

Corelații, corelații, corelații ... perspectiva SEM¹ asupra măsurării în psihologie

Balázsi Róbert²

Universitatea Babeș-Bolyai, Cluj Napoca

Măsurarea a reprezentat și reprezintă în continuare o problemă esențială atât a cercetării în psihologie, cât și a psihologiei aplicate, astfel construcția unor instrumente de măsură, respectiv teste, standardizate reprezintă un imperativ al ambelor domenii. Calitatea măsurării (implicit calitatea testelor) se reflectă în indicatorii de validitate și fidelitate a instrumentelor de măsură. Există și alte caracteristici importante ale unui instrument de măsură, oferite de exemplu de analiza de itemi, dar în discuția din acest număr ne vom limita doar la validitate și fidelitate pentru că ele se bazează în general pe valoarea indicilor de corelație calculată între două sau mai multe variabile observate (sau proceduri mai complexe bazate pe analiza corelațională).

Întrebarea la care încercăm să oferim un răspuns este, în ce măsură interpretarea efectuată în cadrul conceptual tradițional al unor indici de corelație între două variabile observate pot oferi, sau nu, informațiile necesare în acest sens. Ceea ce ne face să ne îndoim de eficiența practicilor efectuate în spiritul psihometriei clasice este perspectiva, relativ nouă oferită de metodologia SEM. Nu ne propunem să punem sub semnul întrebării psihometria clasică în întregimea ei (în acest sens vezi asumptiile Teoriei Răspunsului la Item, Embretson & Hershberger, 1999), dar vom discuta două aspecte care să arate clar cum modelarea prin ecuații structurale, care și ea se bazează pe analiza corelațiilor sau a covarianțelor, nuanțează problematica validității și fidelității.

SEM reprezintă o modalitate de analiză structurală a covarianțelor, având ca scop verificarea ipotezei conform căreia matricea de covarianțe a populației variabilelor observate este o funcție a parametrilor liberi necunoscuți ai modelului teoretic structurat (Breckler, 1990; Joreskog & Sorbom, 1984).

Modelul teoretic structurat practic specifică relații între diferite variabile măsurate și variabile latente.

În general modelul SEM presupune două componente: componenta de măsurare și componenta structurală. Componenta de măsurare specifică relațiile existente între variabilele latente și indicatorii acestora (variabilele măsurate), în timp ce componenta structurală specifică relațiile existente între variabilele latente (Anderson & Gerbing, 1988).

De asemenea, se face o distincție între variabile exogene și variabile endogene. Variabilele exogene reprezintă acele variabile spre care nu indică nici o săgeată unidirecționată în graficul modelului testat. Variabilele endogene sunt cele spre care sunt orientate săgeți unidirecționate. În esență săgețile unidirecționate în model sugerează faptul că variabilele exogene determină variabilele endogene care la rândul lor pot, sau nu, determina alte variabile endogene (Cohen & Cohen, 1983).

Primul model, cel de măsurare este de obicei specificat prin următoarele două ecuații:

$$x = \Lambda \xi + \delta$$

x reprezentând vectorul variabilelor exogene măsurate, Λ fiind matricea coeficienților ce indică influența variabilelor latente exogene ξ asupra variabilelor măsurate. δ reprezintă matricea erorilor de măsurare a variabilelor exogene măsurate. A doua ecuație:

$$y = \Lambda \eta + \epsilon$$

surprinde aceleași relații în cazul variabilelor endogene. y reprezintă matricea variabilelor măsurate endogene, Λ matricea coeficienților ce indică influența variabilelor latente endogene, notate cu η , asupra variabilelor endogene măsurate.

$$\eta = \beta \eta + \Gamma \xi + \zeta$$

unde β și Γ reprezintă influența variabilelor latente endogene (notate cu η) și exogene

¹ SEM (Structural Equation Modeling) se traduce Modelare prin Ecuații Structurale.

² Adresa de corespondență: robertbalazsi@psychology.ro

(notate cu ξ) asupra variabilei latente endogene și ζ marchează prezența altor variabile externe care afectează variabila latentă endogenă (Bollen, 1989).

Demersul SEM constă în structurarea unui model, stabilirea parametrilor liberi și setați al acestuia urmând ca pe baza matricei de covarianță a datelor și a matricii ipotetice elaborate pe baza modelului structurat să se calculeze valorile parametrilor modelului și indicii de potrivire (Bentler & Bonett, 1980). Să vedem în continuare cum acest demers schimbă perspective noastră asupra demersului de validare relativă la construct și a demersului de evaluare a consistenței interne ca indice al fidelității testului.

Conform definiției oferite de Anastasi (1976) validitatea reprezintă măsura în care un test măsoară într-adevăr ceea ce își propune să măsoare. Aceeași sursă bibliografică oferă o descriere detaliată a principalelor modalități de validare a unui instrument de măsură. Printre cele propuse un rol important îl are validarea relativă la construct. O procedură, larg acceptată, de validare relativă la construct o reprezintă cea propusă de Campbell și Fisk (1959), validarea convergentă și discriminantă. Cel mai frecvent aceasta presupune derularea unui studiu MTMM³ ce presupune măsurarea a mai multor variabile latente cu metode diferite de evaluare (la nivel general am putea spune că se măsoară mai multe constructe prin scale sau itemi diferiți). Ulterior în matricea de corelație, dacă testele implicate au validitate convergentă și discriminantă, se observă corelații mai mari între diferitele măsurări ale aceluiași construct (între diferitele metode aplicate aceleiași variabile latente) față de corelațiile înregistrate între aceeași metodă aplicată diferitelor trăsături (între aceeași metodă aplicată diferitelor variabile latente).

Criteriile de interpretare a matricii oferite de Campbell și Fiske (1959) au fost intens combătute într-o serie de articole (Althaus & Heberlein, 1970; Alwin, 1974; Widaman, 1985). În principal problema majoră o reprezintă faptul că valoarea calculată a unui coeficient de corelație este afectată în mare măsură de fidelitatea scalei. Or, aceste erori de măsurare, respectiv impactul acestora asupra concluziilor studiului, nu pot fi detectate printr-o singură inspecție a matricii de corelații.

O altă problemă semnalată de Bollen și Lennox (1991) este că analiza doar a

coeficienților de corelație nu ține cont de relația existentă între diferitele modalități de operaționalizare ale aceluiași construct, respectiv de relațiile existente între variabilele latente implicate în studiu. Ca și criteriu general este acceptat că diferitele măsurători al aceluiași construct trebuie să coreleze mai mult decât valorile pentru măsurătorile prin aceeași metodă a două constructe diferite.

În termeni formali putem afirma că,

$$y_1 = \lambda_{11}\hat{\eta}_1 + \epsilon_1$$

$$y_2 = \lambda_{21}\hat{\eta}_1 + \epsilon_2$$

scorul la o variabilă măsurată a unui construct (ex. un item al unei scale de inteligență) este determinată de puterea relației dintre item și construct ($\lambda_{11}\hat{\eta}_1$ și $\lambda_{21}\hat{\eta}_1$) și eroarea de măsurare asociată aceluia item (ϵ_1 și ϵ_2). Conform psihometriei clasice media erorilor de măsurare este zero și acestea nu corelează cu nivelul de abilitate ($\hat{\eta}_1$ și $\hat{\eta}_2$) fiind independente și între ele (Nunnally, 1978). Aceeași formulă poate fi scrisă și pentru un al doilea construct ($\hat{\eta}_2$).

$$y_3 = \lambda_{32}\hat{\eta}_2 + \epsilon_3$$

$$y_4 = \lambda_{42}\hat{\eta}_2 + \epsilon_4$$

În acest context corelația dintre y_1 și y_2 este determinată de produsul dintre $\lambda_{11}\lambda_{21}$; acest produs va determina valoarea corelațiilor ce indică validitatea convergentă. În schimb corelația dintre trăsăturile aparținând de două variabile latente diferite (y_2 și y_3) este determinată de produsul $\lambda_{21}\lambda_{32}\rho_{21}$, adică depinde de puterea relației fiecărui item cu constructul care îl determină dar și de puterea relației dintre cele două constructe (ρ_{21}).

În aceste condiții prescripția inițială a psihometriei clasice, corelații mari între diferite metode de măsurare ale aceluiași construct și corelații mici între valorile rezultate în urma aplicării aceleiași metode de măsurare la constructe diferite, este discutabilă. Dacă λ_{11} este mai mic decât $\lambda_{32}\rho_{21}$ atunci inevitabil corelația dintre diferitele măsurători ale aceluiași construct va fi mai mică decât corelația dintre două măsurători, cu aceeași metodă, ale unor constructe diferite.

Să presupunem că între cele două variabile latente $\hat{\eta}_1$ și $\hat{\eta}_2$ există o corelație redusă $\rho_{21} = 0.1$ și că încărcarea pe factor a itemilor este $\lambda_{11} = 0.6$, $\lambda_{21} = 0.6$, respectiv $\lambda_{32} = 0.6$, $\lambda_{42} = 0.6$. În acest caz va fi adevărată afirmația, $r_{12} > r_{23}$. Însă dacă valoarea lui

³ MTMM (Multi Trait Multi Method) se traduce Metoda Trăsăturilor Multiple și Metodelor Multiple.

crește (de ex. $\rho_{21} = 0.5$) și încărcarea diferitelor măsurători pe factori diferă foarte mult $\lambda_{11} = 0.3$, $\lambda_{21} = 0.3$ și $\lambda_{32} = 0.8$, $\lambda_{42} = 0.8$ afirmația $r_{12} > r_{23}$ nu va mai fi adevărată. Aceasta nu se datorează faptului că testul nu are validitate convergentă, ci se datorează legăturii existente între cele două constructe.

Este clar că exemplul prezentat operează cu valori extreme, valorile din cercetările reale situându-se undeva între aceste două extreme. Ceea ce este important de reținut este că asumția psihometriei clasice funcționează perfect doar în cazul în care variabilele latente sunt necorelate (Bollen, 1984). Referindu-ne la constructele vizate prin cercetările psihologice ne este greu de acceptăm această asumție, deoarece în psihologie există o oarecare corelație între marea majoritate a variabilelor (puține variabile pot fi considerate perfect ortogonale). Eroarea pe care o riscă un cercetător limitându-se doar la analiza vizuală a unei astfel de matrici de corelație depinde de fidelitatea măsurătorilor și de valoarea corelației existente între variabilele latente. Din păcate niciuna dintre aceste două componente nu sunt direct prezente în matricea de corelație.

Fidelitatea reprezintă gradul de consistență al unui scor (Anastasi, 1976). Dacă azi măsurăm inteligența cuiva, ne așteptăm ca la o reluare a măsurării să obținem un scor similar, asta în condițiile în care trăsătura evaluată este una stabilă. Există mai multe modalități de estimare a fidelității, printre care cel mai frecvent utilizat este indicele de consistență internă (Cronbach & Meehl, 1955).

Bollen și Lennox (1991) arată că premisele procedurii de calcul și modul de interpretare a indicelui de consistență internă, reflectă prescripțiile psihometriei clasice și, ca urmare, acordă itemilor doar statutul de variabile efect. Adică singura posibilitate luată în considerare este ca variabila latentă să determine variabilele măsurate. În acest sens, există o serie de studii care arată că, în anumite situații variabilele măsurate determină nivelul variabilei latente. Pentru a oferi un exemplu pentru ambele situații, să ne închipuim două variabile latente: inteligența și predispoziția spre discriminare (Bollen & Lennox, 1991). În primul caz capacitatea subiectului de a rezolva anumite probleme de un anumit grad de dificultate se datorează cantității variabilei latente. Cu cât o persoană este mai inteligentă cu atât este mai probabil că va putea rezolva și itemii cu un grad ridicat de dificultate. Aceasta este imaginea generală a psihometriei clasice asupra relației existente

între construct și diferitele modalități de operare a acestuia. În schimb variabilele care determină predispoziția spre discriminare, rasă, vârstă, gen și disabilități nu sunt determinate de variabila latentă. Din contră, acestea sunt cele care determină probabilitatea de a fi victima unui act discriminator.

Analizând exemplul de mai sus nu ne este greu de acceptat că indicele de fidelitate are sens doar în cazul în care toate variabilele măsurate (itemii testului) reprezintă indicatori efect al constructului investigat. Dacă în schimb unele dintre ele reprezintă indicatori cauzali ai constructului, în cazul lor nu ne putem aștepta să existe corelații prea ridicate, așa cum nici între variabilele constructului predispoziție spre discriminare nu ne așteptăm să existe corelații.

În cercetările reale este foarte probabil să apară și variabile măsurate care reprezintă efectul constructului (de ex. singurătatea ca măsură a depresiei) dar intervin și ca indicatori cauzali (de ex. izolarea ca și cauză a depresiei). În cazul acestor variabile latente modelul de măsurare ce urmează a fi supus confirmării trebuie astfel configurat încât să reflecte și acest aspect al fenomenului (Bollen, 1984).

Concluzii

În acest studiu ne-am propus să oferim o imagine foarte sumară asupra metodologiei SEM și asupra modului în care această metodologie schimbă (sau trebuie să schimbe) perspectiva noastră asupra procesului de măsurare în psihologie. Marele avantaj al metodologiei SEM este că impune specificarea și testarea unui model explicit de măsurare. În psihometria clasică în majoritatea cazurilor acest aspect este doar unul asumat, fără a fi verificat. Din păcate utilizarea doar a coeficienților de corelație nu ne oferă informații asupra relației existente între variabilele latente (aspect ce influențează indicii validității de construct) și nici asupra naturii relației dintre indicatori și variabilă latentă (ce va afecta interpretarea indicelui de consistență internă). În ambele cazuri specificarea unui model de măsurare și verificarea acestuia în cadrul metodologiei SEM va duce la îmbunătățirea măsurării și la creșterea validității concluziilor cercetării.

Bibliografie

- Anastasi, A. (1976). *Psychological testing*. New York: MacMillan Publishing.
- Alwin, D.F. (1974). Approaches to the interpretation of relationships in the multitrait-multimethod matrix. In H. L. Costner(Ed.), *Sociological methodology 1973-1974* (pp. 79-105). San Francisco: Jossey-Bass.
- Althauser, R.P., & Heberlein, T.A. (1970). A causal assessment of validity and the multitrait-multimethod matrix. In E. Borgatta (Ed.), *Sociological methods* (pp. 151-169). San Francisco: Jossey-Bass.
- Anderson, J.C., & Gebing, D.W. (1988). Structural Equation Modeling in Practice: A Review and Recommended Two-Step Approach. *Psychological Bulletin*, 103, 3, 411-423.
- Bentler, P.M., & Bonett, D.G. (1980). Significance tests and goodness-of-fit in the analysis of covariance structures. *Psychological Bulletin*, 88, 588-606.
- Bollen, K.A. (1984). Multiple indicators: Internal consistency or no necessary relationship? *Quality and Quantity*, 18, 377-385.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K.A. & Lennox, R. (1991). Conventional Wisdom on Measurement: A Structural Equation Perspective. *Psychological Bulletin*, 110, 2, 305-314.
- Breckler, S.J. (1990). Applications of Covariance Structure Modeling in Psychology: Cause for Concern? *Psychological Bulletin*, 107, 2, 260-273.
- Campbell, D.T., & Fiske, D.W. (1959). Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Cohen, J. & Cohen, P. (1983). *Applied multiple regression/correlation analysis for the behavioral sciences*. London: Lawrence Erlbaum Associates.
- Cronbach, L.J. & Meehl, P.E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, 52, 281-302.
- Embretson, S.S., & Hershberger, S.L. (1999). *The New Rules of Measurement What Every Psychologist and Educator Should Know*. London: Lawrence Erlbaum Associates, Publishers.
- Joreskog, K.G., & Sorbom, D. (1984). *LISREL vi: Analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.
- Nunnally, J.C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.
- Widaman, K.F. (1985). Hierarchically nested covariance structure models for multitrait-multimethod data. *Applied Psychological Measurement*, 9, 1-26.



Organizational Diagnosis & Development
 Brand Research
 Academic & Institutional Evaluation
 Organizational Simulation & Gaming

Focus on people and opportunities.
 For reliable outcomes
contact@aphorme.ro; www.aphorme.ro