

(Sekali Lagi) Korelasi Aitem Total Bukanlah Validitas Item

Wahyu Widhiarso

Fakultas Psikologi UGM | 2012

Sering kali kita menjumpai sebuah hasil penelitian yang menggunakan kalimat seperti ini: *"dari nilai korelasi aitem total didapatkan bahwa validitas aitem bergerak antara 0,43 hingga 0,67"*. Dalam konteks ini validitas aitem ditunjukkan oleh korelasi aitem total (r-it). Saya mencoba melacak penggunaan r-it sebagai koefisien validitas, dan menemukan bahwa di tahun 1936, prosedur itu telah dipakai. Saya mendapatkannya dari tulisan Swineford (1936), berjudul *Validity of test items* dimuat dalam *Journal of Educational Psychology*. Kata validitas aitem ini relatif baru jika dibandingkan dengan validitas skor tes, karena skor tes merupakan satu kesatuan aitem-aitem di dalamnya, maka validitas tes lebih ditekankan pada tes secara keseluruhan, dalam hal ini diwakili oleh skor tes.

Meski Swineford menggunakan istilah validitas aitem, namun alenia awal tulisannya, dia buru-buru menegaskan bahwa validitas butir sejatinya ditunjukkan dengan korelasinya dengan tes lain yang ditetapkan sebagai kriteria yang telah valid. *"One view of the validity of a test item is its correlation with some valid criterion which is independent of the test"* (Swineford, 1936).

Selanjutnya, Swineford memperkenalkan bahwa r-it dapat dipakai untuk menunjukkan validitas aitem, akan tetapi syaratnya adalah tes tersebut secara keseluruhan sebelumnya telah dibuktikan valid. *"If the test as a whole, however, is known to be valid, then the rating of any item may be measured by its correlation with the test of which it is an element. Such a correlation may be thought of not only as a measure of the consistency between the item and the test total, but also as the validity of the item, since it shows the extent to which the item measures whatever function the test as a whole measures"* (Swineford, 1936). Dari pernyataan tersebut dapat diketahui bahwa r-it selain menjelaskan konsistensi aitem dan skor total tes, juga menunjukkan validitas aitem. Pernyataan tersebut tidak berhenti di situ, ketika muncul kata *"since it shows..."*, yang jika diartikan akan menjelaskan bahwa aitem tersebut sesuai dengan fungsi ukur keseluruhan tes atau skala.

Korelasi Aitem Total Sebagai Inisial dari Validitas Aitem

Fishman dan Galguera (2003) memang mengatakan bahwa korelasi aitem total dapat menunjukkan validitas. Dikatakan olehnya, *"initial expression of item validity can be the correlation of each item's response with total test score minus the the score for the item in question"*. Jika dibaca lebih lengkap di alenia sebelumnya, validitas menurut mereka lebih tepat direfleksikan oleh skor tes. Korelasi butir total hanyalah sebagai informasi pengantar saja mengenai informasi validitas skor tes.

Korelasi Aitem Total Sebagai Indeks Validitas Aitem

Urbina (2004) juga mensinyalir bahwa korelasi aitem total terkait dengan validitas aitem. Dikatakan dalam bukunya, *"Psychometricians usually refer to item validity statistics as indexes of item discrimination, because their role is to reveal the extent to which an item accurately differentiates among test takers with regard to the traits or behaviors that the test is designed to assess"*. Ditambahkan olehnya bahwa dalam proses pengembangan tes, diskriminasi aitem, yang juga dikenal sebagai indeks validitas aitem, dapat diperoleh dengan menggunakan beberapa kriteria atau indikator level individu pada konstruk yang diukur oleh tes. Kriteria yang digunakan untuk tujuan ini dapat berupa tiga jenis. (a) kriteria internal dengan pada tes yang sedang dikembangkan (misalnya, skor total tes), (b)

kriteria eksternal berupa tes yang mengukur hal yang sama dan telah tervalidasi, atau (c) kombinasi dari kriteria internal dan eksternal.

Indeks Validitas Aitem

Menurut Allen & Yen (2002), meskipun banyak metode telah diusulkan untuk melakukan analisis aitem, ada empat metode yang banyak digunakan: (1) analisis tingkat kesulitan aitem dan indeks diskriminasi aitem (korelasi aitem-total skor tes), (2) reliabilitas dan indeks validitas aitem (3) karakteristik kurva aitem dan (4) analisis faktor. Terlihat di sini bahwa korelasi aitem total merupakan bagian dari properti indeks diskriminasi aitem dan bukan indeks validitas aitem. Indeks validitas aitem memang terkait dengan korelasi aitem total, akan tetapi ada penambahan deviasi standar dalam penghitungannya. Indeks validitas aitem merupakan perkalian antara korelasi aitem total (r_{it}) dengan deviasi standar aitem (s_i).

Kesimpulan

Buku dasas psikometri seperti Nunnally & Bernstein (1994), Allen & Yen (2002), dan Crocker & Algina (1986) tidak menuliskan bahwa korelasi butir total adalah koefisien validitas. Juga buku-buku psikometri terbaru, misalnya Raykov (2011). Korelasi aitem total yang tinggi disebabkan oleh korelasi antar butir di dalam tes yang tinggi. Dan korelasi antar butir yang tinggi tidak berkaitan dengan validitas, melainkan konsistensi internal (alpha).

Kita bisa membuat tiga butir tes yang memiliki korelasi aitem tinggi, tetapi belum tentu valid. Tes yang kita kembangkan tersebut dikatakan valid ketika ada dasar eksternal yang dipakai untuk meninjau. Bisa dari teori yang dipakai untuk membuat butir (validitas isi) atau tes lain yang sudah valid (validitas kriteria). Kita juga bisa menggunakan internal tes sebagai dasar, misalnya dengan analisis struktur faktor (validitas konstruk).

Dalam penelitian, misalnya skripsi biasanya jenis validitas dilakukan dengan menggunakan validitas isi. Ketika butir yang dikembangkan sesuai dengan teori yang dipakai dan disetujui oleh pakar di bidang tersebut, maka hal itu sudah menunjukkan validitas alat ukur yang dikembangkan. Prosedur kemudian dilanjutkan dengan melakukan analisis butir dari sampel uji coba untuk memastikan butir yang ditulis mampu membedakan karakteristik sampel berdasarkan atribut yang diukur. Proses ini dilakukan dengan menghitung korelasi aitem-total sebagai koefisien yang menunjukkan daya diskriminasi aitem. Semakin banyak aitem di dalam alat ukur yang memiliki korelasi aitem-total yang tinggi, semakin tinggi reliabilitas pengukuran yang dilakukan. Jadi, korelasi aitem total tidak terkait dengan validitas, namun lebih terkait dengan reliabilitas.

Referensi

- Allen, M.J., & Yen, W.M. (2002). *Introduction to Measurement Theory*. Long Grove, IL: Waveland Press.
- Crocker, L., & Algina, J. (1986). *Introduction to Classical and Modern Test Theory*. New York: Harcourt Brace Jovanovich College Publishers.
- Fishman, J.A., & Galguera, T. (2003). *Introduction to Test Construction in the Social and Behavioral Sciences: A Practical Guide*. Lanham, MD: Rowman & Littlefield Publishers, Inc.
- Nunnally, J.C., & Bernstein, I.H. (1994). *Psychometric theory (3rd Edition)*. New York: McGraw Hill Inc.
- Swineford, F. (1936). Validity of test items. *Journal of Educational Psychology*, 27(1), 68-78.
- Urbina, S. (2004). *Essentials of psychological testing*. Hoboken, NJ.: John Wiley & Sons, Inc.