

Hasil Uji Statistik dan Penulisan Butir yang Kurang Tepat

Ketika mendapat pertanyaan apa penyebab hasil uji statistik yang dilakukan tidak signifikan, saya seringkali merasa *lemes* dan *speechless*. Entah itu hasil uji korelasi yang tidak signifikan atau hasil uji model (SEM) yang signifikan. Pertanyaan itu yang bisa menjawab adalah peneliti itu sendiri, karena dia yang menguasai dasar teori, alat ukur yang dipakai dan karakteristik data penelitian.

Jawaban standar saya biasanya adalah silahkan membaca tulisan ini^[4]. Namun beberapa kali saya mendapati bahwa meski beberapa saran di tulisan itu sudah diikuti, hasilnya tetap tidak seperti harapan peneliti. Kalau itu yang terjadi, maka kesimpulannya adalah memang variabel yang dikorelasikan benar-benar tidak memiliki hubungan yang signifikan.

Kalau memang tidak signifikan, mengapa dipaksakan signifikan? Mungkin teori yang kita pakai kurang kuat dan belum banyak diuji dalam penelitian. Kalau teorinya kuat, mungkin teori tersebut tidak berlaku pada sampel kita. Misalnya ada kajian teoritis yang mengatakan bahwa evaluasi kerja secara eksplisit akan meningkatkan performansi kerja. Penelitian di barat menunjukkan kedua variabel itu memiliki signifikan, belum teori ini berlaku di Indonesia. Orang Indonesia kalau diberi evaluasi secara eksplisit kemungkinan akan merasa down, karena hal itu seperti halnya aib kita dibuka pada khalayak.

Namun saya berkesimpulan bahwa para peneliti telah memahami hal ini, dan mereka sangat selektif dalam memilih teori yang kira-kira jika diterapkan pada sampel orang Indonesia akan berlaku. So, tidak ada kesalahan dalam menggunakan teori dan merasionalisasi hubungan antar variabel sebelum keduanya dihipotesiskan. Teori yang dipakai lebih banyak bersifat umum sehingga tidak ada masalah dalam hal pemilihan teori.

Investigasi saya berlanjut pada alat ukur yang dipakai oleh peneliti. Hasilnya, kebanyakan cara peneliti dalam mengoperasionalkan variabel (atribut) menjadi perilaku yang teramati (item di dalam skala), kurang tepat. Reliabilitas yang tinggi tidak menjamin skala yang dipakai valid, karena reliabilitas yang tinggi bisa jadi muncul dari skala yang kurang valid. Korelasi aitem-total yang tinggi yang dihitung melalui SPSS juga bukan koefisien validitas. Korelasi aitem-total menunjukkan bahwa butir-butir di dalam skala memiliki kesamaan domain ukur, namun memiliki kesamaan dalam mengukur apa tidak dijelaskan oleh nilai korelasi itu. Hal inilah yang menyebabkan saya beberapa kali mengingatkan bahwa korelasi aitem-total bukan koefisien validitas.

Kembali lagi ke masalah kekurangakuratan dalam mengoperasionalkan konsep teoritik menjadi indikator perilaku. Ada sebuah kasus penelitian yang menemukan bahwa hubungan antara harga diri dan ekstraversi tidak signifikan. Sekilas, memang agak janggal karena orang yang mempersepsi dirinya memiliki sejumlah kelebihan cenderung lebih terbuka, proaktif dan bebas dalam mengekspresikan diri. Prof. Sofia Retnowati pernah mengajari saya untuk melihat apakah dua variabel itu memiliki keterkaitan, bisa dilihat dari butir-butir kedua variabel. Dalam kasus hubungan antara harga diri dan ekstraversi ini kita bisa melihat dari dua sampel butir dari kedua variabel tersebut.

Ilustrasi

Misalnya harga diri diwakili oleh butir "*Saya memiliki sejumlah kelebihan yang saya banggakan*". Sebaliknya butir ekstraversi diwakili oleh butir "*Saya tidak malu mengeluarkan pendapat saya pada orang lain*". Kedua butir menyediakan opsi respons dari *sangat tidak setuju* hingga *sangat setuju*. Hampir dipastikan, secara logis, subjek yang menjawab setuju pada pada sampel butir harga diri, juga akan menjawab setuju pada sampel butir ekstraversi di atas. Walhasil, semakin banyak subjek yang sama-sama setuju pada kedua butir, maka korelasi antar skor yang dihasilkan kedua variabel akan signifikan.

Ketika mahasiswa datang kepada saya dan menanyakan mengapa hasil uji statistik yang mereka lakukan tidak signifikan. Pertama kali saya menyarankan untuk mengecek data (kemungkinan salah penyekoran), membuang outliers dsb. Kedua, ketika saya pandang teori yang dipakai untuk menyusun hipotesis cukup kuat dan logis, saya meminta mahasiswa menunjukkan seperti butir-butir dalam skala mereka. Saya ambil beberapa sampel aitem dari kedua variabel yang diukur, lalu saya mengisinya (membayangkan saya adalah subjek penelitian mereka). Dari proses ini saya bisa menemukan penyebab hasil uji (i.e korelasi) yang tidak signifikan. Kebanyakan yang saya temui adalah hasil uji tidak signifikan disebabkan oleh mahasiswa kurang cermat dalam menjabarkan konsep teoritik menjadi butir pernyataan.

Kasus I

Sebuah penelitian menguji hubungan antara tingkat kemandirian individu dengan konformitas terhadap tekanan kelompok. Secara logis dapat kita ketahui bahwa orang yang mandiri memiliki sifat bebas dan tidak terpengaruh oleh intervensi dari luar, dengan demikian mereka cenderung tidak terpengaruh oleh tekanan, alias memiliki konformitas terhadap tekanan yang rendah. Jika dinalar, kedua variabel ini akan memiliki korelasi yang signifikan. Tapi ternyata hasil uji statistik mengatakan sebaliknya. Jika kita lihat pada sampel butir kemandirian, yaitu "*Saya bebas menentukan apa tujuan saya*" dan butir konformitas, "*Teman-teman saya melihat bahwa saya tidak mudah dipengaruhi*". Jika saya dikenai dua pernyataan tersebut, untuk butir pertama saya memilih opsi *sangat setuju*, sedangkan pada pernyataan kedua saya memilih opsi *netral*. Teman-teman melihat saya tidak dapat dipengaruhi, bisa jadi berbeda dengan apa yang saya rasakan. So, wajar saja hasil uji tidak signifikan, karena pernyataan menyangkut diri saya sedangkan yang satunya pandangan orang lain tentang saya.

Kasus II

Contoh lain, hubungan antara kualitas pelayanan dengan kepuasan pelanggan. Secara logis bisa diterima, semakin berkualitas pelayanan yang diberikan maka pelanggan kita akan semakin puas. Tapi butir kualitas pelayanan diwujudkan dalam butir berikut, "*Sebaiknya keluhan pelanggan diatasi secepatnya*", sedangkan butir kepuasan diwujudkan dalam, "*Secara umum saya puas dengan pelayanan di sini*". Hampir dipastikan, setiap subjek akan memilih opsi setuju pada butir pertama, akan tetapi belum tentu mereka juga setuju dengan butir yang kedua. So, korelasi yang tidak signifikan tidak terelakkan. Butir pertama yang mengukur kualitas pelayanan merupakan opini normatif yang hampir semua orang menyetujuinya, sebaliknya belum tentu mereka setuju juga pada butir kedua.

Kasus III

Sebuah penelitian menemukan hubungan yang tidak signifikan antara motivasi berprestasi dan sikap terhadap perubahan organisasi. Contoh butir pengukuran motivasi adalah "*Saya siap menghadapi segala jenis tantangan*" sedangkan contoh butir pengukuran sikap adalah "*Saya mendukung setiap perubahan organisasi di tempat saya bekerja*". Silahkan anda membayangkan menjadi subjek penelitian yang dikenai dua butir tersebut. Apakah korelasi antara motivasi berprestasi dan sikap terhadap perubahan organisasi akan tinggi (signifikan). Ya, karena orang yang menyukai tantangan akan cenderung menerima perubahan.

Rekomendasi

Kasus di atas menunjukkan bahwa operasionalisasi konsep teoritik yang menjadi butir yang kurang tepat bertanggung jawab terhadap hasil uji yang tidak signifikan. Beberapa rekomendasi yang dapat saya berikan adalah: (1) buatlah butir yang benar-benar merepresentasikan atribut yang diukur secara konsisten. (2) Jangan terlalu jauh dalam memilih indikator perilaku, tetapkan fokus pada indikator yang telah dijelaskan oleh teori. (3) Sebelum mengambil data, anda bisa melakukan simulasi dengan menjadikan anda sebagai subjek penelitian untuk memperkirakan gambaran data yang didapatkan. (4) Lakukan proses validasi alat ukur yang lebih ketat, misalnya melibatkan panel ahli (lebih dari satu) ketika menguji validitas isi. Reliabilitas dan korelasi aitem-total yang tinggi bukan jaminan alat ukur kita valid.

Catatan

Tulisan ini dibuat dengan tujuan untuk menunjukkan pentingnya masalah penulisan butir dalam menghasilkan alat ukur yang valid. Ketika alat ukur tidak valid, maka hipotesis penelitian yang kita ajukan tidak akan didukung oleh data, meski dibangun berdasarkan teori yang mapan. Proses melakukan simulasi dengan membayangkan diri kita sebagai subjek penelitian kemudian mengisi butir pada alat ukur penelitian, yang ditulis di sini, hanyalah sekedar ilustrasi saja. Proses ini bukan proses standar yang harus dilakukan dalam penelitian. Cuma sekedar menunjukkan apa penyebab hasil uji hipotesis ditolak. Jangan lantas kemudian ketika kita menulis aitem pada skala yang mengukur atribut ukur yang berbeda kemudian dihubung-hubungkan, subjek merespons secara konsisten pada kedua alat ukur, dengan tujuan agar hasil uji statistik yang dihasilkan menunjukkan hubungan yang signifikan.

Jena, 2012

Wahyu Widhiarso