

## GENERALISASI VALIDITAS DALAM PENELITIAN KUANTITATIF

Validity Generalisation in Quantitative Research

Andi Ulfa Tenri Pada<sup>1,2</sup>

<sup>1</sup>Program Studi Doktor Penelitian dan Evaluasi Pendidikan UNY, Karangmalang Yogyakarta

<sup>2</sup>Program Studi Pendidikan Biologi FKIP Unsyiah, Darussalam Banda Aceh

e-mail : andiulfa.usk@gmail.com

### Abstrak

Tujuan utama dari penelitian generalisasi validitas adalah untuk mengestimasi mean dan variansi validitas sesungguhnya dari suatu tes seleksi. Prinsip dasar dari model generalisasi validitas adalah pemartisian variansi, yaitu identifikasi dan pemilahan terhadap variansi dari validitas teramati yang bertanggung jawab terhadap *error* sistematis yang berhubungan dengan artifak statistik dan *error* dari *random sampling*. Model dasar untuk generalisasi validitas didasarkan pada suatu struktur persamaan untuk koefisien korelasi teramati antara prediktor X dan kriteria Y, dimana R adalah korelasi teramati,  $\rho$  adalah korelasi populasi yang dibatasi, dan *e* adalah peluang *error* yang berhubungan dengan *sampling*.

**Kata kunci** : generalisasi validitas, tes, meta-analisis, *error* pengukuran.

### Abstract

The main objective of the validity generalization study is to estimate the mean and variance of the true validity of a selection test. The basic principle of the validity generalization model is partitioning variance, namely the identifying and sorting on the variance of observed validity that responsible towards systematic errors related to statistics artifacts and the error of random sampling. The basic model for validity generalization is based on an equation structure of observed correlation coefficients between X predictor and Y criterion, where R is observed correlation,  $\rho$  is the limited population correlation, and *e* is error probability related to sampling.

**Keywords**: validity generalization, test, meta-analysis, measurement error.

### PENDAHULUAN

Salah satu isu penting dalam prosedur penelitian pendidikan adalah sejauh mana bukti validitas yang diperoleh dalam suatu situasi dapat digeneralisasi untuk situasi lain tanpa studi lebih lanjut terhadap validitas dalam situasi yang baru tersebut. Salah satu fokus penelitian kuantitatif adalah pada ketetapan bahwa hasil penelitian pada suatu situasi tertentu dapat digeneralisasikan pada situasi yang lebih luas, artinya bahwa kesimpulan yang diperoleh dari hasil penelitian terhadap sampel atau situasi yang terbatas itu dapat ditarik kesimpulan umum yang berlaku bagi suatu populasi yang diwakili oleh sampel itu dalam situasi yang luas. Oleh karena itu persoalan yang pokok dihadapi oleh peneliti kuantitatif ialah persoalan representativitas sampel.

Keterbatasan utama studi validasi adalah dapat dengan mudah mengalami masalah metodologis yang tak terlihat. Dengan membandingkan validasi dan temuan di beberapa penelitian, maka memungkinkan untuk menentukan apakah validitas kriteria terkait tes relatif stabil atau tes ini hanya valid dalam situasi

tertentu. Dengan menggunakan teknik meta-analisis, prosedur perbandingan ini disebut generalisasi validitas dalam tes seleksi dan literatur psikometri.

### Teori Generalisasi Validitas

Para peneliti yang menggunakan rancangan eksperimen juga terkait dengan persoalan generalisasi. Bila penemuannya diberlakukan bagi kelompok diluar eksperimen, maka muncul persoalan validitas eksternal. Artinya sekalipun hasil dari suatu eksperimen dianggap memiliki validitas eksternal, namun bila ditarik lebih jauh untuk melihat tingkat generalisasinya, maka validitas eksternalnya juga diragukan. Rendahnya validitas eksternal itu disebabkan misalnya oleh pengaruh *pre-test* baik terhadap kelompok eksperimen maupun kelompok kontrol dan juga oleh pengaruh lainnya (misalnya pengalaman pribadi, persepsi, pedewasaan) yang dapat dipandang mempengaruhi hasil eksperimen. Setiap pengaruh yang masuk dalam eksperimen itu diberi sebutan *reactive effect*.

Persoalan validitas eksternal lebih mencolok lagi di dalam rancangan penelitian eksperimental

sosial didalam laboratorium. Kritik utama adalah bahwa penelitian laboratorium itu bersifat *artificial* (dibuat-buat) dan berlangsung tidak natural apa adanya. Kehidupan sosial sebenarnya tidak sama dengan kehidupan dalam laboratoriaum. Terlepas dari kelemahannya, Mook (1983) menyatakan bahwa penelitian manusia dengan menggunakan laboratorium bukanlah bertujuan untuk memperoleh validitas hasil penelitian tetapi untuk tingkah laku yang akan terjadi atau akibat yang diramalkan yang dapat terjadi.

Penelitian eksperimental dan survei telah menaruh perhatian yang besar pada persoalan generalisasi. Bukti yang berasal dari penelitian validitas terdahulu merupakan informasi yang berharga dalam memutuskan apakah suatu tes sesuai untuk digunakan untuk tujuan tertentu. Konsep ini telah dikenal di berbagai bidang dimana tes digunakan untuk tujuan seleksi, termasuk dalam bidang pendidikan, perekrutan tenaga kerja, maupun militer. Secara tradisional di berbagai bidang ini diyakini bahwa validitas suatu instrumen yang digunakan untuk seleksi mungkin situasinya bersifat spesifik. Sehingga diyakini bahwa suatu tes harus tetap divalidasi dimana pun ia akan digunakan. Pandangan ini diterima secara luas dalam bidang psikologi personalia. Namun pandangan yang mengatakan bahwa generalisasi tidak dapat dilakukan terhadap validitas suatu instrumen ini dibantah oleh beberapa peneliti seperti Guion (2006), yang menyatakan bahwa meskipun sulit untuk membuat inferensi bagi suatu aturan yang bersifat umum dari perilaku, tetapi masih sangat mungkin untuk membuat prediksi yang akan bermanfaat bagi kepentingan praktis. Pernyataan seperti ini juga disampaikan oleh banyak peneliti lainnya sebagai tantangan bagi para ahli psikologi personalia untuk mengembangkan suatu metode yang dapat digunakan untuk meneliti aturan umum beserta bukti-bukti yang mendukungnya.

Meta-analisis dari kebanyakan studi validitas yang dilakukan pada kemampuan kognitif umum juga telah berulang kali menunjukkan bahwa validitas kemampuan kognitif umum dalam memprediksi keberhasilan dalam pekerjaan tertentu sedikit berbeda dari satu kondisi ke kondisi yang lain (Schmidt & Hunter, 1981). Dengan demikian, ada bukti yang signifikan bahwa hasil validasi untuk tes kemampuan kognitif umum dapat digeneralisasikan pada seluruh kondisi. Sehingga tidak perlu untuk melakukan studi validitas untuk memprediksi keberhasilan dalam pekerjaan tertentu pada setiap kondisi. Validitas kemampuan kognitif umum untuk memprediksi kinerja peserta didik dalam satu kondisi, misalnya, dapat disimpulkan dari validitas yang ditemukan pada ratusan penelitian sebelumnya.

Keterbatasan lain dari studi validasi lokal yang spesifik adalah akurasi statistik yang dihasilkan (Schmidt, Hunter & Urry, 1976). Statistik yang akurat membutuhkan ukuran sampel yang besar. Kriteria terkait validitas tes dalam studi validasi lokal biasanya disimpulkan hanya jika besarnya temuan mencapai tingkat tertentu disebut yang disebut signifikansi statistik. Semakin kecil ukuran sampel, maka semakin tinggi koefisien validitas teramati yang diperlukan untuk menyimpulkan tingkat validitas yang dapat diterima. Contohnya, untuk memperoleh prediksi yang akurat terhadap pemilihan presiden tidak mungkin akan diperoleh dengan polling sampel hanya 15 pemilih. Sebagian besar jajak pendapat mewawancarai 1.000 pemilih atau lebih. Ada sampling error besar dalam studi validasi individu yang dilakukan dengan sampel kecil, sehingga data tidak dapat digunakan untuk menarik kesimpulan yang akurat.

### **Pendekatan Terhadap Generalisasi Validitas**

Penelitian konvensional tentang generalisasi validitas bermula ketika Schmidt & Hunter (1977) meneliti power statistik dari penelitian-penelitian validitas terpilih dan mempertimbangkan suatu prosedur untuk menguji hipotesis mengenai kekhususan yang bersifat situasional. Prosedur baru yang digunakan adalah teknik meta analisis untuk mengestimasi variansi kesalahan melalui sejumlah hasil penelitian. Dengan menggunakan kumpulan penelitian terdahulu serta menggunakan teknik meta analisis, temuan bisa didapatkan dengan cara membuatnya memiliki manfaat praktis yang besar (Glass, McGaw & Smith, 1981).

Tujuan utama dari penelitian generalisasi validitas adalah untuk mengestimasi mean dan variansi validitas sesungguhnya dari suatu tes seleksi (Suen, H.K. & Kim, J.H., 2006). Pengembangan metodologi serta argumen pendukungnya telah banyak dilakukan di area ini. Argumen-argumen tersebut bertujuan agar metode dilakukan dengan pendekatan tertentu yang lebih rinci dibandingkan dengan konsep-konsep dasar yang hanya dicantumkan dalam tinjauan pustaka.

Pendekatan yang berasal Schmidt & Hunter (1977) menjelaskan tentang tujuh sumber variansi yang potensial menyebabkan variasi pada koefisien validitas teramati, namun hanya empat yang diestimasi. Pendekatan ini menggunakan teknik meta analisis dan metode quasi-bayesian untuk mengestimasi mean dan variansi dari distribusi validitas sesungguhnya, serta untuk mendapatkan suatu interval yang dapat dipercaya mengenai validitas sesungguhnya. Sebagai hasil tambahan dari analisis ini, persentase dari variansi teramati yang dihitung dari sumber-sumber artifisial juga turut diestimasi.

Pendekatan Schmidt & Hunter digambarkan sebagai quasi-bayesian dikarenakan memerlukan informasi mengenai efek ketidakreliabelan, pembatasan jangkauan, dan faktor-faktor lainnya yang kesemuanya dijelaskan dalam suatu istilah yang disebut dengan distribusi awal atau yang diasumsikan. Ada tiga distribusi yang diperlukan bagi kebanyakan prosedur yaitu reliabilitas tes, reliabilitas kriteria, dan pembatasan jangkauan antar penelitian validitas yang digunakan. Distribusi-distribusi ini jarang tersedia dalam laporan-laporan baik yang dipublikasikan maupun tidak, sehingga beberapa distribusi hipotetik kemungkinan tidak sama dengan distribusi yang sesungguhnya dalam dunia nyata dan para pengguna metode generalisasi validitas harus menerima berbagai perbedaan yang ada sebagai sumber bias dalam prosedur yang digunakan.

Ukuran perbedaan ini menentukan akurasi dari berbagai implementasi praktis dari prosedur-prosedur generalisasi validitas. Tanpa menjelaskan secara rinci macam-macam metode generalisasi validitas, dapat dikatakan bahwa hasil analisis generalisasi validitas mungkin tergantung pada kebenaran dari distribusi hipotetik yang digunakan, metode khusus yang digunakan (seperti penjumlahan atau perkalian), dan data dalam bentuk penelitian validitas yang tersedia untuk dianalisis.

Dibutuhkan banyak usaha, waktu, dan dana untuk menunjukkan kriteria-terkait validitas menggunakan pendekatan konkuren atau prediktif. Sehingga lebih mudah menggunakan pendekatan generalisasi validitas. Sebelumnya diyakini bahwa validitas uji terbatas pada situasi tertentu, hal ini disebut kekhususan situasional. John Hunter dan Frank Schmidt (1990) melakukan penelitian untuk menguji keyakinan tersebut. Teknik generalisasi validitas dilakukan baik secara statistik dan konseptual terkait dengan teknik meta-analisis yang dikembangkan bersama. Bahkan, ide tentang meta-analisis dimulai sebagai masalah yang berkaitan dengan generalisasi validitas. Kesimpulannya adalah bahwa validitas cenderung menggeneralisasi seluruh situasi, dan ada banyak penelitian yang mendukungnya (Schmidt et al., 1993). Bukti bahwa validitas dapat menggeneralisasi seluruh situasi ini sangat meyakinkan, tetapi hal ini tidak sederhana, namun banyak kritik dan kekhawatiran telah dikemukakan terhadap metodologi statistik yang digunakan dalam generalisasi validitas.

### Prosedur Dasar Generalisasi Validitas

Prinsip dasar dari model generalisasi validitas adalah pemartisian variansi, yaitu identifikasi dan pemilahan terhadap variansi dari validitas teramati yang bertanggung jawab terhadap *error* sistematis yang berhubungan dengan artifak

statistik dan *error* dari *random sampling*. Secara khusus Pearlman, Schmidt, Hunter, Linn & Dunbar menjelaskan dalam literatur generalisasi validitas bahwa efek sistematis ini dapat dibaca pada bahasan-bahasan tentang perbedaan antara penelitian-penelitian dalam hal reliabilitas tes, kriteria reliabilitas, hal batasan interval, jumlah dan jenis dari kriteria kontaminasi dan defisiensi, struktur faktor dari tes yang konstruk yang sama, dan perbedaan dalam penelitian-penelitian dalam hal struktur faktor dari kriteria pengukuran (Schmidt et al., 1993). Dengan demikian variansi yang tidak terkait dengan ketujuh faktor ini akan mencerminkan variabilitas dari validitas yang sesungguhnya. Model dasar untuk generalisasi validitas didasarkan pada suatu struktur persamaan untuk koefisien korelasi teramati antara prediktor X dan kriteria Y yang dapat dituliskan sebagai berikut:

$$R = \rho' + e$$

Dimana R adalah korelasi teramati,  $\rho'$  adalah korelasi populasi yang dibatasi, dan *e* adalah peluang *error* yang berhubungan dengan *sampling*. Jika dalam hal ini ukuran sampelnya tak terhingga, maka korelasinya akan sebesar  $\rho'$ . Namun dengan ukuran sampel yang aktual (tidak tak terhingga) maka ada perbedaan antar korelasi teramati (R) yang dihitung dari data. Korelasi yang dibatasi dihitung dari data. Korelasi yang dibatasi berbeda dari korelasi populasi yang berhubungan dengan ketidakreliabelan prediktor, ketidakreliabelan kriteria, dan pembatasan jangkauan. Persamaan struktural yang lengkap yang digunakan di hampir keseluruhan penelitian generalisasi validitas dapat ditulis sebagai berikut:

$$r = abc\rho' + e$$

Dimana:

- $\rho'$  = korelasi populasi
- r* = korelasi teramati
- a* =  $(r_{yy'})^{1/2}$  = akar kuadrat dari reliabilitas kriteria
- b* =  $(r_{xx'})^{1/2}$  = akar kuadrat dari reliabilitas kriteria
- c* =  $(u/(1+(u-1)^2 a^2 b^2))^{1/2}$ , adalah faktor pembatasan jangkauan
- e* = kesalahan *sampling*
- u* = rasio dari SD terbatas terhadap SD tak terbatas dari X.

### SIMPULAN

Ada dua manfaat utama studi generalisasi validitas. Pertama, hasil penelitian generalisasi dapat berfungsi untuk menarik kesimpulan ilmiah tentang hubungan antara variabel. Salah satu contoh generalisasi validitas yang baik adalah kesimpulan yang dibuat oleh Hunter dan Schmidt (1981) bahwa tes kemampuan kognitif yang paling

sering digunakan adalah valid untuk semua jenis pekerjaan. Selain itu mereka juga menyimpulkan bahwa bahwa validitas tes kognitif yang diteliti adalah tidak spesifik untuk suatu situasi atau spesifik untuk pekerjaan tertentu. Temuan ini dapat meningkatkan pemahaman kita bahwa test/kriteria memungkinkan aplikasi yang lebih berguna sebagai prediktor skor.

Kedua, bukti validitas terkait kriteria yang diperoleh dari penelitian sebelumnya dapat digunakan untuk mendukung penggunaan tes

dalam situasi yang baru. Aplikasi teori generalisasi validitas ini memiliki potensi besar bagi pendidik dan pengusaha yang tidak memiliki ukuran sampel yang cukup atau sumber daya dalam suatu organisasi, namun ingin menerapkan program pengujian yang terbukti valid. Proses transfer tes dari satu situasi di mana tes telah terbukti valid dengan situasi atau lokasi lain yang serupa, sering disebut sebagai “transportabilitas validitas” dari satu situasi ke situasi lainnya.

#### DAFTAR PUSTAKA

- Glass, G. V., McGaw, B., dan Smith, M. L. 1981. *Meta-analysis in social research*. Sage Library of Social Research. Sage Publications, Michigan University.
- Guion, R. M., dan Highhouse, S. 2006. *Essentials of Personnel Assessment and Selection*. Mahwah, NJ, Lawrence Erlbaum Associates, Inc.
- Hunter, John E., dan Schmidt, Frank L. 1990. *Methods of Meta-Analysis: Correcting Error and Bias in Research Findings*. Newbury Park, CA: Sage, 592.
- Linn, R. L., dan Dunbar, S. B. 1985. Validity generalization and predictive bias. In R. A. Berk (Ed.), *Performance assessment: Methods & applications*. Baltimore, MD, The Johns Hopkins University Press.
- McDaniel, M. A. 2006. Validity Generalization as a Test Validation Approach. <http://www.people.vcu.edu/~mamdani/Publications/McPhail.c05.pdf>.
- Mook, Douglas G. 1983. In Defense of External Invalidity. *American Psychologist*, April: 379-387.
- Schmidt, FL, dan Hunter, JE. 1977. Development of a general solution to the problem of validity generalization. *Journal of Applied Psychology*.
- Schmidt, F.L., dan Hunter, J.E. 1981. Employment testing: Old theories and new research findings. *American Psychologist*, 36:1128-1137.
- Schmidt, F. L., Law, K., Hunter, J. E., Rothstein, H. R., Pearlman, K., dan McDaniel, M. 1993. Refinements in validity generalization methods: Implications for the situational specificity hypothesis. *Journal of Applied Psychology*, 78:3-12.
- Schmidt, F.L., Hunter, J.E., dan Urry, V.W. 1976. Statistical power in criterion-related validation studies. *Journal of Applied Psychology*, 61:473-485.
- Suen, H.K. ( ), dan Kim, J.H. ( ). 2006. (Validity generalization: Accumulation of criterion-related evidence). (*Educational Measurement*), 3(2):3-14. <http://suen.ed.psu.edu/~hsuen/papers.html>