

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Paul Sârbescu¹

Abstract

In Romania, the use of exploratory factor analysis (EFA) in the development or cultural adaptation of psychological tests is a phenomenon that gained a very large amplitude in recent years. Many of the recommendations that apply when using EFA for variables will lead to erroneous results if EFA is used for items. The purpose of this article is to provide relevant information about issues to be taken into account when EFA is used for items. Basically, this article presents a series of recommendations about decisions on determining the need to implement EFA, the sample used, choosing a factor extraction method, selecting the optimal number of factors, factor rotation and identifying false factors. Finally, a synthesis of recommendations made throughout the article is presented, along with the main techniques for forming item parcels and a guide for conducting a parallel analysis using SPSS syntax.

Keywords: Exploratory factor analysis, item parcels, parallel analysis

Résumé

En Roumanie, l'utilisation de l'analyse factorielle exploratoire (EFA) dans le processus de développement ou d'adaptation culturelle des épreuves psychologiques représente un phénomène qui a pris un grand essor pendant les dernières années. Beaucoup de recommandations valables pour l'utilisation de la EFA pour les variables vont mener à des résultats erronés dans le cas de l'utilisation de la EFA pour les items. Le but de cet article c'est de fournir des informations significatives sur les aspects qui doivent être pris en considération lors qu'il s'agit d'utiliser la EFA pour les items. Pratiquement, le présent article vient illustrer une série de recommandations liées aux décisions concernant la détermination du besoin d'application de la EFA, l'échantillon utilisé, le choix d'une méthode d'extraction des facteurs et l'identification des faux facteurs. A la fin, il y a une synthèse des recommandations mentionnées tout au long de l'article, en présentant les principales techniques pour la formation des parcelles d'items aussi bien qu'un guide pour la réalisation d'une analyse parallèle dans la syntaxe du logiciel SPSS.

Mots-clés: l'analyse factorielle exploratoire, parcelles d'items, analyse parallèle

Rezumat

Utilizarea analizei factoriale exploratorii (eng. „exploratory factor analysis” sau EFA) în procesul de dezvoltare sau de adaptare culturală a probelor psihologice reprezintă un fenomen ce a căpătat o foarte mare amploare în ultimii ani, în România. Multe dintre recomandările care sunt valabile pentru utilizarea EFA în cazul variabilelor vor duce la rezultate eronate în cazul utilizării EFA pentru itemi. Scopul acestui articol este de a oferi informații relevante despre aspectele care trebuie luate în calcul atunci când EFA este utilizată în cazul itemilor. Practic, acest articol prezintă o serie de recomandări legate de deciziile privind determinarea necesității de a aplica EFA, eșantionul utilizat, alegerea unei metode de extracție a factorilor, selectarea numărului optim de factori, rotirea factorilor și identificarea factorilor falși. La final este realizată o sinteză a recomandărilor făcute pe par-

¹ Universitatea de Vest din Timișoara.
Adresă de corespondență: paul.sarbescu@gmail.com

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

cursul articolului, sunt prezentate principalele tehnici pentru formarea pachetelor de itemi, cât și un ghid pentru realizarea unei analize paralele din sintaxa programului SPSS.

Cuvinte-cheie: Analiza factorială exploratorie, pachete de itemi, analiza paralelă

Introducere

Analiza factorială exploratorie (EFA) este una dintre cele mai utilizate tehnici statistice din științele sociale. Scopul principal al acesteia este identificarea numărului și a naturii factorilor care stau la baza unui set de variabile manifeste (Little, Cunningham, Shahar și Widaman, 2002; Bandalos, 2008; Sava, 2011).

Cu toate că tehnica a fost dezvoltată acum mai bine de un secol și a fost aplicată constant în diverse scopuri (dezvoltarea unor teorii ale personalității, construirea unor instrumente psihometrice), mulți autori au criticat modalitatea de aplicare a EFA (Armstrong, 1967; Gould, 1981). De asemenea, unele revizuri ale cercetărilor publicate în literatura de specialitate ce folosesc EFA au scos la iveală carențe evidente la nivelul deciziilor luate de autori în derularea acestei tehnici statistice (Ford, MacCallum și Tait, 1986; Fabrigar, Wegener, MacCallum, și Strahan, 1999).

În cazul utilizării analizei factoriale exploratorii pentru dezvoltarea/adaptarea culturală a unor instrumente psihometrice, există mai multe aspecte ce trebuie luate în considerare, pentru a evita obținerea unor rezultate distorsionate. Utilizarea EFA în cazul itemilor în același fel ca în cazul variabilelor poate conduce, destul de des, la obținerea unor structuri factoriale inadecvate. Acest lucru se datorează, în principal, particularităților specifice ale itemilor (Gorsuch, 1997).

Articolul de față cuprinde o trecere în revistă a aspectelor specifice ale folosirii EFA în cazul itemilor, urmată de o analiză a deciziilor ce trebuie luate în cazul utilizării EFA cu itemi. Acest decizii sunt legate de: determinarea necesității de a aplica EFA, eșantionul utilizat, alegerea unei metode de extracție a factorilor, selectarea numărului optim de factori, rotirea acestora și identificarea factorilor falși. La final sunt prezentate câteva informații legate de formarea pachetelor de itemi, iar încheierea este formată dintr-un set de concluzii și recomandări.

Aspecte specifice ale folosirii analizei factoriale exploratorii în cazul itemilor

Utilizarea EFA în cazul itemilor trebuie făcută ținând cont de faptul că itemii au proprietăți diferite față de scale, și necesită astfel tehnici specifice, pentru evitarea obținerii unor rezultate eronate. Prezentăm în continuare principalele diferențe dintre itemi și scale, ce pot influența rezultatele unei analize factoriale exploratorii (Gorsuch, 1997):

a) Itemii au o fidelitate mai scăzută decât scalele; combinarea unui set de itemi cu fidelitate scăzută reduce eroarea itemilor prin combinarea varianței comune. Astfel, este mult mai dificil de lucrat cu corelațiile dintre itemi, deoarece acestea sunt mai scăzute decât corelațiile dintre scale. Din această cauză, factorii sunt mai instabili în cazul utilizării EFA cu itemi decât în cazul utilizării EFA cu scale.

b) Itemii conțin deseori varianță străină constructului care este măsurat. Aproape orice item din orice scală conține o frază sau un cuvânt care poate fi interpretat de cineva într-o manieră idiosincronică. Scalele diminuează aceste erori minore prin însumarea confuziilor aferente itemilor, astfel normalizându-le.

c) Distribuțiile itemilor diferă frecvent unele de altele, de multe ori în mod intenționat. Pentru a putea surprinde toate variațiile unui construct este necesară prezența unor itemi cu medii care variază de-a lungul scalei de răspuns. Itemii cu mediile situate la extreme au, în general, distribuții asimetrice; aceste distribuții influențează corelațiile dintre itemi deoarece o corelație poate fi puternică și pozitivă doar dacă itemii care sunt corelați au aceeași distribuție. Astfel, prezența diferitelor distribuții reduce valoarea corelațiilor dintre itemi. Scalele sunt formate prin însumarea unor itemi, normalizând astfel distribuțiile; deci, scalele prezintă în general corelații mai ridicate decât itemii.

d) Răspunsurile la itemi sunt aproape tot timpul formate dintr-o serie de categorii (dihotomice sau scale de tip Likert cu 5-7 variante), și

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

foarte rar dintr-o scală de valori continuă. Din cauza faptului că valorile nu au o amplitudine mai mare (ex. de la 0 la 50), răspunsurile la un item pot reflecta doar diferențe majore între indivizi, privind constructul măsurat (ex. chiar dacă 40 de persoane obțin valoarea 4 la un anumit item, vor exista anumite diferențe între ei, privind constructul măsurat, iar aceste diferențe ar putea fi identificate de un alt item). Deoarece scorurile scalelor reprezintă suma itemilor componenți, scalele funcționează mai degrabă ca o variabilă continuă, având astfel o mai bună posibilitate de diferențiere între indivizi.

Din cauza acestor diferențe, realizarea unei analize factoriale exploratorii pe un set de itemi în același fel ca pe un set de variabile va duce, foarte des, la obținerea unor rezultate distorsionate. Sugestiile din literatura de specialitate pentru folosirea EFA în cazul itemilor (mai ales atunci când distribuțiile itemilor deviază substanțial de la normalitate) sunt următoarele (Bernstein și Teng, 1989; Panter, Swygart, Dahlstrom și Tanaka, 1997; O'Connor, 2000): pentru datele dihotomice, este recomandată efectuarea analizei factoriale folosind matricea corelațiilor de tip tetrachoric dintre itemi, în locul matricii de corelații Pearson. Pentru itemii ce folosesc scale de tip Likert, este recomandată efectuarea analizei factoriale folosind matricea corelațiilor de tip polychoric dintre itemi, în locul matricii de corelații Pearson.

Corelațiile de tip tetrachoric și polychoric se bazează pe presupunerea că toate categoriile de răspuns (dihotomice sau scală de tip Likert) sunt de fapt indicatori pentru variabile latente, care au o distribuție normală. Analizele factoriale a matricelor de corelație de tip tetrachoric sau polychoric sunt, în esență, analize factoriale a relațiilor dintre variabilele care stau la baza datelor brute, și care se presupune că sunt continue și au o distribuție normală (Panter și colab., 1997). Din păcate, SPSS-ul nu oferă posibilitatea realizării EFA asupra unor astfel de matrici de corelație. În schimb, alte programe precum FACTOR* (Lorenzo-Seva și Ferrando, 2006) oferă posibili-

tatea realizării EFA pe matrici de corelație de tip tetrachoric și polychoric.

O altă recomandare pentru situația în care avem de-a face cu date dihotomice (dar nu numai) este formarea pachetelor de itemi. Prin intermediul acestei metode se grupează și se analizează mai mulți itemi împreună, renunțându-se astfel la utilizarea itemului în sine ca unitate de măsură (Bandalos, 2008; Sava, 2011). Mai multe informații despre situațiile când este recomandată formarea pachetelor de itemi, cât și despre principalele tehnici de formare a acestora, vor fi oferite către finalul acestui articol.

Principalele decizii în cazul folosirii analizei factoriale exploratorii

Determinarea necesității de a aplica EFA

Înainte de a trece efectiv la aplicarea EFA, orice cercetător ar trebui să fie sigur că această tehnică statistică este cea necesară pentru atingerea obiectivelor studiului (Fabrigar și colab., 1999). Astfel, în cazul dezvoltării unui nou instrument psihometric, utilizarea EFA este necesară pentru a identifica unidimensionalitatea sau multidimensionalitatea setului de itemi supus analizei, și implicit structura factorială a probei. De asemenea, în cazul adaptării culturale a unui test, EFA poate fi utilizată la primele aplicări ale probei, pentru a verifica în ce măsură structura factorială originală a probei este reprodusă pe populație românească; de asemenea, ulterior aplicării EFA este necesară și aplicarea analizei factoriale confirmatorii (eng. „confirmatory factor analysis” sau CFA), pentru confirmarea modelului factorial al probei respective. În schimb, în cazul în care o probă a fost adaptată cu succes și utilizată intensiv, folosirea EFA nu mai are foarte mare relevanță, în aceste situații fiind mult mai indicată utilizarea CFA, pentru a verifica invarianța instrumentului respectiv (Sava, 2011).

* Programul și instrucțiunile de utilizare sunt disponibile gratuit pe internet la adresa: <http://psico.fcep.urv.es/utilitats/factor/index.html>.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Eșantionul utilizat

Pentru selectarea unui eșantion adecvat trebuie avute în vedere două aspecte: caracteristicile și volumul său (Gorsuch, 1997). Legat de caracteristicile eșantionului, soluția optimă ar consta în selectarea unui lot de omogenitate medie, deoarece eșantioanele foarte omogene sau foarte eterogene pot conduce la obținerea unor soluții factoriale incorecte. De asemenea, foarte importantă este și similaritatea dintre lotul utilizat și populația țintă pe care va fi aplicată proba (ex. dacă testul va fi aplicat pe femei de peste 50 de ani, atunci lotul utilizat pentru dezvoltarea/adaptarea culturală a probei ar trebui să fie format din femei de peste 50 de ani; dacă testul urmează să fie aplicat elevilor de liceu, atunci lotul utilizat ar trebui să fie format din elevi de liceu). Din păcate, în majoritatea studiilor de acest gen sunt utilizate eșantioane formate din studenți, în special cei de la științe sociale, probabil din cauza gradului mare de accesibilitate (ITC, 2010). Eșantioanele astfel construite au în general anumite caracteristici (vârste între 19 și 25 de ani, preponderența genului feminin etc.), care nu sunt întotdeauna adecvate pentru dezvoltarea/adaptarea oricărui tip de instrument psihometric.

Legat de volumul eșantionului, trebuie menționat faptul că orice regulă fixă nu este valabilă și nici foarte folositoare (Gorsuch, 1997; MacCallum, Widaman, Zhang și Hong, 1999). Atât numărul minim de participanți, cât și raportul minim participanți-itemi necesare pentru a asigura o analiză optimă variază în funcție de anumite caracteristici ale itemilor incluși în studiu; mai exact, acestea variază în principal în funcție de nivelul comunalității itemilor, care se află într-o relație invers proporțională cu volumul eșantionului. În cazul utilizării EFA cu variabile (care au de obicei un nivel al comunalității mai ridicat decât itemii), soluții factoriale optime au fost identificate pe eșantioane mai mici de 100 de persoane, având un raport participanți-variabile de numai 3:1. În cazul utilizării EFA cu itemi, sunt în general necesare eșantioane mai mari (minim 100 de participanți, preferabil peste 200), și un raport participanți-itemi de cel puțin 5:1 (MacCallum și colab., 1999; de Winter, Dodou și Wieringa, 2009; Sava, 2011). Evident, după cum menționam

anterior, aceste indicații nu trebuie interpretate ca niște reguli rigide, peste care nu se poate trece sub nici o formă (ex. structura factorială a unui test de 12 itemi ce măsoară 2 fațete ale aceluiași factor, are mari șanse să fie reprodusă cu succes chiar și pe un eșantion de 70-80 de participanți; de asemenea, structura factorială a unui test multidimensional de 110 itemi este foarte probabil să fie identificată pe un lot de 400 de participanți, adică la un raport participanți-itemi mai mic de 4:1)

Alegerea unei metode de extracție a factorilor

Una dintre cele mai importante diferențe care trebuie menționată atunci când discutăm despre metodele de extracție a factorilor este aceea dintre analiza componentelor principale (eng. „principal component analysis” sau PCA) și analiza factorilor comuni (AFC; eng. „common factor analysis”). O primă diferență dintre cele două se observă la nivelul modului de analiză a datelor. AFC identifică trei surse ale variației datelor: factorii comuni, factorii specifici și erorile de măsurare, gruparea ultimelor două formând varianța unică. În opoziție, PCA nu diferențiază între varianța comună și varianța unică, dispersia datelor având la bază două surse: dispersia explicată și cea datorată erorilor de măsurare. De asemenea, teoria de la baza PCA nu specifică existența unor factori comuni, folosind ca substitut termenul de componentă principală (Fabrigar și colab., 1999; Sava, 2011).

O altă diferență dintre cele două metode este identificată la nivelul scopului fiecăreia. Dacă obiectivul AFC este de a explica corelațiile dintre o serie de variabile manifeste prin determinarea unui număr restrâns de factori care să le influențeze, scopul PCA este de explica cât mai mult din totalul dispersiei variabilelor manifeste prin identificarea unui număr restrâns de componente principale (Hogarty, 2005; Sava, 2011).

Astfel, având în vedere faptul că obiectivul PCA nu este acela de a identifica factori comuni care să influențeze o serie de variabile manifeste, iar componentele principale nu sunt variabile latente, putem afirma faptul că PCA nu este o metodă de analiză factorială și, în consecință, nu ar trebui utilizată atunci când dorim

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

realizarea unei analize factoriale exploratorii* (Fabrigar și colab., 1999).

Cele mai utilizate metode de extragere specifice EFA sunt analiza factorilor principali (eng. „principal axis factoring”), factorizarea Alpha (eng. „Alpha factoring”) și analiza factorială bazată pe estimarea verosimilității maxime (eng. „maximum likelihood factor analysis” sau ML FA). Prima prezintă avantajul de a nu fi foarte pretențioasă în ceea ce privește prezența unei distribuții multivariate normale a datelor; cea de-a doua este orientată pe maximizarea fidelității (α Cronbach), iar cea de-a treia, deși sensibilă la forma distribuției datelor, prezintă marele avantaj de a testa gradul de semnificație statistică a saturației factorilor și corelației dintre factorii extrași (Fabrigar și colab., 1999, Sava, 2011). Din păcate, din multitudinea de indicatori descriptivi pe care metoda îi poate oferi, SPSS-ul oferă doar testul de estimare a raportului verosimilității (eng. „likelihood ratio test” sau LRT), bazat pe o distribuție χ^2 , foarte sensibil la volumul eșantionului, care tinde să supraestimeze numărul de factori care trebuie extrași (Hayashi, Bentler și Yuan, 2007). În schimb, programul CEFA** (Browne, Cudeck, Tateneni și Mels, 2008) oferă toți indicatorii aferenți acestei metode, fiind astfel indicat în cazul în care dorim să folosim metoda ML FA.

Selectarea numărului optim de factori

Probabil cea mai cunoscută și utilizată metodă pentru selectarea numărului de factori este criteriul lui Kaiser. Metoda presupune selectarea unui număr de factori comuni egal cu numărul de *eigenvalue* mai mari decât 1. Chiar dacă metoda funcționează acceptabil atunci când EFA este folosită în cazul variabilelor, în cazul utilizării EFA pentru itemi metoda va produce în majoritatea situațiilor soluții eronate, supraestimând numărul de factori care trebuie selectați (Fabrigar și colab., 1999; Henson, 2006). Astfel, folosirea acestei metode atunci

când utilizăm EFA pentru dezvoltarea/adaptarea culturală a unor probe este contraindicată.

O altă metodă la fel de cunoscută este criteriul grafic al lui Cattell. În această procedură, valorile *eigenvalue* calculate sunt reprezentate grafic în ordine descrescătoare, iar graficul este inspectat pentru a se identifica ultima scădere substanțială a intensității *eigenvalue* (așa-numitul „punct de cotitură”), după care se selectează un număr de factori comuni egal cu numărul de *eigenvalue* situați deasupra punctului de cotitură (Fabrigar și colab., 1999). Metoda funcționează bine în situațiile în care factorii comuni prezenți sunt foarte bine reprezentați. Dezavantajele ei apar atunci când punctul de cotitură nu este foarte evident, sau în momentul în care există mai multe scăderi bruște ale intensității *eigenvalue*, deoarece nu există o explicație obiectivă a ceea ce înseamnă o „scădere substanțială” în intensitate (Sava, 2011).

Una dintre cele mai precise metode pentru selectarea numărului optim de factori este metoda analizei paralele (Horn, 1965). Aceasta se bazează pe compararea intensității *eigenvalue* dintre datele reale și un set de date aleatoare, urmată de selectarea unui număr de factori comuni egal cu numărul valorilor *eigenvalue* reale care sunt mai mari decât valorile *eigenvalue* aleatoare (Hayton, Allen și Scarpello, 2004). Chiar dacă metoda oferă rezultate valide în majoritatea situațiilor, există și anumite situații particulare (ex. număr foarte mare de itemi, multe comunalități foarte scăzute: $< .30$) când analiza paralelă tinde să supraestimeze ușor numărul de factori care trebuie selectați (Buja și Eyuboglu, 1992). Un dezavantaj major legat de accesibilitatea analizei paralele este reprezentat de faptul că metoda nu se poate realiza din programul SPSS fără apelul la sintaxă (Anexa 1 prezintă un ghid simplificat pentru realizarea unei analize paralele folosind sintaxa SPSS-ului).

O metodă care câștigă din ce în ce mai mult teren în fața abordărilor clasice este selectarea

* Programul SPSS nu permite realizarea PCA în varianta clasică; sub denumirea de „principal component method” programul realizează o analiză factorială a componentelor principale. Diferența dintre cele două este minimă, opțiunea din SPSS introducând o sursă suplimentară de eroare, lucru ce o face și mai puțin dezirabilă.

** Programul împreună cu manualul de utilizare sunt disponibile gratuit pe internet la adresa: <http://faculty.psy.ohio-state.edu/browne/software.php>.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

numărului de factori în funcție de indicatorii descriptivi oferiți de analiza factorială bazată pe estimarea verosimilității maxime (ML FA). Această metodă oferă indicatori prezenți și în analiza factorială confirmatorie. Cel mai potrivit indicator absolut este RMSEA (eng. „root mean squared error of aproximation”), care reprezintă o estimare a discrepanței dintre modelul testat și datele folosite, pe gradele de libertate ale modelului (Sass, 2009). În plus, este oferit și un interval de încredere pentru valoarea calculată. Interpretarea sa se face în felul următor: dacă limita maximă a intervalului nu depășește 0.08 avem un model bun, dacă nu depășește 0.10 avem un model acceptabil, iar dacă depășește 0.10 avem un model deficitar (Fabrigar și colab., 1999). Un alt indicator este ECVI (eng. „expected cross validation index”), care estimează gradul de generalizare al soluției obținute la alte eșantioane. Deși nu există valori absolute pentru interpretarea ECVI, acesta este util în compararea modelelor alternative, valorile mai mici arătând un grad de generalizare mai ridicat. Un indicator considerat problematic și, în consecință, contraindicat pentru selectarea numărului de factori, este testul de estimare a raportului verosimilității (LRT), din cauza supraestimării numărului de factori care trebuie selectați (Hayashi, Bentler și Yuan, 2007).

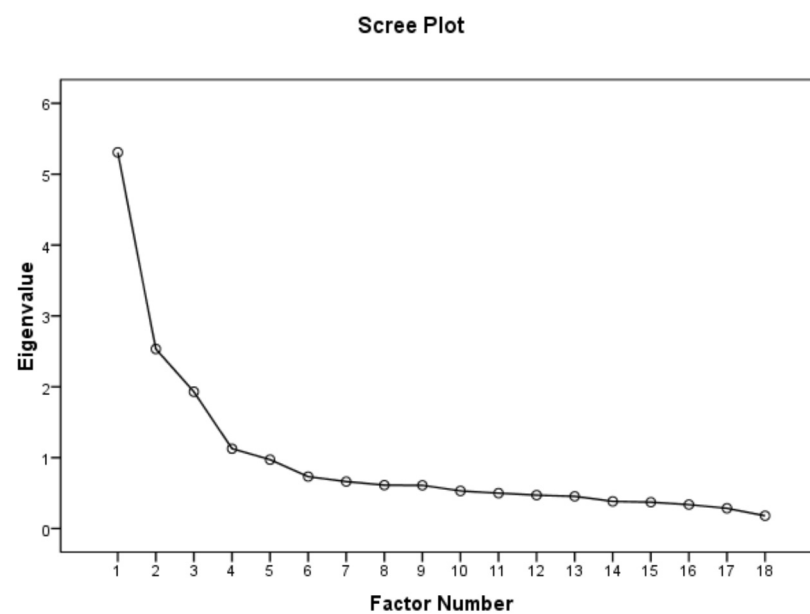
Prezentăm în continuare o situație practică în care criteriul lui Kaiser supraestimează

numărul de factori ce trebuie extrași, în timp ce criteriul grafic al lui Cattell și analiza paralelă sugerează soluția optimă. Exemplul este extras din procesul de validare al *Scalei de îndeplinire a nevoilor la locul de muncă (Need Satisfaction at Work Scale – NSWS)*; Van den Broeck, Vansteenkiste, De Witte, Lens & Soenens, 2010) în România. Autorii scalei au identificat o structură factorială formată din trei factori: *nevoia de autonomie, de competență și de relaționare*.

Primele 5 valori *eigenvalue* extrase sunt: 5.31, 2.53, 1.93, 1.27 și 0.97. *Scree plot*-ul obținut este prezentat în figura 1, iar rezultatele analizei paralele sunt prezentate grafic în figura 2.

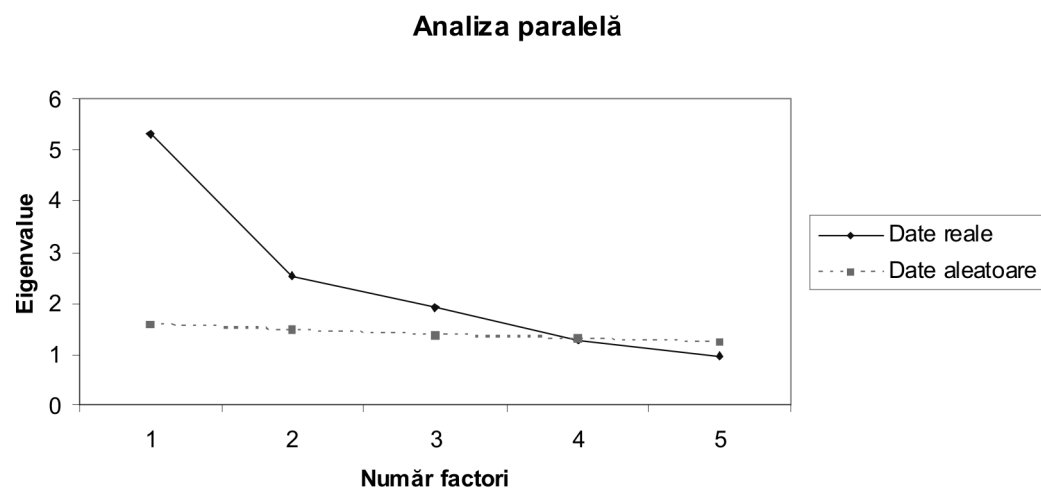
După cum se poate observa în cazul *scree plot*-ului, „punctul de cotitură” este situat după cel de-al treilea factor. De asemenea, analiza paralelă arată că primele trei valori *eigenvalue* reale sunt mai mari decât valorile *eigenvalue* aleatoare. Astfel, ambele metode sugerează o soluție cu trei factori. În schimb, criteriul lui Kaiser sugerează o soluție cu patru factori (supraestimând numărul de factori care trebuie extrași), deoarece există patru valori *eigenvalue* mai mari decât 1. Un ultim aspect ce trebuie discutat se referă la situațiile când este selectat un număr greșit de factori. În literatura de specialitate există un consens cu privire la acest aspect: subestimarea numărului de factori este o eroare mult mai gravă decât supraestimarea acestora (Fava și Velicer, 1992; Wood, Tataryn și Gor-

Figura 1. *Scree plot*-ul pentru NSWS



Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Figura 2. Rezultatele analizei paralele pentru NSWS



such, 1995). Când se produce subestimarea, factorii selectați vor conține un nivel de eroare considerabil, deoarece itemii care ar fi trebuit să încarce factorii care nu au fost extrași pot încărca incorect factorii extrași. De asemenea, nivelul de saturație al itemilor care ar încărca în mod normal factorii extrași poate fi afectat. Astfel, nivelul de eroare regăsit în model crește direct proporțional cu numărul de factori subestimați. În schimb, când se produce supraestimarea, nivelul de saturație pentru factorii autentici conține, în general, mai puțină eroare decât în cazul subestimării. De asemenea, supraestimarea cu doi, trei sau chiar patru factori induce aproximativ la fel de multă eroare în model ca și supraestimarea cu un singur factor (Wood și colab., 1995).

Rotirea factorilor

Pentru a ajunge la o structură factorială cât mai adecvată trebuie utilizat criteriul structurii simple definit de Thurstone. Acest lucru înseamnă că din multitudinea de rotații posibile, cea care va avea cea mai bună structură simplă va avea o semnificație psihologică reală, va fi cea mai ușor de interpretat și de replicat. Practic, structura simplă se referă la o soluție în care fiecare item obține saturații ridicate într-un singur factor și saturații reduse în ceilalți factori (Fabrigar și colab., 1999).

Pentru obținerea structurii simple au fost dezvoltate o serie de rotiri ale factorilor. Cea

mai importantă distincție între acestea le împarte în rotiri ortogonale și rotiri oblice (Browne, 2001). Rotirile ortogonale constrâng factorii să rămână necorelați, pe când rotirile oblice permit factorilor să coreleze între ei.

Din categoria rotirilor ortogonale, rotirea Varimax este de departe cea mai populară și cea mai des utilizată. Cu toate că în foarte multe situații această rotire a produs rezultate adecvate, folosirea ei (și a oricărei rotiri de tip ortogonal) în situația utilizării EFA în cazul itemilor nu este indicată (Gorsuch, 1997). Principalul motiv care conduce la această recomandare este legat de faptul că, în general, factorii măsurati de un anumit instrument psihometric tind să coreleze între ei. Chiar și în situațiile în care factorii au o anumită independență, tot există corelații de magnitudini scăzute (până la .30) între aceștia. De aceea, este mult mai indicată utilizarea rotirilor de tip oblic pentru a ajunge la cea mai simplă structură factorială.

Spre deosebire de rotirile ortogonale, nu există nici o rotire oblică care să domine în cercetările psihologice. Cele mai utilizate rotiri oblice sunt: Direct Quartimin, Promax și rotirea orto-oblică Harris-Kaiser (Fabrigar și colab., 1999; Browne, 2001). Rotirile Promax și Harris-Kaiser se regăsesc în SPSS în mod direct, iar rotirea Direct Oblimin (din SPSS) cu valoarea delta prestabilită la zero corespunde unei rotiri Direct Quartimin. Atât programul CEFA cât și programul FACTOR oferă o gamă mult mai largă de rotiri.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Mai trebuie menționat faptul că utilizarea unei rotiri oblice nu necesită ca factorii să fie corelați. Dacă soluția cu cea mai bună structură simplă implică factori ortogonali, o rotire oblică va oferi valori ale corelațiilor dintre factori apropiate de zero și va produce o soluție factorială apropiată de cea a unei rotiri ortogonale (Fabrigar și colab., 1999). Deci, în cazul utilizării EFA pentru dezvoltarea/adaptarea culturală a unei probe este mult mai indicată folosirea unei rotiri de tip oblic, decât a uneia de tip ortogonal.

Identificarea factorilor falși

În situația în care după efectuarea rotirii apar factori falși, aceștia trebuie identificați și respinși. În prezent, nu există reguli clare pentru identificarea acestor factori falși; totuși, prezentăm patru indicii care v-ar putea ajuta în luarea unei decizii legate de autenticitatea factorilor extrași (Wood și colab., 1995):

a) gradul de saturație al itemilor – este mult mai probabil ca factorii în care itemii sunt puternic saturați să fie autentici, decât factorii ce prezintă saturații scăzute ale itemilor;

b) valorile corelațiilor dintre itemii care încarcă într-un factor – este foarte probabil ca un factor să fie autentic dacă itemii care îl încarcă sunt corelați între ei;

c) semnificația factorului – dacă itemii care sunt puternic saturați într-un factor prezintă asemănări evidente între ei, atunci este mult mai probabil ca factorul să fie autentic, decât dacă similaritățile între itemii respectivi sunt aproape inexistente;

d) replicabilitatea factorului – dacă un factor apare în mai multe analize factoriale realizate pe eșantioane diferite, este destul de probabil ca acesta să fie autentic. Factorii falși ar trebui să fie instabili în eșantioane diferite, pe când factorii autentici ar trebui să fie replicați.

Practic, dacă nici unul dintre elementele prezentate mai sus nu este prezent, atunci este foarte probabil ca factorul respectiv să fie fals, și trebuie eliminat. Dacă cel puțin două elemente sunt prezente, factorul are șanse foarte mari de a fi autentic, și ar trebui reținut. Dacă doar un singur element este prezent, atunci factorul respectiv ar trebui etichetat ca fiind „discutabil”, necesitând cercetări ulterioare (Wood și colab., 1995).

Formarea pachetelor de itemi

O tehnică foarte des utilizată, atât în cazul analizei factoriale exploratorii cât și în cazul celei confirmatorii, este formarea pachetelor de itemi. Aceasta presupune gruparea mai multor itemi împreună după anumite criterii, și folosirea ulterioară a acestora ca variabile manifeste în EFA sau CFA (Little și colab., 2002; Sava, 2011). Formarea pachetelor de itemi depinde în exclusivitate de obiectivele studiului și de instrumentele folosite. Dacă scopul cercetării este de a identifica efectele unei variabile latente la un nivel superior de generalitate, atunci formarea pachetelor de itemi poate reduce sau neutraliza efectele perturbatoare ale factorilor aflați la un nivel inferior de generalitate. În aceste situații, formarea pachetelor de itemi este indicată, poate chiar necesară. În schimb, dacă scopul studiului este de a verifica particularitățile constructului măsurat la nivelul itemilor individuali, atunci minimizarea efectelor de nivel inferior ar tinde să ascundă exact efectele dorite a fi studiate. În aceste situații, formarea pachetelor de itemi este contraindicată (Little și colab., 2002).

În cazul în care pentru atingerea obiectivelor studiului este necesară formarea pachetelor de itemi, există un aspect foarte important care trebuie luat în considerare: dimensionarea setului de itemi trebuie determinată înaintea formării pachetelor (Little și colab., 2002). Atât în cazul adaptării culturale a unui test, și cu atât mai mult în cazul dezvoltării unui nou instrument psihometric, este necesară aplicarea unei analize factoriale exploratorii pentru a determina dacă setul de itemi este unidimensional sau multidimensional. Acest lucru este important deoarece tehnicile de formare a pachetelor de itemi diferă în funcție de dimensionarea setului de itemi supuși analizei. Vom prezenta în continuare cinci tehnici de formare a pachetelor de itemi, primele trei fiind pentru seturi unidimensionale de itemi, iar ultimele două pentru seturi multidimensionale de itemi (Little și colab., 2002):

a) Repartizarea aleatoare – constă în distribuirea fiecărui item, într-un mod aleator, către unul dintre pachete. Această tehnică ar trebui să ducă la formarea unor pachete care conțin, în

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

aproximativ aceeași măsură, dispersie datorată factorului comun.

b) Ponderarea item-construct – are ca scop construirea unor pachete echilibrate, din punct de vedere al relației item-construct. Se realizează în felul următor: presupunem că avem un set unidimensional de 12 itemi, și dorim să formăm 3 pachete; folosind nivelul de saturație al itemilor ca și criteriu, se aleg primii trei itemi cu cel mai ridicat nivel de saturație (ex. itemii 5, 8 și 11) și se distribuie (în ordine) celor trei pachete (itemul 5 pachetului 1, itemul 8 pachetului 2 și itemul 11 pachetului 3). Se continuă prin selectarea următorilor trei itemi (ex. itemii 1, 7 și 12) și distribuirea lor în cele trei pachete, dar în ordine inversă (itemul cu saturația cea mai ridicată se distribuie pachetului ce conține itemul din seria anterioară cu saturația cea mai scăzută: itemul 1 pachetului 3, itemul 7 pachetului 2 și itemul 12 pachetului 1). Algoritmul continuă până la epuizarea numărului de itemi analizați.

c) Reducerea erorii de aprobare – constă în gruparea în același pachet a unor itemi formulați pozitiv cu itemi formulați negativ, după ce aceștia din urmă au fost codați invers (ex. dacă avem un set unidimensional de 16 itemi, dintre care 8 sunt formulați pozitiv și 8 negativ, și dorim să formăm 4 pachete, vom include în fiecare pachet 2 itemi formulați pozitiv și 2 formulați negativ). Scopul acestei tehnici este reducerea bias-ului negativ vs. pozitiv, legat de constructul care este măsurat.

d) Abordarea consistenței interne – constă în crearea pachetelor folosind fațetele constructului măsurat ca și criteriu de grupare, și se realizează astfel: presupunem că avem un set multidimensional de 16 itemi ce conține 4 fațete (A, B, C și D), fiecare măsurată prin 4 itemi (A: 1 - 4, B: 5 - 8, C: 9 - 12, D: 13 - 16); primul pachet ar reflecta fațeta A, și ar fi format din itemii 1, 2, 3 și 4; cel de-al doilea pachet ar reflecta fațeta B, și ar fi format din itemii 5, 6, 7 și 8, ș.a.m.d..

e) Abordarea domeniului caracteristic – constă în combinarea itemilor ce aparțin unor fațete diferite pentru a forma pachete (folosind exemplul anterior, fiecare pachet ar conține câte un item din fiecare fațetă măsurată: primul pachet ar conține itemii 1, 5, 9 și 13, cel de-al

doilea itemii 2, 6, 10 și 14, etc.). Această tehnică încearcă să explice multidimensionalitatea prin crearea unor pachete care cuprind nu numai varianța comună (ca în abordarea consistenței interne), dar și varianța aferentă fațetelor unice ale multiplelor dimensiuni.

Încheiem această secțiune prin menționarea unor aspecte complementare legate de formarea pachetelor de itemi: numărul de itemi aferenți unei pachet poate varia între 3 și 6 (Little și colab., 2002; Bandalos, 2008). Alegerea numărului de pachete ce urmează a fi formate depinde atât de numărul total de itemi utilizați, cât și de raportul itemi – factori. De exemplu, nu ar avea nici un sens ca pentru un test unidimensional de 12 itemi să formăm 6 pachete a câte 2 itemi fiecare, fiind mult mai indicat să formăm 3 pachete a câte 4 itemi fiecare. De asemenea, pentru un test multidimensional de 40 itemi – câte 10 pentru fiecare fațetă, am putea: a) să folosim abordarea consistenței interne, și am forma 8 pachete – câte două pentru fiecare fațetă, fiecare fiind compusă din 5 itemi; b) să folosim abordarea domeniului caracteristic, și atunci am putea forma câte pachete dorim, dar ar trebui să ne asigurăm că fiecare pachet conține itemi din fiecare fațetă și, preferabil, fiecare fațetă este prezentă în fiecare pachet în aceeași măsură (Little și colab., 2002).

De asemenea, mai trebuie să menționăm faptul că tehnicile pentru seturi multidimensionale de itemi pot fi aplicate în condițiile în care instrumentul măsoară fațete ale aceluiași construct (ex. agresivitatea măsurată prin fațetele agresivitate fizică, verbală, furie și ostilitate). În cazul în care instrumentul măsoară factori care au o anumită independență (ex. un test care măsoară factorii modelului Big Five de personalitate, fără a măsura însă și fațetele lor), se recomandă folosirea unor tehnici de formare a pachetelor unidimensionale pentru itemii aferenți fiecărui factor în parte. În schimb, în cazul unui test ce măsoară și fațetele modelului Big Five, se poate aplica una dintre tehnicile multidimensionale de formare a pachetelor pentru fiecare factor (cu fațetele aferente), după modelul prezentat în exemplul anterior.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Concluzii

Scopul acestui articol a fost oferirea de informații relevante despre aspectele care trebuie luate în considerare atunci când analiza factorială exploratorie este utilizată în cazul itemilor. Toate informațiile și sugestiile prezentate au ca scop îmbunătățirea calității utilizării EFA, în procesul de adaptare culturală sau dezvoltare a testelor. Astfel, acest articol poate fi folositor atât cercetătorilor care se ocupă cu dezvoltarea sau adaptarea culturală a instrumentelor psihometrice pe populație românească, cât și studenților (în principal masteranzi sau doctoranzi) a căror teme de disertație/doctorat includ astfel de analize.

La finalul acestui articol prezentăm un set de recomandări sub formă de ghid, bazat pe informațiile anterioare, legat de modalitatea optimă de aplicare a EFA în cazul dezvoltării/ adaptării culturale a instrumentelor psihometrice:

1. Verificați dacă EFA este tehnica statistică de care aveți nevoie pentru atingerea obiectivelor studiului.

2. Selectați un eșantion care să respecte atât criteriile legate de volum (un raport participanți-itemi de minim 5:1) cât și legate de caracteristici (omogenitate medie și similaritate între eșantion și populația țintă a probei).

3. Verificați (pentru precauție) valorile Skewness și Kurosis ale itemilor înainte de a începe analiza. Dacă valorile indică pentru mai mulți itemi distribuții ce deviază substanțial de la normalitate, puteți continua analiza din programul SPSS (asumându-vă riscul de a obține o structură factorială neadecvată), sau puteți apela la programul FACTOR.

4. Realizați o EFA din programul SPSS, urmând următoarele recomandări: metodă de extracție – analiza factorilor principali; metodă de selecție – criteriul lui Kaiser (deși nu este o soluție adecvată, se poate folosi în cazul realizării unei EFA incipientă, doar pentru a verifica câți factori selectează programul fără nici o constrângere); rotire – Direct Oblimin (cu valoarea zero pentru delta).

5. Verificați *scree plot*-ul obținut și apoi realizați o analiză paralelă din sintaxa programului SPSS, folosind ghidul din Anexa 1. Dacă soluția indicată de analiza paralelă este susținută

și de *scree plot*, puteți fi destul de sigur că e soluția corectă. Dacă cele două diferă, veți alege soluția care are cel mai ridicat grad de interpretabilitate. În cazul în care ambele soluții au un nivel scăzut de interpretabilitate, puteți încerca realizarea EFA utilizând unul dintre programele indicate în secțiunile anterioare.

6. Realizați din nou o EFA, selectând de această dată numărul de factori indicat de soluția aleasă anterior. Matricea structurii factoriale va conține structura pe care o veți analiza.

7. Nu uitați ca în momentul în care veți raporta rezultatele obținute să menționați toate alegerile făcute, incluzând o justificare succintă pentru fiecare.

Bibliografie

- Armstrong, J. S. (1967). Derivation of theory by means of factor analysis or Tom Swift and his electric factor analysis machine. *The American Statistician*, 21(1), 17-21.
- Bandalos, D. L. (2008). Is Parceling Really Necessary? A Comparison of Results From Item Parceling and Categorical Variable Methodology. *Structural Equation Modeling: A Multidisciplinary Journal*, 15(2), 211-240.
- Bernstein, I. H., Teng, G. (1989). Factoring items and factoring scales are different: Spurious evidence for multidimensionality due to item categorization. *Psychological Bulletin*, 105, 467-477.
- Browne, M. W. (2001). An Overview of Analytic Rotation in Exploratory Factor Analysis. *Multivariate Behavioral Research*, 36(1), 111-150.
- Browne, M. W., Cudeck, R., Tateneni, K. & Mels G. (2008). CEFA: Comprehensive Exploratory Factor Analysis, Version 3.03 [Computer software and manual]. Găsit la adresa: <http://faculty.psy.ohio-state.edu/browne>.
- Buja, A., & Eyuboglu, N. (1992). Remarks on parallel analysis. *Multivariate Behavioral Research*, 27, 509-540.
- Dinno, A. (2009). Exploring the Sensitivity of Horn's Parallel Analysis to the Distributional Form of Random Data. *Multivariate Behavioral Research*, 44(3), 362-388.
- Fabrigar, L. R., Wegener, D. T., MacCallum, R. C., & Strahan, E. J. (1999). Evaluating the use of exploratory factor analysis in psychological research. *Psychological Methods*, 4(3), 272-299.
- Fava, J. L., & Velicer, W. F. (1992). The effects of overextraction on factor and component analysis. *Multivariate Behavioral Research*, 27, 387-415.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

- Ford, J.K., MacCallum, R.C. & Tait, M. (1986). The Application of Exploratory Factor Analysis in Applied Psychology: A Critical Review and Analysis. *Personnel Psychology*, 39(2), 291-314.
- Gorsuch, R. L. (1997). Exploratory Factor Analysis: Its Role in Item Analysis. *Journal of Personality Assessment*, 68(3), 532-560.
- Gould, S. J. (1981). *The mismeasure of man*. New York: Norton.
- Hayashi, K., Bentler, P. M., & Yuan, K.-H. (2007). On the Likelihood Ratio Test for the Number of Factors in Exploratory Factor Analysis. *Structural Equation Modeling: A Multidisciplinary Journal*, 14(3), 505-526.
- Hayton, J. C., Allen, D. G., & Scarpello, V. (2004). Factor Retention Decisions in Exploratory Factor Analysis: a Tutorial on Parallel Analysis. *Organizational Research Methods*, 7(2), 191-205.
- Henson, R. K. (2006). Use of Exploratory Factor Analysis in Published Research: Common Errors and Some Comment on Improved Practice. *Educational and Psychological Measurement*, 66(3), 393-416.
- Hogarty, K. Y. (2005). The Quality of Factor Solutions in Exploratory Factor Analysis: The Influence of Sample Size, Communalities, and Overdetermination. *Educational and Psychological Measurement*, 65(2), 202-226.
- Horn, J. L. (1965). A rationale and technique for estimating the number of factors in factor analysis. *Psychometrika*, 30, 179-185.
- International Test Commission (2010). *International Test Commission Guidelines for Translating and Adapting Tests*. Găsit la adresa: <http://www.intescom.org>.
- Little, T. D., Cunningham, W. A., & Shahar, G. (2002). To Parcel or Not to Parcel : Exploring the Question, Weighing the Merits. *Structural Equation Modeling: A Multidisciplinary Journal*, 9(2), 151-173.
- Lorenzo-Seva, U., & Ferrando, P. J. (2006). FACTOR: a computer program to fit the exploratory factor analysis model. *Behavior research methods*, 38(1), 88-91.
- MacCallum, R.C., Widaman, K.F., Zhang, S. & Hong, S. (1999). Sample size in factor analysis. *Psychological Methods*, 4(1), 84-99.
- O'Connor, B. P. (2000). SPSS and SAS programs for determining the number of components using parallel analysis and velicer's MAP test. *Behavior research methods, instruments, & computers*, 32(3), 396-402.
- Panter, A. T., Swygert, K. A., Dahlstrom, W. G., & Tanaka, J. S. (1997). Factor analytic approaches to personality item-level data. *Journal of Personality Assessment*, 68, 561-589.
- Sass, D. (2009). Factor Loading Estimation Error and Stability Using Exploratory Factor Analysis. *Educational and Psychological Measurement*, 70(4), 557-577.
- Sava, F.A. (2011). *Analiza datelor în cercetarea psihologică*. Cluj-Napoca: Editura ASCR.
- De Winter, J. C. F., Dodou, D. & Wieringa, P.A. (2009). Exploratory Factor Analysis With Small Sample Sizes. *Multivariate Behavioral Research*, 44(2), 147-181.
- Van den Broeck, A., Vansteenkiste, M., De Witte, H., Soenens, B., Lens, W. (2010). Capturing autonomy, competence, and relatedness at work: construction and initial validation of the Work-related Basic Need Satisfaction scale. *Journal of Occupational and Organizational Psychology*, 83 (4), 981-1002.
- Wood, J. M., Tataryn, D. J., & Gorsuch, R. L. (1996). Effects of under- and overextraction on principal axis factor analysis with varimax rotation. *Psychological Methods*, 1(4), 354-365.

Anexa 1: Ghid pentru realizarea unei analize paralele folosind sintaxa programului SPSS

Prezentăm în continuare 10 pași simpli pentru a realiza o analiză paralelă folosind sintaxa programului SPSS*:

Pasul 1: După deschiderea programului, accesați sintaxa făcând click pe: „File › New › Syntax”.

Pasul 2: Copiați următoarele comenzi în fereastra aferentă sintaxei:

```
set mxloops=9000 printback=off width=80 seed = 1953125.  
matrix.
```

Pasul 3: Introduceți numele și locația fișierului ce conține datele pe care doriți să le analizați (fișierul trebuie să fie realizat în SPSS), după comanda „FILE = ” (ex. „FILE = C:\Date\studiu.sav”). Dacă specificați „FILE = * ”, programul va analiza fișierul SPSS care este deschis în acel moment, acest lucru fiind mult mai la îndemână. De asemenea, trebuie să specificați variabilele din baza de date ce vor fi folosite pentru analiză (ex. dacă ați numerotat itemii pe care doriți să-i analizați cu it1, it2, it3, ..., it20, atunci veți scrie „VAR = it1 to it20”). Comanda aferentă descrierii de mai sus este:

```
GET raw / FILE = * / missing=omit / VAR = it1 to it20.
```

Pasul 4: Introduceți numărul seturilor de date paralele ce doriți să fie analizate (în general, 1000 de seturi sunt suficiente; în cazul în care soluția obținută nu este îndeajuns de clară, puteți utiliza mai multe seturi de date):

```
compute ndatsets = 1000.
```

Pasul 5: Introduceți valoarea percentilului dorit pentru realizarea analizei (de obicei, percentilul 95):

```
compute percent = 95.
```

Pasul 6: Alegeți metoda de extracție a factorilor pe care doriți să o utilizați („1” pentru analiza componentelor principale, sau „2” pentru analiza factorilor principali). Indiferent de metoda de extracție pe care ați folosit-o, veți selecta opțiunea 1, deoarece selectarea celei de-a doua opțiuni tinde să supraestimeze numărul de factori care trebuie extrași (Buja și Eyuboglu, 1992; Dinno, 2009):

```
compute kind = 1.
```

Pasul 7: Alegeți modalitatea de generare a seturilor de date paralele („1” pentru generarea unor date aleatoare cu distribuție normală, sau „2” pentru permutarea datelor brute). Alegeți prima variantă în cazul în care itemii supuși analizei au distribuții încadrate în limitele normalității. Cea de-a doua variantă este foarte precisă în situațiile în care datele brute nu au o distribuție normală, dar durează mai mult timp. Dacă itemii utilizați în analiză încalcă condițiile de normalitate, realizați mai întâi o analiză paralelă folosind prima metodă (pentru a avea un punct de reper), iar apoi folosiți cea de-a doua metodă.

```
compute randtype = 1.
```

* Ghidul a fost inspirat din indicațiile Prof. Univ. Dr. Brian O'Connor despre analiza paralelă, disponibile pe internet la adresa: <https://people.ok.ubc.ca/briocconn/nfactors/nfactors.html>.

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

Pasul 8: Introduceți următoarele comenzi în fereastra aferentă sintaxei:

```

compute ncases = nrow(raw).
compute nvars = ncol(raw).

do if (kind = 1 and randtype = 1).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute realeval = eval(d * vcv * d).
compute evals = make(nvars,ndatsets,-9999).
loop #nds = 1 to ndatsets.
compute x = sqrt(2 * (ln(uniform(ncases,nvars)) * -1) ) &*
      cos(6.283185 * uniform(ncases,nvars) ).
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
end if.

do if (kind = 1 and randtype = 2).
compute nm1 = 1 / (ncases-1).
compute vcv = nm1 * (sscp(raw) - ((t(csum(raw))*csum(raw))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute realeval = eval(d * vcv * d).
compute evals = make(nvars,ndatsets,-9999).
loop #nds = 1 to ndatsets.
compute x = raw.
loop #c = 1 to nvars.
loop #r = 1 to (ncases - 1).
compute k = trunc( (ncases - #r + 1) * uniform(1,1) + 1 ) + #r - 1.
compute d = x(#r,#c).
compute x(#r,#c) = x(k,#c).
compute x(k,#c) = d.
end loop.
end loop.
compute vcv = nm1 * (sscp(x) - ((t(csum(x))*csum(x))/ncases)).
compute d = inv(mdiag(sqrt(diag(vcv)))).
compute evals(:,#nds) = eval(d * vcv * d).
end loop.
end if.

```

Folosirea analizei factoriale exploratorii în cazul itemilor: Aspecte specifice și recomandări

```

compute num = rnd((percent*ndatsets)/100).
compute results = { t(1:nvars), realeval, t(1:nvars), t(1:nvars) }.
loop #root = 1 to nvars.
compute ranks = rnkorder(evals(#root,:)).
loop #col = 1 to ndatsets.
do if (ranks(1,#col) = num).
compute results(#root,4) = evals(#root,#col).
break.
end if.
end loop.
end loop.
compute results(:,3) = rsum(evals) / ndatsets.

print /title="PARALLEL ANALYSIS:".
do if (kind = 1 and randtype = 1).
print /title="Principal Components & Random Normal Data Generation".
else if (kind = 1 and randtype = 2).
print /title="Principal Components & Raw Data Permutation".
end if.
compute specifics = {ncases; nvars; ndatsets; percent}.
print specifics /title="Specifications for this Run."
/rlabels="Ncases" "Nvars" "Ndatsets" "Percent".
print results
/title="Raw Data Eigenvalues, & Mean & Percentile Random Data Eigenvalues"
/clabels="Root" "Raw Data" "Means" "Prcntyle" /format "f12.6".

compute root = results(:,1).
compute rawdata = results(:,2).
compute percntyl = results(:,4).

save results /outfile=* / var=root rawdata means percntyl .

end matrix.

TSPLIT VARIABLES= rawdata means percntyl /ID= root /NOLOG.

```

Pasul 9: Executați sintaxa făcând click pe: „Run > All”.

Pasul 10: Alegeți numărul optim de factori ce trebuie reținuți, comparând valorile *eigenvalue* din datele reale cu cele din datele aleatoare. Trebuie să menționăm faptul că, în varianta originală a procedurii (Horn, 1965), valorile *eigenvalue* medii din datele aleatoare erau utilizate pentru comparație. O recomandare curentă este utilizarea valorilor *eigenvalue* din datele aleatoare ce corespund percentilui ales (O'Connor, 2000).