar (2012), sebagai kriteria pemilihan item berdasarkan korelasi item-total, biasanya digunakan batasan koefisien > 0.30 . Semua item yang mencapai koefisien korelasi minimal 0.30 data pembedanya dinyatakan memenuhi syarat psikometrik sebagai bagian dari test. Namun kriteria ini tentu saja tidak menjadi patokan tunggal dalam menentukan item mana yang akhirnya diikuti sebagai bagian test final. Perlu juga memperhatikan cakupan isi domain (*content of the test domain*) dan tujuan penggunaan hasil test (Azwar, 2012).

Koreksi Terhadap Efek Spurious Overlap

Terkait dengan korelasi item-total ini, lanjut Azwar (2012) memberikan penjelasan bahwa bahwa dalam kaitannya dengan hasil komputasi koefisien korelasi antara korelasi item dengan skor total ada kekhawatiran terkait dengan jumlah item yang ada dalam tes yang dapat

Hendryadi

Uji Validitas dengan korelasi item total ?

mengakibatkan *over estimate* dikarenakan besarnya kontribusi item dalam menentukan skor test. Dengan kata lain, korelasi item-total adalah korelasi antara skor item dengan skor total yang merupakan bagian dari item itu sendiri. Hal ini selanjutnya menyebabkan koefisien korelasinya cenderung lebih tinggi daripada kalau korelasi itu dihitung antara skor item dengan skor test yang tidak mengandung item yang bersangkutan. Keadaan inilah yang disebut *spurious overlap* (Guilford, dalam Azwar, 2012:158).

Akibat terjadinya *spurious overlap* ini adalah *over estimate* terhadap korelasi item yang bersangkutan dengan skor test. Untuk menghilangkan efek *spurious overlap* tersebut maka koefisien korelasi item total yang dihitung dengan korelasi *product moment pearson* dikoreksi dengan nilai simpangan baku (*standard deviation*). Formula ini kemudian menghasilkan koefisien yang dikenal dengan *corrected item-total correlation coefficient*. Pada analisis ini berguna untuk menyeleksi item berdasarkan daya beda yang menghendaki akurasi tinggi (Azwar, 2012).

Dimana posisi korelasi item-total ini digunakan ?

Dalam SPSS, kita dapat melihat bahwa korelasi item-total merupakan bagian dari uji reliabilitas (konsistensi internal) dimana menu untuk mengaktifkan analisis ini tergabung dari pengujian reliabilitas.

Reliability Statistics	
Cronbach's Alpha	N of Items
.786	5

Item Statistics			
	Mean	Std. Deviation	N
Gaji	4.21	.725	400
Bonus	4.22	.732	400
Peralatan_kerja	4.08	.735	400
Rekan_Kerja	4.14	.727	400
Pengawasan	4.21	.734	400

Item-Total Statistics				
	Scale Mean if Item Deleted	Scale Variance if Item Deleted	Corrected Item-Total Correlation	Cronbach's Alpha if Item Deleted
Gaji	16.64	4.823	.581	.741
Bonus	16.63	4.820	.574	.743
Peralatan_kerja	16.77	4.854	.557	.749
Rekan_Kerja	16.70	4.916	.544	.753
Pengawasan	16.64	4.853	.559	.748

Gambar 2.

Contoh Uji Reliabilitas (Konsistensi internal) di SPSS

Pada tampilan Gambar 2 di atas kita dapat melihat bahwa bahwa seluruh item yang digunakan memiliki nilai *corrected item-total correlation* lebih besar dari 0.30. Hal ini mengindikasikan bahwa item yang digunakan memenuhi syarat konsistensi internal, dan mendukung pengujian koefisien Cronbach Alpha sebesar 0.788 yang sudah memenuhi

standar reliabilitas yang baik (> 0.70). Dengan demikian dapat disimpulkan bahwa lima item yang digunakan untuk mengukur kepuasan kerja telah memenuhi syarat reliabilitas dari sisi konsistensi internal. Jadi pada dasarnya, korelasi item-total merupakan uji konsistensi internal (keselarasan item) dan lebih tepat digunakan untuk pengujian reliabilitas. Secara jelas, LoBiondo-Wood & Haber (2013) dalam buku "*Nursing Research: Methods, Critical Appraisal & Utilization*" memasukkan korelasi item-total sebagai bagian dari uji konsistensi yang terkait dengan homogenitas (lihat LoBiondo-Wood & Haber, 2013, p. 297).

Untuk memastikan hal ini, saya juga akan menunjukkan berbagai artikel sebagai rujukan empiris dari penjelasan tentang kegunaan uji korelasi item total. Saya akan menampilkan beberapa kalimat yang dikutip dari berbagai artikel jurnal internasional yang memiliki reputasi dan sitasi tinggi dari berbagai publisher major di mesin pencari Google Scholar. Berikut ini disajikan kutipan asli dari berbagai artikel yang menunjukkan penggunaan teknik korelasi item-total:

1. "...Reliability was determined by calculating Pearson correlation coefficients and intra-class correlation coefficients in addition to Cronbach's alpha and item-total correlation at the beginning of the study...." (Schoppink et al., 1996)
2. "....Internal consistency reliability was calculated using item-total correlation and Cronbach α" (Quadrelli et al., 2009)
3. ".....To establish reliability, the item answers of all questionnaires were analysed by using Cronbach's Alpha correlation coefficient and corrected item total correlation....." (Westphal et al., 2002)
4. ".....Reliability was determined by internal consistency. Internal consistency was measured by calculating Cronbach's alpha and item-total correlation" (Bicer et al., 2004)
5. ".....Pearson's Moment Correlation Coefficient and Cronbach's Alpha were also used to establish reliability" (Aygör et al., 2018);
6. ".....The total item correlation coefficient was between 0.47–0.69 which shows that the items were in the reliability level..." (Şahin et al., 2020);
7. ".....internal consistency, item-total correlation, and test–retest reliability. Internal consistency refers to the extent to which the items are interrelated ... " (Akbulut et al., 2020);
8. ".....The Cronbach's α and the item-total correlation of each subscale of the PEACE questionnaire were calculated to assess internal consistency reliability" (Okamura et al., 2021)

Lalu teknik apa yang tepat untuk menguji validitas ?

Seperti di jelaskan pada awal tulisan ini, ada tiga jenis validitas yaitu *content validity* (validitas isi), *criterion validity* (validitas kriteria), dan *construct validity* (validitas konsep). Terkait dengan validitas isi, beberapa teknik statistik yang umum digunakan adalah Koefisien Validitas Isi – Aiken's V dan Koefisien Validitas Isi – Lawshe's CVR (lihat Hendryadi, 2017).

Validitas berdasarkan kriteria atau *criterion-related validity* merupakan sebuah ukuran validitas yang ditentukan dengan cara membandingkan skor-skor tes dengan kinerja tertentu pada sebuah ukuran luar. Ukuran luar ini seharusnya memiliki hubungan teoritis dengan variabel yang di ukur oleh tes itu. Misalnya, tes inteligensi mungkin berkorelasi dengan rata-rata nilai akademis. Dengan kata lain, pengujian ini dapat dilakukan dengan mengkorelasikan antara dua skala yang berbeda.

Untuk validitas konsep, kita dapat merujuk evaluasi pengukuran pada referensi buku *Struktural Equation Modeling*. Misalnya, validitas konstruk diperiksa berdasarkan pendekatan validitas konvergen dan validitas diskriminan. Validitas konvergen dinilai melalui loading factor

(λ) dan average variance extract (AVE) (Hair et al., 2014). Sebuah tes memiliki validitas konvergen jika memiliki korelasi yang tinggi dengan tes lain yang mengukur konstruk yang sama. Sebaliknya, validitas tes yang berbeda ditunjukkan melalui korelasi yang rendah dengan tes yang mengukur konstruk yang berbeda. Pendekatan *Multitrait-Multimethod* (MT-MM) merupakan pendekatan yang banyak dilakukan untuk menguji validitas konvergen dan validitas diskriminan.

Teknik lain yang sering digunakan untuk menguji validitas konsep adalah analisis factor yang dapat diaplikasikan dengan SPSS dan software sejenis lainnya. Analisis factor merupakan salah satu teknik yang paling populer untuk menguji validitas konsep dikarenakan ia dapat dipakai untuk mengidentifikasi dan mengakses kekuatan relative berbagai ciri-sifat psikologis yang berbeda. Analisis faktor juga dapat dipakai dalam merancang tes untuk mengidentifikasi faktor atau faktor-faktor primer melalui serangkaian test yang berbeda. Ada dua jenis analisis faktor yaitu *Exploratory Factor Analysis* (EFA) dan *Confirmatory Factor Analysis* (CFA). Dalam analisis faktor eksplorasi, seseorang peneliti ingin mengeksplorasi data empiris untuk menemukan dan mendeteksi karakteristik dan hubungan antar variabel tanpa menerapkan model yang pasti pada data. Analisis faktor eksplorasi dapat berupa pemodelan, atau pembuatan hipotesis. Dalam analisis faktor konfirmatori, di sisi lain, seseorang membangun sebuah model yang diasumsikan untuk menggambarkan, atau menjelaskan data empiris dengan parameter yang relatif sedikit (lihat Hair et al., 2010; 2012).

Terakhir, pendekatan baru yang dikembangkan oleh Hair et al. (2020) adalah *confirmatory component analysis* (CCA). Teknik ini mengusulkan pendekatan alternatif yang diterapkan untuk mengkonfirmasi model pengukuran ketika menggunakan model SEM-PLS, CCA juga berguna untuk mengembangkan ukuran baru, dan dapat berfungsi untuk model pengukuran eksplorasi dan konfirmasi (Hair et al., 2020).

KESIMPULAN

Teknik statistik untuk menguji validitas skor, terutama yang berasal dari kuesioner dapat menggunakan tiga pendekatan, yaitu pendekatan SEM, analisis faktor tunggal (EFA maupun CFA), Multitrait-Multimethod (MT-MM) untuk menguji validitas konvergen dan validitas diskriminan, dan *confirmatory component analysis* (CCA) yang menggunakan pendekatan SEM-PLS.

Referensi

- Akbulut, S., Aydinli, F. E., Kuşçu, O., Özcebe, E., Yilmaz, T., Rosen, C. A., & Gartner-Schmidt, J. (2020). Reliability and validity of the Turkish reflux symptom index. *Journal of Voice*, 34(6), 965-e23.
- Aygör, H. E., Fadiloğlu, Ç., Şahin, S., Aykar, F. Ş., & Akçiçek, F. (2018). Validation of edmonton frail scale into elderly turkish population. *Archives of gerontology and geriatrics*, 76, 133-137.
- Azwar, S. (2012). *Reliabilitas dan Validitas*. Ed. 4. Yogyakarta, Pustaka Pelajar
- Bicer, A., Yazici, A., Camdeviren, H., & Erdogan, C. (2004). Assessment of pain and disability in patients with chronic neck pain: reliability and construct validity of the Turkish version of the neck pain and disability scale. *Disability and rehabilitation*, 26(16), 959-962.
- Gary Growth – Marnat. (2010). *Handbook of Psychological Assessment*. Yogyakarta : Pustaka Pelajar
- Hair Jr, J. F., Howard, M. C., & Nitzl, C. (2020). Assessing measurement model quality in PLS-SEM using confirmatory composite analysis. *Journal of Business Research*, 109, 101-110.

- Hair, J. F., Sarstedt, M., Pieper, T. M., & Ringle, C. M. (2012). The use of partial least squares structural equation modeling in strategic management research: a review of past practices and recommendations for future applications. *Long range planning*, 45(5-6), 320-340.
- Hair, J.F., Black, W.C., Babin, B.J. and Anderson, R.E. (2010), *Multivariate Data Analysis*, 7th ed., Prentice Hall, Upper Saddle River, NJ
- Hendryadi, H. (2017). Validitas isi: tahap awal pengembangan kuesioner. *Jurnal Riset Manajemen dan Bisnis (JRMB) Fakultas Ekonomi UNIAT*, 2(2), 169-178.
- Hendryadi, H., Tricahyadinata, I., & Zannati, R. (2019). Metode Penelitian: Pedoman Penelitian Bisnis dan Akademik. *Jakarta: LPMP Imperium*.
- LoBiondo-Wood, G., & Haber, J. (2013). Methods and critical appraisal for evidence-based practice. *Nursing Research: Text and Study Guide Package*.
- Okamura, M., Fujimori, M., Hata, K., Mori, M., Mack, J. W., Prigerson, H. G., & Uchitomi, Y. (2021). Validity and reliability of the Japanese version of the Peace, Equanimity, and Acceptance in the Cancer Experience (PEACE) questionnaire. *Palliative & Supportive Care*, 1-7.
- Quadrelli, S., Davoudi, M., Galíndez, F., & Colt, H. G. (2009). Reliability of a 25-item low-stakes multiple-choice assessment of bronchoscopic knowledge. *Chest*, 135(2), 315-321.
- Şahin, G., Buzlu, S., Kuğuoğlu, S., & Yılmaz, S. (2020). Reliability and Validity of the Turkish Version of the Simulation Effectiveness Tool–Modified. *Florence Nightingale Journal of Nursing*, 28(3), 250.
- Schoppink, L. E., van Tulder, M. W., Koes, B. W., Beurskens, S. A., & de Bie, R. A. (1996). Reliability and validity of the Dutch adaptation of the Quebec Back Pain Disability Scale. *Physical therapy*, 76(3), 268-275.
- Westphal, T., Piatek, S., Schubert, S., Schuschke, T., & Winckler, S. (2002). Reliability and validity of the upper limb DASH questionnaire in patients with distal radius fractures. *Zeitschrift für Orthopädie und ihre Grenzgebiete*, 140(4), 447-451.
- Widhiarso, W. (2012). "(Sekali Lagi) Korelasi Aitem Total Bukanlah Validitas Item". <http://widhiarso.staff.ugm.ac.id/files/Korelasi%20Aitem%20Total%20Bukanlah%20Koefisien%20Validitas%20Aitem.pdf>
- Yudko, E., Lozhkina, O., & Fouts, A. (2007). A comprehensive review of the psychometric properties of the Drug Abuse Screening Test. *Journal of substance abuse treatment*, 32(2), 189-198.

Lihat Juga

- Allen, M., & Yen, W. (2002). *Introduction to measurement theory*. Long Grove, IL: Waveland Press, Inc.
- Frey, B. (2018). Classical Test Theory. *The SAGE Encyclopedia Of Educational Research, Measurement, And Evaluation*.
- Nunnally, J. C. (1994). *Psychometric theory 3E*. Tata McGraw-hill education.
- Polit, D.F., Beck, C.T., (2006). *Essentials of Nursing Research*, 6th ed. Lippincott Williams &Wilkins, Philadelphia.
- Traub, R. (2005). Classical Test Theory in Historical Perspective. *Educational Measurement: Issues And Practice*, 16(4), 8-14.
- Wright, B. (2005). A History of Social Science Measurement. *Educational Measurement: Issues And Practice*, 16(4), 33-45.