KONFIRMATÍV FAKTORANALÍZIS ROP-R-REL

Vargha András
KRE BTK Pszichológiai Intézete
ELTE PPK Pszichológiai Intézete
vargha.andras@kre.hu

Összefoglaló

A pszichológiai tesztek szerkezeti érvényességének, strukturális validitásának a vizsgálatában kiemelt fontosságú módszer a konfirmatív (megerősítő) faktoranalízis, szokásos rövidítéssel CFA. Ezt a módszert nem is olyan régen még csak borsos árú szoftverek segítségével lehetett elvégezni. Ma már szerencsére több – mindenki által elérhető – ingyenes szoftver is tartalmaz CFA-modult. Cikkünk a CFA elméleti hátterét is felvázolva mutatja be a legújabb ilyen szoftver, a ROP-R használatát a CFA végrehajtására, kitérve arra is, hogy milyen lehetőségek állnak rendelkezésre a szintén ingyenes jamovi és JASP szoftver esetében. A CFA-elemzés lényegének megértését a cikkben több szemléltető példa könnyíti meg, amelyek egy valódi pszichológiai kutatásból származnak.

Kulcsszavak: konfirmatív faktoranalízis, CFA, ROP-R, JASP, jamovi

Pszichometriai háttér

Számos pszichológiai konstrukció mérésére használunk pszichológiai teszteket. Például a személyiségvonások ötfaktoros (extraverzió, barátságosság, lelkiismeretesség, érzelmi stabilitás és nyitottság) terét szokásosan különböző Big Five tesztekkel mérik. Ilyen többek között a Caprara és munkatársai (1993) által szerkesztett *Big Five Questionnaire* (BFQ) vagy a John és Srivastava (1999) nevéhez fűződő *Big Five Inventory* (BFI). Hasonlóképpen a mentális egészség konstrukcióját mérhetjük például a Mentális Egészség Teszt (MET) segítségével, melynek 5 skálája a jóllét, a savoring, az alkotó-végrehajtó hatékonyság, az önreguláció és a reziliencia (Vargha et al., 2020).

Egy megadott skálaegyüttessel rendelkező teszt szerkezeti érvényességének – más szavakkal strukturális validitásának – az igazolásához azt kell kimutatni, hogy létezik egy olyan faktormodell, amelyre a teszt skálái jól illeszkednek. Ehhez szükséges, hogy a teszt tételeinek terében létezzen egy pontosan annyi dimenziós faktorstruktúra, ahány skálája van a tesztnek, és hogy az

DOI: 10.17627/ALKPSZICH.2024.1.149

egyes tételek skálabesorolásuk szerint illeszkedjenek ezekre a dimenziókra.

Mindezek részletei akkor lesznek világosak és érthetők, miután ismertettük a faktoranalízis fogalmát általánosságban, majd ezen belül a konfirmatív (megerősítő, modelltesztelő) faktoranalízist,[1] mely névre a továbbiakban a hazai és a nemzetközi irodalomban is meghonosodott CFA rövidítéssel hivatkozunk.

STATISZTIKAI HÁTTÉR: A FAKTORANALÍZIS ALAPFOGALMAI

50-60 évvel ezelőtt még a statisztikusok többsége is azt gondolta, hogy a standard faktoranalízis egy főkomponens-analízis (FKA),[2] amelynél az 1-nél nagyobb sajátértékű főkomponenseken Varimax-rotációt hajtunk végre. Nem kevesen még ma is ezt gondolják, amihez az olyan népszerű statisztikai szoftverek is hozzájárulnak, mint például az SPSS. Belenézve ugyanis az SPSS legújabb, 29. verziójának menürendszerébe, faktoranalízist a Dimension Reduction elemzési menüpont Factor nevű moduljával tudunk végrehajtani, de ez alapértelmezés szerint egy sima FKA-t jelent. Ezzel összhangban egy ilyen FKA-elemzés outputjának fejlécén is a „Factor Analysis" címet találjuk. Mi tehát a különbség az FKA és a faktoranalízis általános modellje között?

Az FKA lényege, hogy az eredeti változók súlyozott összegeiként kevés számú korrelálatlan összetevőt, ún. főkomponenst *(principal components)* hozunk létre, amelyek az eredeti változók teljes varianciájának lehető legnagyobb részét megmagyarázzák (Tabachnick & Fidell, 2013; Vargha, 2019). A főkomponenseket egymás után, egyenként hozzuk létre az alábbi algoritmus szerint.

1. Az első főkomponens a változók olyan súlyozott összege, amely a változókkal a lehető legszorosabb lineáris kapcsolatban van, azaz amelynél a lehető legnagyobb az egyes változókkal számított korrelációk négyzetösszege.
2. Minden további főkomponensre is ugyanez igaz, csak azon feltétel mellett, hogy az újabb főkomponensek korrelálatlanok az összes előttük kiszűrt főkomponenssel.
3. Adott főkomponens esetén az egyes változókkal számított korrelációk négyzetösszegét a főkomponens sajátértékének nevezzük, és ez gyakorlatilag azt mutatja, hogy a főkomponens hány változónyi varianciát magyaráz meg, illetve fed le.
4. Az egyértelműség kedvéért a főkomponenseket standardizált (0 átlagú és 1 szórású) alakban definiáljuk.

Logikájában az FKA hasonlít a többszörös lineáris regresszió módszeréhez. A különbség: az FKA-ban nem egyetlen függő változót, hanem az összes vizsgált változót szeretnénk „megmagyarázni", és nem független változók egy csoportja, hanem saját magukból gyártott új változók segítségével, amelyek tömörebben tartalmazzák az általuk lefedett információt. Az FKA fő célja tehát a változók számának csökkentése a takarékosság jegyében, hogy további elemzésekben már elég legyen ezzel a kevés számú új változóval dolgozni. Elvárjuk, hogy a megtartott főkomponensek a változók teljes

[1] Angolul *confirmatory factor analysis* (CFA).
[2] Angolul *principal component analysis,* szokásos rövidítéssel PCA.

varianciájának minimum 50-60%-át magyarázzák, de jónak csak a 70% fölötti megmagyarázott varianciájú megoldásokat tekintjük (lásd pl. Shaharudin et al., 2018).

Az FKA-val szemben a faktoranalízis alapgondolata, hogy létezik egy közvetlenül nem megfigyelhető, látens faktorstruktúra, mely valamilyen módon kifejezi, képviseli vizsgált változóink (korrelációs vagy kovarianciamátrixuk által tartalmazott) közös lényegét, és megfigyelt változóink értékei lényegében ezen látens faktorok hatásaként alakulnak. A faktoranalízis egyik fő célja e faktorstruktúra feltárása, amit feltáró, exploratív faktoranalízis[3] (röviden EFA) segítségével hajthatunk végre (lásd pl. Vargha, 2019). A faktoranalízis másik fő célja pedig ennek ellenőrzése, megerősítése, ha van egy konkrét elképzelésünk a faktormodellről, például a faktorok száma és a faktorokra illeszkedő változók (teszttételek) tekintetében. Ezt az elemzést a CFA segítségével végezhetjük el.

A faktoranalízis modelljében az elemzésbe bevont minden változó varianciája két összetevőre bontható:

1. a több változóban is jelen levő, közös faktorok által meghatározott rész;
2. a csak erre a változóra jellemző, s ily módon a többi változótól független egyedi rész (egyediség vagy reziduális), mely tartalmazza az ezen változó mérésével kapcsolatos hibát is.

Az FKA-val ellentétben a faktoranalízis csak a változók által közösen lefedett részre alkot modellt, nem célja a változók teljes varianciáját megmagyarázni. Az FKA esetében a hangsúly az eredeti (megfigyelt, manifeszt) változókon van, a faktoranalízis esetében viszont a közvetlenül nem megfigyelhető, látens faktorokon, amelyek a dolgok igazi lényegét fejezik ki, megbújnak a háttérben, és onnan hatnak mért változóinkra. Az FKA főkomponensei nem látensek, azok az eredeti változók rögzített szabály szerint képezett transzformáltjai (súlyozott összegei), konkrét mintán személyenként egyértelműen meghatározható konkrét értékűek. Ezzel szemben a faktoranalízis látens faktorait konkrét mintán is csak becsülni tudjuk, ahogy a köztük lévő korrelációkat vagy faktorsúlyaikat, s ezek értékei a becslési módszertől is függő közelítéseknek tekinthetők.

Az EFA-ban több módszer is létezik a faktorstruktúra feltárására, ezek közül a leggyakoribb a *maximum likelihood* (ML), a főfaktorelemzés (*principal axis factoring* vagy PAF) és a minimum reziduális módszer (*minimum residual* vagy MinRes) (lásd Osborn, 2014; Tabachnick & Fidell, 2013; Vargha, 2019).

Mivel az EFA-ban a változóknak csak a közös (legalább két változó által lefedett) részét faktorizáljuk, egy jó faktorstruktúra létezésének elengedhetetlen feltétele, hogy a változók közös része, amit a Kaiser-Meyer-Olkin (röviden KMO) mutatóval szoktak leggyakrabban mérni, kellően nagy legyen (vö. Vargha, 2019). Fentiekből következően az EFA-ban feltárt faktorok az eredeti változók varianciájának csak kisebb részét magyarázzák, mint az FKA főkomponensei.

A kapott faktorstruktúra értelmezhetősége mind az FKA, mind az EFA esetében hasznos és fontos szempont. Bár az FKA esetében nem modellt, hanem csupán a változóinkat megfelelő mértékben helyettesíteni képes komponenseket keresünk, ezek értelmes és jelentéssel felruházható volta

[3] Angolul *exploratory factor analysis,* szokásos rövidítéssel EFA.

megkönnyíti a velük végzett későbbi elemzéseket. Ezt az elsődlegesen feltárt struktúra, a megtartott főkomponensek, illetve faktorok forgatásával (rotációjával) lehet elérni. A forgatás egy matematikai művelet valamilyen többdimenziós térben, melynek technikai részletei pszichológiai szempontból nem lényegesek. A lényeg az, hogy a forgatás eredményeképpen létrehozott komponensek[4] (FKA), illetve faktorok (EFA) jobban illeszkednek az eredeti változókra, mint forgatás előtti megfelelőik, így azok az értelmezést megkönnyítő szakmai jelentéssel is felruházhatók. A forgatás (angolul *rotation*) technikai részleteiről mindössze annyit érdemes tudni, hogy több iterációs lépésben történik, minden lépésben törekedve arra, hogy a változók komponens-, illetve faktorsúlyai abszolút értékben vagy az 1 felé közelítsenek (ilyenek alkalmasak a komponensek, illetve a faktorok értelmezésére), vagy 0 közeliek legyenek.

Fontos, hogy a forgatott főkomponensek, illetve faktorok a változók varianciájából együttesen pontosan akkora részt magyaráznak, mint a nem forgatott főkomponensek, illetve faktorok, vagyis a forgatással az összes megmagyarázott rész aránya nem változik. A forgatással tehát nem többet, hanem jobban tudunk megmagyarázni a változókból. A forgatott faktorok jobban szeparálódnak egymástól, mint az eredetiek, azaz a változók többsége (esetleg az összes) csak egy-egy faktorra ad erős töltést. Optimális esetben a mért változók tehát a faktoroknak megfelelő, nem átfedő csoportokat alkotnak.

Az EFA-ban használt legfontosabb fogalmak, amelyek részben a CFA-ban is szerepet kapnak, az alábbiak.

- *Elsődleges faktorsúlymátrix (primary factor loading matrix):* standardizált regressziós együtthatók, amelyeket az EFA-ban a választott módszerrel elsődlegesen feltárt látens faktorok együttese mint független változók segítségével az elemzésbe bevont változók regressziós előrejelzésére kapunk, s ami a standardizálás és a látens faktorok korrelálatlansága (ortogonális szerkezete) miatt megegyezik a változók és a látens faktorok becsült korrelációs mátrixával.
- *Forgatott faktorsúlymátrix (rotated factor loading matrix)*, más néven *mintázatmátrix (pattern matrix):* standardizált regressziós együtthatók, amelyeket a forgatott látens faktorok együttese, mint független változók segítségével a CFA-ba bevont változók regressziós előrejelzésére kapunk. Ez ortogonális (Varimax, Equamax stb.) forgatás esetén megegyezik a változók és a látens faktorok becsült korrelációs mátrixával, de ferde (pl. Promax vagy Oblimin) forgatás esetén a két mátrix különbözik. Elsődlegesen e mátrix alapján szokták értelmezni a feltárt faktorstruktúrát.
- *Struktúramátrix (structure matrix):* a változók és a látens faktorok becsült korrelációs mátrixa.
- *Faktorkorrelációk táblázata (matrix of factor correlations):* a forgatott látens faktorok becsült korrelációs mátrixa. Ferde forgatás esetén ebből olvashatjuk ki, hogy a forgatott faktorok milyen szoros kapcsolatban vannak egymással.
- *Kommunalitás (communality):* a végső struktúrában a faktorok által egy változó varianciájából megmagyarázott varian-

[4] Forgatás után a főkomponensek már elveszítik maximális megmagyarázó képességüket, ezért használjuk ilyenkor velük kapcsolatban a „fő” előtag nélküli komponens elnevezést.

ciaarány (a mintázatmátrixszal kapcsolatban említett, adott változóra vonatkozó regressziós becslés R^2-értéke). Egy változó 0 és 1 közötti értékű kommunalitása tehát varianciájának azon hányada, amelyet a közös faktormodell megmagyaráz. *Egyedisége* vagy *unicitása (uniqueness)* pedig az, ami a kommunalitást 1-re kiegészíti, ez az ún. *reziduális varianciaarány*. A változók kommunalitását a forgatás nem változtatja meg. A kommunalitások összege rögzített faktorszám esetén a változókból a faktorok által megmagyarázott összvariancia, melyet a változók számával leosztva az összvariancia-arányt kapjuk. Ez az arány mindig kisebb, mint ugyanazon a mintán ugyanazokkal a változókkal FKA-elemzést végezve, az azonos számú megtartott főkomponenshez tartozó kumulatív százalékarány.

A CFA lényege

Az FKA-ban főkomponensekként a változók olyan súlyozott összegeit keressük, amelyek a változók teljes varianciájának a lehető legnagyobb részét megmagyarázzák. Az EFA-ban olyan faktormodellt keresünk, amelynek faktorai a változók *közös részét* magyarázzák a lehető legnagyobb mértékben. Ennek egyik kritériuma lehet például az, hogy a faktormodell mennyire képes reprodukálni a változók közös részét képviselő kovarianciamátrixot, illetve korrelációs mátrixot (a standardizált változók kovarianciamátrixát). A CFA-ban is az EFA kritériumait alkalmazzuk, csak az alábbi plusz megszorítások mellett.

1. Minden látens faktor csak azokból a változókból (tételekből) építkezhet, amelyeket számára kijelölünk. Így a látens faktorok csak saját tételeikkel korrelálhatnak, más faktorok tételeivel nem, vagyis kereszttöltéseket nem engedünk meg.
2. A végső modellben a tételek reziduálisai egymással nem korrelálhatnak (sem egyazon faktoron belül, sem különböző faktorok között).

Ezeken a megszorításokon a modellillesztés javítása céljából olykor enyhíthetünk, ha ezt szakmai érvekkel is alá tudjuk támasztani.

Egy konkrét mintán a CFA-elemzés abból áll, hogy keresünk egy olyan faktormodellt, amely eleget tesz a fenti megszorításoknak és a lehető legjobban illeszkedik adatainkra, illetve az azokból kiszámítható korrelációs vagy kovarianciamátrixra. Az alkalmas faktormodell feltárása nem olyan egyszerű és egyértelmű, mint az FKA-ban a főkomponensek meghatározása. A faktormodellt különböző módszerek (pl. ML, MLR, MLMV, ULS, WLS[5] stb.) segítségével, minden esetben csak több lépésben, a modellt lépésről lépésre javítva lehet becsülni.

Egy faktormodell jóságának a mérésére a CFA-ban különböző együtthatók (RMSEA, CFI, TLI stb.) állnak rendelkezésre, a tesztelést pedig a más statisztikai eljárásokból is ismert χ^2-próbás illeszkedésvizsgálattal lehet elvégezni. Például RMSEA,[6] mint egy abszolút illeszkedési mutató a tételek kovarianciamátrixát és a CFA modelljéből becsült kovarianciamátrixot hasonlítja össze. Ehhez kiszámítjuk a két mátrixból az elemenkénti

[5] Az ML-lel kezdődő módszerek mind maximum likelihood becslés alapúak, az LS végződésűek pedig a legkisebb négyzetes *(Least Squares)* regressziós becslésen alapulnak.

[6] *Root Mean Square Error of Approximation*

négyzetes eltérések átlagát, az RMSEA ennek a négyzetgyöke (Kline, 2010). A CFI[7] ugyanakkor egy relatív összehasonlító mutató, mely azt méri, hogy a faktormodell tesztelésére alkalmazott χ^2 illeszkedési statisztika mennyivel kisebb (tehát a modell mennyivel jobb illeszkedésű), mint annak a lehető legrosszabb, ún. alapmodellnek a χ^2-értéke, ahol a változók és a látens faktorok együttesében minden korreláció 0, vagyis ahol semmi nem függ össze semmivel (Vargha, 2019). Az illeszkedés jóságának mérésére használt főbb mutatók összefoglalását lásd például Vargha (2019), Schreiber és munkatársai (2006), illetve Kenny (2020) munkájában.

Ha egy faktormodell a CFA-ban jónak bizonyul, még további komplexebb kérdéseket is meg lehet fogalmazni. Például igaz-e, hogy a teszt skálái egy közös konstrukcióra illeszkednek (pl. a MET 5 skálája a mentális egészség konstrukciójára)? Vagy igaz-e az, hogy egy megerősített faktormodell ugyanúgy érvényes különböző csoportokban? Előbbi kérdést például a másodrendű (Schivinski, 2013) vagy a bifaktoros (Dunn & McCray, 2020) CFA-modellek segítségével lehet megvizsgálni, utóbbit pedig a mérési invariancia *(measurement invariance)* módszerével (Chen, 2007; Kim & Yoon, 2011; Vandenberg & Lance, 2000).

Faktorstruktúrák ábrázolása, különböző faktormodellek

Egy faktormodellt a CFA-ban faktordiagram[8] segítségével szoktak ábrázolni, melynek több formája is létezik. Közös bennük, hogy a diagramon a látens faktort vagy faktorokat ellipszis vagy kör, a megfigyelt (manifeszt) változókat pedig téglalap alakú keretbe helyezik. Például egy sima egydimenziós (egyskálás) faktormodell az *1. ábrá*n látható. Az ábra a 17 tételes Globális Jóllét Kérdőív (lásd Vargha et. al., 2020) egyfaktoros struktúráját ábrázolja, melyen a kérdőív 17 tétele (GJ01É, GJ02É, …, GJ17Sp) mind egy Globális jóllét (GJ) elnevezésű közös faktorra illeszkedik. Az Érzelmi (GJÉ), a Pszichológiai (GJP), a Szociális (GJSz) és a Spirituális (GJSp) jóllét skálájával a kérdőív a jóllét bio-pszicho-szocio-spirituo modelljét operacionalizálja, hangsúlyozva, hogy a teljes jóllét feltétele, hogy az ember humán természetének minden aspektusában jól működjön, miközben jól érzi magát a bőrében (Oláh, 2021).

Az *1. ábrá*n a tételek nevének végén – a sorszám jelzete után – látható betűk a skálahovatartozásra utalnak, bár ebben az egydimenziós konstrukcióban a négy skálát külön nem vonjuk be a modellbe. Az ábrán a GJ látens faktortól a tételekhez vezető egyirányú nyilak (útvonalszakaszok) fölötti számok egy konkrét faktormodell standardizált faktorsúlybecslései (egyben a látens faktor és a tételek közti korrelációk becslései), melyeket egy Magyarország Boldogságtérképét elemző kutatás[9] 2022-es, N = 1003 fős (26% férfi, 74% nő) adatállományán elvégzett CFA-elemzés során kaptunk. A kutatás az ELTE PPK Pozitív Pszichológiai kutatócsoportja keretében (vezető: Prof. Oláh Attila), a Jobb Veled a Világ Alapítvány közreműködésével történt. Az *1. ábrá*n a téte-

7 *Comparative Fit Index*

8 Szokták a path-diagram elnevezést is használni.

9 Vö. http://boldogsagprogram.hu/magyarorszag-boldogsagterkepe-2021/

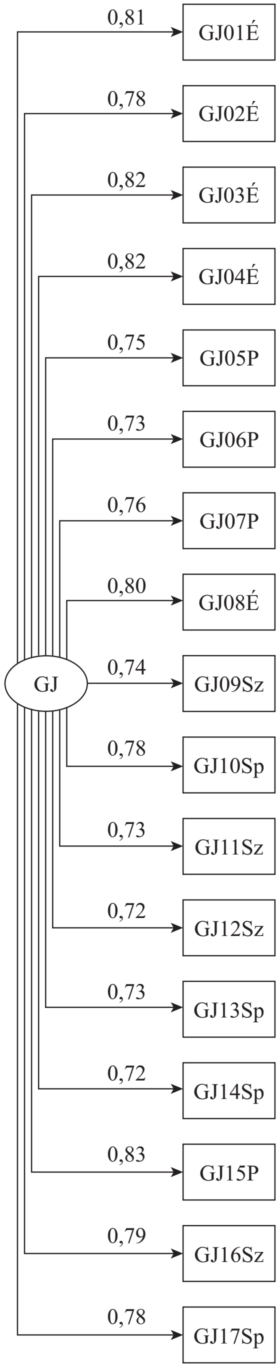
lek között azért nem látunk nyilakat, mert ebben az egydimenziós modellben alapértelmezés szerint feltételezzük, hogy minden, ami a tételekben közös, az benne van a látens faktorban, a tételek egyediségei, reziduális részei pedig egymástól függetlenek, azaz nem korrelálnak egymással. Ezen a megszorításon indokolt esetben enyhíteni szoktak, ami látszólag ellentmond annak a korábbi kijelentésünknek, hogy az EFA-ban a változók teljes közös, legalább két változó által lefedett részét faktorizáljuk. A kérdés az, hogy miért nem része a közös faktornak az, ami csak két tételt köt össze? A válasz: az EFA megpróbálja a változókban lévő összes közös részt feltérképezni és csoportosítani, a CFA pedig a közös szakmai lényeget azonosítani. Ha két tételt az a formai jellemző köt össze, hogy verbális megfogalmazásuk nagyon hasonló, és ez inspirál esetükben hasonló válaszokat, és eredményez a két tétel között pozitív korrelációt, érdemes kapcsolatukat mint egyedi összefüggést bevonni a modellbe, mely független a teszt fő konstruktumától.

Az EFA egymás mellé helyezve, nem hierarchikusan tárja fel a faktorképző tényezőket. A CFA-ban azonban lehetőségünk van arra, hogy az adatainkat egy ennél komplexebb faktormodellre illesszük. Már maga az is a komplexitás irányába tett lépés, hogy egyes tételpárok kovarianciáit beépítjük a modellbe. Az alábbiakban ismertetésre kerülő elsőrendű többfaktoros, másodrendű, majd bifaktoros modellek a komplexitás különböző szintjeit és típusait képviselik, ezek teljes spektrumát az általános SEM-modellek tartalmazzák (lásd pl. Hoyle, 2023). Előfordulhat, hogy bizonyos formai vonás (pl. több tétel fordított, negatív megfogalmazása) a komplex modellben is faktorképző tényező lesz (pl. a bifaktoros modellben[10] egy külön specifikus faktort eredményez).

Elsőrendű faktormodell több korreláló faktorral

Ha a Globális Jóllét Kérdőív tételeinek skálabesorolását is figyelembe vesszük, feltételezve, hogy a 17 tétel a négy skálának megfelelő, négy egymással korreláló látens faktorra illeszkedik, akkor a *2. ábra* modelljéhez, egy elsőrendű többfaktoros faktormodellhez *(first order multifactor model)* jutunk.

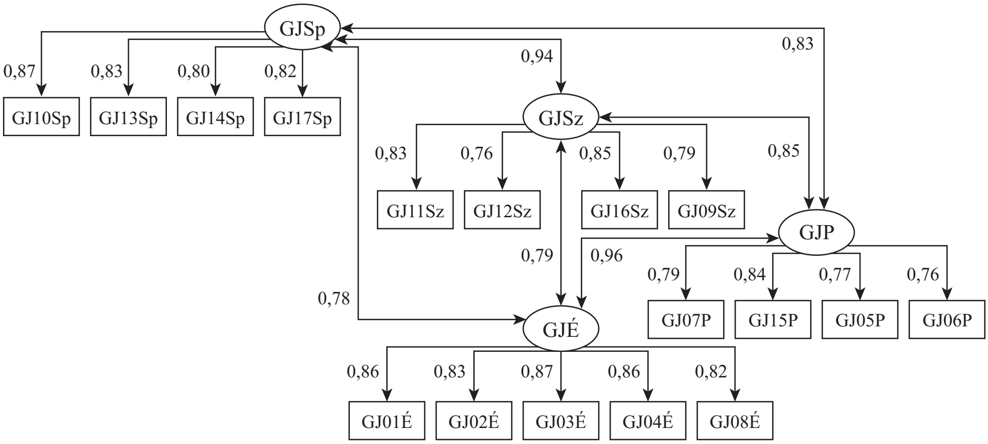
[10] Lásd egy későbbi alfejezetben.



1. ábra. A CFA grafikonja egyetlen elsőrendű látens faktor (GJ) esetén

Ilyen modelleket szoktak a standard CFA-elemzésekkel megvizsgálni (tesztelni és az illeszkedés jóságát mérni). A *2. ábrá*n a négy látens faktortól (GJÉ, GJP, GJSz és GJSp) a saját tételekhez vezető egyirányú nyilak fölötti számok itt is egy konkrét faktormodell standardizált faktorsúlybecslései, melyeket a fent említett boldogságkutatás adatain végzett elsőrendű CFA-elemzés során kaptunk. A látens faktorokat összekötő kétirányú nyilak fölötti számok a látens faktorok közötti páronkénti korrelációk (a standardizált kovarianciák) becslései.

A *2. ábrá*n sem látunk nyilakat a tételek között, mert alapértelmezés szerint ebben a modellben is feltételezzük, hogy minden, ami a tételekben közös, az benne van a négy, egymással kisebb-nagyobb mértékben korreláló látens faktorban, a tételek reziduálisai pedig egymástól függetlenek. Nyilakat a látens faktorok és a nem saját tételek között sem látunk. Ez annak a feltételezésnek felel meg, hogy minden tétel csak a saját skálájának megfelelő látens faktorra illeszkedik, a többi faktorral nem korrelál, vagyis nincsenek kereszttöltések. Indokolt esetben ezeken a megszorításokon enyhíteni szoktak.



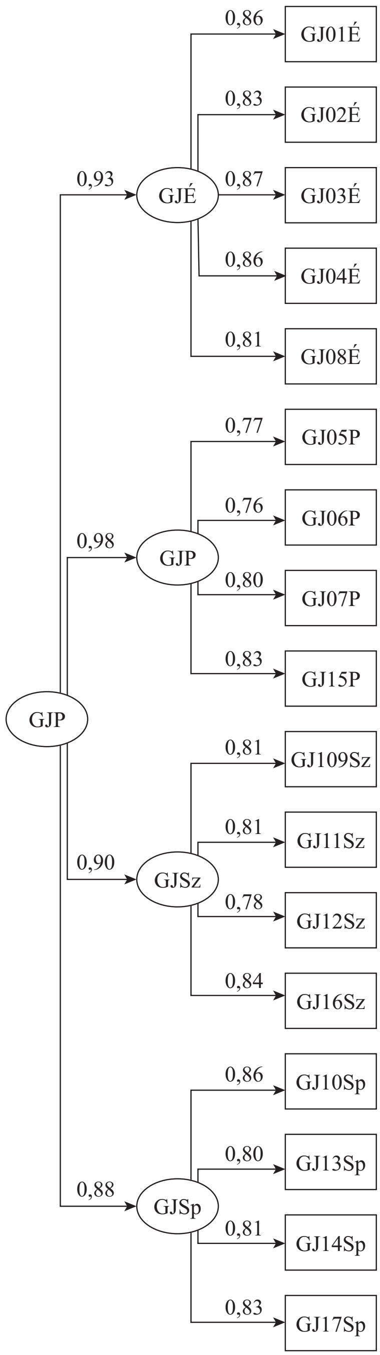
2. ábra. A CFA grafikonja négy elsőrendű látens faktor (GJÉ, GJP, GJSz és GJSp) esetén

Másodrendű faktormodell

Ha feltételezzük, hogy az elsőrendű látens faktorok egy közös másodrendű látens faktorra illeszkednek, vagyis hogy ők egy általuk kifeszített közös konstruktum különböző megnyilvánulási formái, akkor egy másodrendű faktormodellhez[11] *(second order factor model)* jutunk. A modell értelmes voltához elvárás, hogy az elsőrendű faktorok száma legalább három legyen, de ha bizonyos korlátozásokat teszünk,[12] a modell két faktor esetén is becsülhető. A másodrendű faktor képviseli az elsőrendű faktorok közös lényegét, ami nem minden elsőrendű struktúra esetén létezik. Például az 5-skálás MET esetén elképzelhető, hogy a skálák által képviselt öt látens faktor a mentális egészségnek egy másodrendű konstruktumára illeszkedik, de például a Big Five modell öt fő faktora (Extraverzió, Barátságosság, Lelkiismeretesség, Érzelmi stabilitás, Kultúra/Intellektus) esetében értelmes másodrendű faktor lehetősége fel sem merül.

A Globális Jóllét Kérdőív esetében jogos a gondolat, hogy a négy skála (Érzelmi jóllét, Pszichológiai jóllét, Szociális jóllét, Spirituális jóllét) által képviselt négy faktor egy magasabb rendű konstrukció, a globális jóllét négy megnyilvánulási formája, ezért itt a másodrendű faktormodell feltételezése szakmailag értelmes. Ezt a másodrendű modellt láthatjuk a *3. ábrá*n.

[11] Szokták a hierarchikus faktormodell *(hierarchical factor model)* elnevezést is használni.
[12] Például úgy, hogy 1-re állítjuk be az elsőrendű látens faktorok varianciáját.

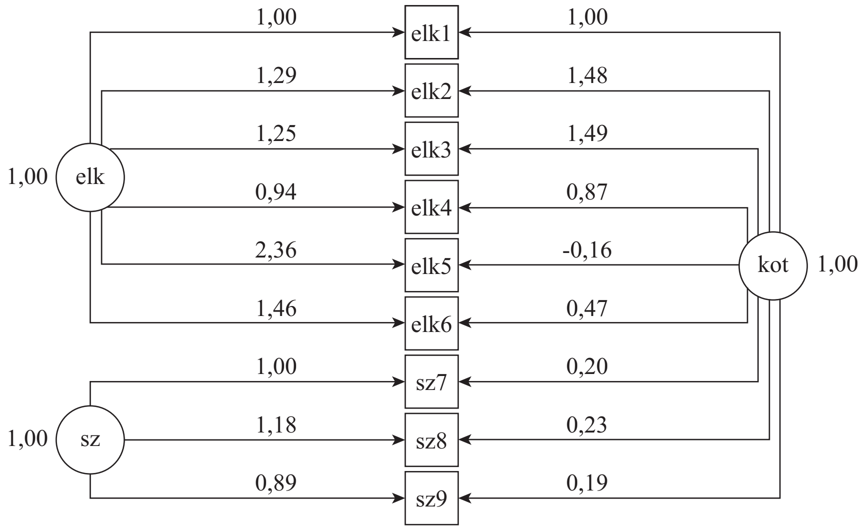


3. ábra. A CFA grafikonja négy elsőrendű (GJÉ, GJP, GJSz és GJSp) és egy másodrendű látens faktor (GJ) esetén

A *3. ábrá*n a GJ másodrendű látens faktor magyarázza a négy elsőrendű látens faktor (GJÉ, GJP, GJSz és GJSp) mindegyikét. Ez utóbbiak között azért nem látunk nyilakat, mert a modell szerint mindazt, ami bennük közös, a GJ másodrendű faktor tartalmazza. A *3. ábrá*n a nyilak fölötti számok itt is egy konkrét faktormodell standardizált faktorsúlybecslései, melyeket a fentebb már említett boldogságkutatás mintáján végzett másodrendű CFA-elemzés során kaptunk. Az ábrán a teszt 17 tétele nincs összekötve a GJ másodrendű faktorral, a modell szerint ugyanis a másodrendű faktor csak az elsőrendű faktorokon keresztül, azok közvetítésével hat az egyes tételekre. Végül az egyes tételek sincsenek egymással összekötve, mert ami bennük közös, azt saját elsőrendű faktoruk képviseli, reziduálisaik pedig függetlenek egymástól. A másodrendű modellben az elsőrendű faktorokat a másodrendű faktor alfaktorainak vagy facetjeinek tekintjük.

Bifaktoros faktormodell

A bifaktoros modell *(bifactor model)* feltételez egy általános faktort *(general factor),* amelyre minden tétel illeszkedik és több specifikus faktort *(specific factors),* amelyek rendre a tételek egy-egy alcsoportját magyarázzák (Dunn & McCray, 2020). Az általános faktor fejezi ki a tételek közös lényegét, míg a specifikus faktorok a skálák egyediségét, amit az általános faktor nem tartalmaz. Alapértelmezés szerint a specifikus faktorok nem korrelálnak sem egymással, sem az általános faktorral. Ennek a modellnek a sémáját mutatja be a *4. ábra,* ahol középen egy kötődésteszt 9 tétele látható (az első hat elkerülés-, az utolsó három pedig szorongástétel), bal oldalon az elkerülés (elk) és a szorongás (sz) specifikus faktora, a jobb oldalon pedig a kötődés (kot) általános faktora. A látens faktorok mellett látható 1,00 érték a standardizált faktorok modellben rögzített varianciája, a nyilakon látható számok pedig itt is egy konkrét faktormodell standardizált faktorsúlybecslései.



4. ábra. A bifaktoros modell sematikus ábrája az ECR-RS kötődésteszt 9 tételével (*elk* = elkerülés, *sz* = szorongás specifikus faktorok, *kot* = kötődés általános faktor)

Ezt a modellt valószínűsíti, ha a teszttételek FKA-elemzése során az első főkomponens sajátértéke legalább háromszorosa az utána következőnek (jelezve egy általános faktor létezését), és ha a tételeknek létezik egy olyan csoportosítása, amelynél az egyes csoportok mint tesztskálák létezését alkalmas EFA-elemzéssel alá tudjuk támasztani (forgatás után kellően magas faktorsúlyokkal). A „bifaktoros" jelző arra utal, hogy minden tétel két faktorra illeszkedik. Egyrészt a közös általános faktorra, másrészt egy-egy specifikus faktorra (a specifikus faktorok valamelyikére).

A másodrendű és a bifaktoros modell a teszttételek oldaláról abban tér el egymástól, hogy az a konstruktum, ami az összes tétel közös lényegét képviseli, a másodrendű modellben hierarchikus módon, az elsőrendű faktorok közvetítésével mint azok közös faktora kapcsolódik a teszttételekhez, a bifaktoros modell esetén pedig közvetlenül, direkt módon.

Az elsődleges faktorok szemszögéből pedig az a nagy különbség a két modell között, hogy a másodrendű modellben a másodrendű faktort, mely az elsőrendű faktorokon keresztül a tételek közös tartalmát is hordozza, az elsőrendű faktorok feszítik ki, tehát azokkal szoros kapcsolatban van, míg a bifaktoros modellben a specifikus faktorok és az általános faktor között nem tételezünk fel korrelációt. *Ez a legfontosabb különbség a kétféle modell között!* Matematikailag nem teljesen precízen megfogalmazva, a bifaktoros modell szerint minden tételnek van egy közös T tulajdonsága, s emellett a tételek egy részének A, másik alcsoportjának B, harmadiknak C specifikus tulajdonsága is van (pl. 3 alskála esetén), viszont sem A-nak, sem B-nek, sem

C-nek nincs köze T-hez, függetlenek tőle. Egy ilyen specifikus tulajdonság lehet például egy olyan formai dolog is, hogy a tétel átfordítandó, negatív megfogalmazású. Ugyanakkor a másodrendű modell szerint a tételek egy része A, másik része B, harmadik része C tulajdonságú, és létezik egy T tulajdonság úgy, hogy A, B és C egyaránt implikálja T-t (de egyik sem azonos vele), tehát T egy közös, általánosabb tulajdonság. Ez esetben az A, B, C tulajdonságok nem esnek kívül a fő konstruktumon (mint attól független specifikus faktorok), hanem alkotórészei annak.

A harmadik, csupán technikai jellegű eltérés az, hogy a másodrendű modellben legalább három elsőrendű faktorra van szükség, hogy értelme legyen az elsőrendű faktorok közös másodrendű faktorának, míg a bifaktoros modell esetén akár két tételcsoport (skála) is elegendő. Mindamellett speciális megkötésekkel (pl. 1-ben rögzítve a specifikus faktorok varianciáját) kétfaktoros másodrendű faktormodell is tesztelhető. Ha adott esetben szakmai szempontból mindkét modell értelmes, azt szoktuk elfogadni, amelyiknek a modellje jobb illeszkedést mutat az empirikus adatokhoz.

A CFA-elemzés, a faktormodell kiértékelése

A CFA-elemzés lényege, hogy a fentiek szerint felállítunk egy faktormodellt, majd egy megfelelően kiválasztott empirikus minta alapján – egy χ^2-próbát végrehajtva – döntünk arról, hogy a modell érvényességét állító nullhipotézis elfogadható-e (modelltesztelés). Emellett arra alkalmas adekvációs mutatók segítségével megmérjük, hogy a modell mennyire illeszkedik empirikus adatainkra. Mindezt az adott faktormodell kiértékelésének nevezzük. Különböző szóba jöhető modellek közül azt preferáljuk, amelyik a modelltesztelés során a legkevésbé bizonyul rossznak (a legkevésbé szignifikáns), s amelynek egyben az adekvációs (illeszkedési) mutatói a legjobbak.

Tekintve, hogy a modellt tesztelő χ^2-próba szignifikanciája erősen függ a minta nagyságától (pl. 1000 fősnél nagyobb minták esetén szinte mindig szignifikáns), a modell kiértékelésében döntő szerepük van az adekvációs mutatóknak. A könnyebb áttekinthetőség érdekében a legfontosabb modell-adekvációs mutatókat külön is összefoglaltuk (lásd *1. táblázat*), további mutatókkal (GFI, NFI stb.) kapcsolatban lásd Schreiber és munkatársai (2006), illetve Kenny (2020) munkáját.

1. táblázat. A CFA legfontosabb adekvációs mutatói

Mutató	**Jelentés**	**Megfelelőség**
RMSEA	Az inputváltozók kovarianciamátrixát és a CFA modelljéből becsült kovarianciamátrixot hasonlítja össze. Az RMSEA az elemenkénti négyzetes eltérések átlaga, majd ennek az átlagnak a gyöke.	0,06 alatt jó, 0,06–0,08 között elfogadható, 0,08 fölött gyenge.
C90 (RMSEA)	90%-os intervallumbecslés az elméleti RMSEA értékre.	Akkor elfogadható, ha tartalmazza a 0,05-ös értéket, vagy ha az egész intervallum 0,08 alatt helyezkedik el.

Mutató	Jelentés	Megfelelőség
pClose	Annak a próbának a *p*-értéke, mely azt teszteli, hogy RMSEA kisebb-e 0,05-nél: Prob(RMSEA ≤ 0,05).	Akkor jó, ha nagyobb 0,05-nél, vagyis ha nem szignifikáns a próba.
χ^2/f	Takarékossági mutató, melynél χ^2 azt az illeszkedést méri, ami az RMSEA-ban is benne van, „*f*" a modell szabad paramétereinek a száma (a χ^2-statisztika szabadságfoka), így a hányados az egy paraméterre eső átlagos χ^2 komponenst adja meg.	Akkor jó, ha 3,5-nél kisebb, 3,5–5 között elfogadható. Gyengesége, hogy függ a minta *N* nagyságától is (*N* növekedésével nő)
SRMR	A hipotetikus modell és a megfigyelt modell korrelációs mátrixa közötti eltérés négyzetgyöke. Minél közelebb van az SRMR a nullához, annál jobban illeszkedik a modellünk.	0,05 alatt jó, 0,08 alatt elfogadható az illeszkedés.
CFI	Egy relatív mutató, mely azt méri, hogy a modell tesztelésére alkalmazott χ^2-statisztika mennyivel kisebb (tehát a modell mennyivel jobb illeszkedésű), mint annak a lehető legrosszabb, ún. alapmodellnek a χ^2-értéke, ahol a változók és a látens faktorok együttesében minden korreláció 0. Értéke 0 és 1 között lehet (Vargha, 2019).	0,95 felett kiváló, 0,90 felett elfogadható, 0,90 alatt gyenge.
TLI	CFI-hez hasonló módon definiált mutató, mely többnyire kisebb, mint CFI. Értéke legtöbbször 0 és 1 közé esik, ritkán azonban kicsivel 1 fölé is mehet (Hu & Bentler, 1999; Tucker & Lewis, 1973).	0,95 felett kiváló, 0,90 felett jó, 0,85 felett elfogadható, 0,85 alatt gyenge.
AIC, BIC	Akaike[13] és Bayes[14]-féle információs kritériummutató, amelyekkel ML-típusú becsléseknél ugyanazon adatállomány több modellje hasonlítható össze.[15]	A kisebb AIC, illetve BIC értékű modell a jobb.

CFA végrehajtása ROP-R segítségével

A CFA-elemzés végrehajtására számos szoftver alkalmas. A jelen cikk három ingyenesen használható szoftver lehetőségeit tekinti át. Közülük részletesen bemutatjuk a ROP-R erre alkalmas modulját, de kitérünk a jamovi (The jamovi project, 2022) és a JASP (JASP Team, 2023) szoftver CFA-moduljának lehetőségeire is.

ROP-R16 egy R szoftveralapú (R Core Team, 2021), de ROPstat keretben (Vargha, 2016) működő többváltozós statisztikai szoftver, mely Windows operációs rendszerben futtatható. Letöltését és használatát illetően lásd Vargha (2023), Vargha & Bánsági (2022), illetve Vargha és munkatársai (2024). ROP-R magyar-angol kétnyelvű, melyben – egy magyar és egy brit zászló – segítségével bármikor át lehet váltani egyik nyelvről a másikra. Angolra váltva az output is angol nyelvű lesz, aminek a segítségével könnyen

13 Vö.: Akaike (1974) és https://en.wikipedia.org/wiki/Akaike_information_criterion

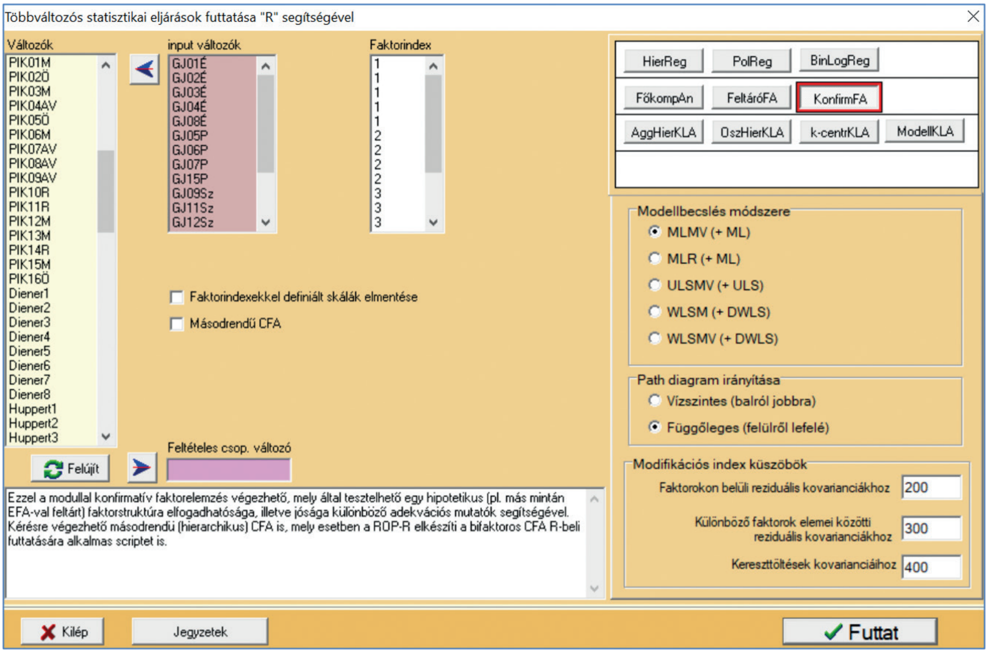
14 Vö.: Schwarz (1978) és https://en.wikipedia.org/wiki/Bayesian_information_criterion

15 Az AIC és a BIC definíciójával, illetve a jelentésük közti különbséggel kapcsolatban lásd Kuha (2004), illetve: https://www.linkedin.com/pulse/aicbic-model-selection-richard-randa

16 Lásd www.ropstat.com

megfeleltethetők egymásnak a magyar és az angol statisztikai szakkifejezések. Emellett a cikkünk végén található mellékletben külön is összefoglaljuk a faktoranalízissel kapcsolatos fogalmakat, angol megfelelőikkel és jelentésükkel együtt.

CFA-elemzéseket ROP-R-ben a *Többváltozós_elemzések_R_segítségével* menüpont *Konfirmatív faktoranalízis* modulja (röviden KonfirmFA vagy CFA) segítségével végezhetünk el (lásd *5. ábra*). Ez a modul egyaránt alkalmas elsőrendű egy- és többfaktoros, másodrendű, valamint bifaktoros elemzések végrehajtására, amelyeket majd alant részletezünk, valódi pszichológiai adatokkal szemléltetve.



5. ábra. A CFA-modul menüablaka ROP-R-ben

A CFA menüablak szolgál többek között az elemzésbe bevonandó változók kiválasztására (inputváltozók ablaka), skálahovatartozásuk jelzésére (Faktorindexablak), a CFA modellbecslés módszerének megválasztására és a modifikációs indexküszöbök beállítására (lásd *5. ábra*). Ezekhez az alábbi eligazítást fűzzük.

Inputváltozók kiválasztása

A CFA-hoz szükséges manifeszt változók, a teszttételek kiválasztásához a Változók feliratú ablakból kell az elemzéshez szükséges tételeket (kijelölésük után akár egy csomagban) az „inputváltozók" ablakba áttenni, a két ablak közti nyílra kattintva. Ezek a változók lesznek a CFA megfigyelt, manifeszt változói. Elméletileg elvárjuk tőlük, hogy folytonos, normális eloszlásúak legyenek, de manapság már léteznek a CFA-ban olyan robusztus módszerek (lásd alább), amelyek ezt a szigorú korlátozást enyhítik, s akár kétértékű nominális (pl. férfi – nő vagy beteg – nem beteg) vagy ordinális változók elemzését is lehetővé teszik. A gyakorlatban a legtöbb kutatás az 5 vagy több kategóriát tartalmazó ordinális változókat folytonosként kezeli, és van némi bizonyíték arra, hogy ez az egyszerűsítés nem gyakorol nagy hatást az eredményekre (vö. Babakus et al., 1987; Dolan, 1994; Johnson & Creech, 1983; Hutchinson & Olmos, 1998). A továbbiakban az egyszerűség kedvéért folytonos változóként hivatkozunk mi is az ilyen változókra, míg kategoriális jelzővel a 2 értékű vagy a 3-4 értékű ordinális változókat illetjük. E terminológia majd a CFA megfelelő módszerének kiválasztása során kap szerepet.

Skálahovatartozás megadása

A Faktorindexablakban egy faktorindex segítségével jelölhető ki, hogy a kiválasztott változók (tételek) rendre mely faktorokhoz (skálákhoz) sorolandók. A faktorindex értéke 1 és 9 közötti egész szám lehet. Az azonos faktorindexű változók azonos faktorhoz tartoznak.

Modellbecslés módszere

A CFA-ban valamely faktormodell elfogadhatóságát szokásosan egy χ^2-statisztikával teszteljük, illeszkedésének jóságát pedig különböző adekvációs mutatók segítségével mérjük. A definiált faktormodell becslésére három alapmódszer áll rendelkezésre (Li, 2016; Shi & Maydeu-Olivares, 2020):

- ML = maximum likelihood módszer;
- ULS = *unweighted least squares,* vagyis a súlyozatlan legkisebb négyzetek módszere;
- DWLS = *diagonally weighted least squares,* vagyis a diagonálisan súlyozott legkisebb négyzetek módszere.

Az ML a strukturális egyenletmodellek, és ezen belül a CFA legszélesebb körben használt faktormodell becslési módszere. Az ML egyaránt alkalmas a modellilleszkedés tesztelésére és az illeszkedés jóságának a kiértékelésére. Az ML egyik hátránya, hogy alkalmazási feltétele a manifeszt változók többdimenziós normális eloszlása, ami a gyakorlatban gyakran nem teljesül. Ha ez a feltételezés sérül, akkor a modelleredmények nem feltétlenül megbízhatóak. A torzítás a becsült paraméterek standard hibáiban és a modelltesztelésre használt χ^2-statisztikában jelenik meg (Li, 2014; Mîndrilă, 2010; Schermelleh-Engel et al., 2003). A normalitás sérülését Curran és munkatársai (1996), illetve Mîndrilă (2010) ajánlása nyomán akkor érdemes komolyan venni, ha a ferdeségi együttható 2-nél, a csúcsossági együttható pedig 7-nél nagyobb abszolút értékű.

A normalitás súlyos sérülése esetén szokták a regressziós elemzésekből is ismert legkisebb négyzetek módszerét alkalmazni súlyozott (WLS) vagy súlyozatlan (ULS) formában, melynek során olyan modellt keresünk, amelyből a kiszámított korrelációs vagy kovarianciamátrix a lehető legkisebb mértékben tér el az adatmintából kiszámított korrelációs vagy kovarianciamátrixtól. A DWLS a WLS módszer olyan változata, amelynél Pearson- helyett polikorikus korrelációt *(polichoric correlation)* számítunk. Itt két kategoriális változó esetén a korrelációt azon feltételezés mellett számítjuk ki, hogy a diszkrét értékek két valójában folytonos normális eloszlású változó kerekített értékei, a polikorikus korreláció pedig e két feltételezett – és ennek megfelelően reprodukált – normális eloszlású változó között számított Pearson-korreláció (Lee et al., 1995).

Az ML, ULS, DWLS alapmódszerekre építve ROP-R-ben öt robusztus becslési variáns érhető el, melyek jól lefedik a CFA- és a SEM-elemzésekben használt módszerek széles spektrumát (Muthén, 1994; Li, 2021; Kyriazos & Poga-Kyriazou, 2023):

- MLMV: ML-becslés robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával. Ez folytonosnak tekinthető, de erősen nem normális eloszlású változók esetén jó alternatívája az ML-nek (vö. Gao et al., 2019; Vargha et al., 2020).
- MLR: ML-becslés Huber-White robusztus standard hibákkal, valamint egy Yuan-Bentler-féle statisztikával aszimptotikusan megegyező tesztstatisztikával. Ez a módszer a normalitás enyhe vagy közepes mértékű sérülése esetén ajánlható az ML helyett folytonos változók esetén (Li, 2016).
- ULSMV: az ULS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával. Ez a módszer kategoriális változók esetén jó alternatívája lehet az ML-nek (Savalei & Rhemtulla, 2013).
- WLSM: a DWLS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot alkalmazó próbastatisztikával. Ez a módszer kategoriális változók esetén szintén jó alternatívája lehet az ML-nek (Savalei & Rhemtulla, 2013), bár nem minden esetben (Navruz, 2016).
- WLSMV: a DWLS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával. Mint a DWLS robusztus változata, kategoriális változók esetén ajánlható, de számos más nem normális eloszlástípus esetén is jobb becslésnek tűnik, mint az MLR (Li, 2016; Holtmann et al., 2016).

A WLSM és a WLSMV összehasonlítása tekintetében Mplus futási tapasztalatok alapján többen megjegyzik,[17] hogy azonos feladatban a WLSM illeszkedése általában gyengébb, mint a WLSMV-é a χ^2-statisztika és az RMSEA, de jobb a CFI és a TLI tekintetében. Muthén és munkatársai (1997) szerint elméletileg a WLSMV-t érdemes preferálni a WLSM-mel szemben.

A CFA futtatása során ROP-R-ben mindig megkapjuk a kiválasztott robusztus módszer mellett a megfelelő ML, ULS vagy DWLS alapbecsléshez tartozó eredményeket is. Nehéz pontos eligazítást adni ahhoz, hogy konkrét esetben melyik módszert válaszszuk a CFA-ban. A kiindulás az lehet, hogy normális eloszlású tételek esetén ML alapú, folytonos nem normális eloszlású tételek esetén ULS alapú, kategoriális (kétértékű vagy ordinális) változók esetén DWLS alapú becslés ajánlott. Jelenleg még hiányoznak az alapos kutatási eredmények az összes lehetséges módszer előnyeiről és hátrányairól. Érdemes ezért adott esetben több módszert is kipróbálni. Ami a legfontosabb: olyan modellt fogadjunk el, amely a kiértékelés során a lehető legjobb illeszkedést mutatja (vö. *1. táblázat*), és szakmailag is értelmes.

A módszerválasztásban szerepet játszhat a hiányzó adatok kezelése is. Ugyanis az MLR kivételével minden módszer csak a komplett, minden változóra érvényes értékkel rendelkező eseteken végez elemzést, az MLR viszont a FIML *(Full Information Maximum Likelihood)* adatkezelési módszer alapján. A FIML módszer az összes rendelkezésre álló információt felhasználja a modell becsléséhez (a nem teljes adatsor-

[17] Lásd pl.: http://www.statmodel.com/discussion/messages/23/121.html?1469548078

ral rendelkező eseteket is figyelembe véve), és olyan faktormodellt konstruál, amely mellett a legnagyobb az adott minta bekövetkezésének a valószínűsége (Collins et al., 2001). Ez esetben is jelen van az alkalmazási feltételek inkonzisztenciája. Miközben az MLR robusztus módszert a normalitás sérülése esetén ajánlják robusztus módszerként, a FIML alkalmazásának fontos feltétele, hogy a változók többdimenziós normális eloszlást kövessenek.

Végezetül a faktormodell felépítésében szerepet kapnak még a modifikációs indexek *(modification indices)* is, amelyeket az alábbiakban tekintünk át.

Modifikációs indexküszöbök

Alapértelmezésben a faktorindexek segítségével definiált faktorstruktúra látens faktorai korrelálhatnak egymással, de az egyes faktorokat alkotó tételek reziduálisai (a faktor által meg nem magyarázott részei) nem – sem az azonos faktorba tartozó tételek, sem a különböző faktorokba tartozó tételek esetén. Ugyanígy alapértelmezésben nincs megengedve a kereszttöltés sem, vagyis az, hogy egy tétel korreláljon valamely nem saját faktorral. Ez a legegyszerűbb faktorstruktúra azonban nem mindig állja meg a helyét. Néha meg kell engedni, hogy a fenti kapcsolatokat képviselő kovarianciák 0-tól különböző értéket is felvehessenek. Erre utaló jelzést a modifikációs indexek magas értékei adhatnak az elsődleges futás eredménylistáján.

Modifikációs indexe a CFA-modell olyan paramétereinek van, amelyekre valamilyen korlátozást teszünk. Például az egy faktorba tartozó tételek reziduálisaira kikötjük, hogy nem korrelálnak (azaz a kovarianciájuk 0). Ugyanez a helyzet a különböző faktorokba tartozó tételekkel, valamint a kereszttöltésekkel is. A modifikációs index megadja, hogy milyen mértékben javul a modell (pontosabban milyen mértékben csökken a modellilleszkedést mérő χ^2-érték), ha egy-egy ilyen korlátozást megszüntetünk, például, ha megengedjük egy kovarianciára, hogy 0-tól különbözzön (MacCallum et al., 1992). A CFA-modul menüablakában ennek alapján beállítható, hogy a faktorokon belüli vagy különböző faktorokba tartozó tételek közötti, illetve a tételek és a nem saját faktorok közötti kovarianciák milyen modifikációs küszöb felett legyenek beépítve egy javított faktormodellbe. Mivel egyetlen paraméter korlátozásának feloldásával a χ^2-érték változása 1 szabadságfokú χ^2-eloszlást követ, egy 3,841 feletti érték már 5%-os szinten szignifikáns. Ugyanakkor azt is meg kell vizsgálni, hogy a legnagyobb modifikációs indexértékek kiemelkednek-e a többi közül. Ha ilyen értékek vannak, először csak 1-2 kovariancia bevonásával érdemes próbálkozni, és azt is mérlegelni kell, hogy a szóban forgó tételek szakmai tartalma, jelentése feljogosít-e erre a bevonásra. A gyakorlatban – saját tapasztalataim alapján – ritkán érdemes 20 alatti modifikációs értékű kovarianciát bevonni a modellbe.

A CFA-modul menüablakában meg lehet választani, hogy az elkészített faktordiagramon (path-diagramon) a tételek függőleges vagy vízszintes elhelyezésűek legyenek-e. A „Faktorindexekkel definiált skálák elmentése” opció bejelölésével lehet kérni, hogy a ROP-R létrehozza, és az adott adatállományhoz illessze a definiált skálákat. A skálák az alkotótételek átlagaként jönnek létre, de ehhez szükséges, hogy az esetlegesen előforduló fordított tételeket még a CFA-elemzés előtt átfordítsuk. Végül a menüablakban a „Másodrendű CFA” opció bejelölése esetén nem sima elsőrendű, hanem másodrendű CFA-elemzést végez a program (feltéve, hogy a definiált faktorok száma legalább kettő). Ez esetben a ROP-R elkészíti a bifaktoros CFA R-beli futtatására alkalmas scriptet is.[18] Ha emellett a bejelölt „Másodrendű CFA" opció alatt a „Bifaktoros CFA végrehajtása" opciót is bejelöljük, akkor ROP-R maga is lefuttatja a létrehozott bifaktoros scriptet. Ha másodrendű CFA-t kérünk, akkor két kijelölt faktor esetén ROP-R 1-ben rögzíti a specifikus faktorok kezdeti varianciáját a több iterációs lépésből álló modellillesztés indulásakor, hogy a CFA érvényesen végrehajtható legyen.

A CFA technikai részleteivel kapcsolatban megjegyezzük, hogy a ROP-R CFA-modulja a *lavaan* (Rosseel, 2012) és a *lavaanPlot* (Lishinski, 2021) R-package-re épül, mégpedig az Mplus szoftver (Muthén & Muthén, 1998–2017) CFA-elemzéséhez igazított beállítással. A CFA eredménylistájának értelmezését konkrét pszichológiai kérdések kapcsán szemléltetjük. Az alábbi fejezetben egy valódi pszichológiai kutatás adatain mutatjuk be, hogy a CFA segítségével hogyan lehet különböző faktormodelleket összehasonlítani, és megtalálni az adatokra legjobban illeszkedő CFA-modellt.

A Globális Jóllét Kérdőív faktorstruktúrája

Első szakmai problémaként megvizsgáljuk, hogy a korábban már említett Globális Jóllét Kérdőív (Vargha et al., 2020) tételei, amelyek feltételezés szerint mind a globális jóllét konstruktumának mérői, milyen faktorstruktúrára illeszkednek. A kérdőív 17 tétele az *1. ábrá*n látható. Nevükben az első két betű (GJ) közös konstruktumuk, a globális jóllét jelzője, ezt követi a tétel sorszáma, majd a név végén látható rövidítés jelzi, hogy a tétel a fő konstruktum melyik skálájába tartozik (É = Érzelmi, P = Pszichológiai, Sz = Szociális, Sp = Spirituális). Például GJ03É, a kérdőív 3. tétele az Érzelmi jóllét skála egyik iteme. A kérdőívet, mely a jóllét bio-pszicho-szocio-spirituo modelljét operacionalizálja, Oláh Attila hozta létre (Oláh, 2021).

A kérdőív faktorstruktúráját a korábban már említett magyarországi boldogságkutatás 1003 fős (26% férfi, 74% nő) adatállományán elemezzük. Az adatfelvétel online formában történt 2022. január 1. és március 15. között, s a kérdőív tételei között nem voltak hiányzó adatok. Az outliereket a ROP-R-ben az FKA-modul „Extrém esetek azonosítása" opciójával azonosítottuk, mely a GMD általánosított Mahalanobis-távolság alapján 3 szélsőséges esetre hívta fel a figyelmet. Ezeket az outliereket az adatállományból töröltük, így minden további elemzést a maradék 1000 fős adatmintán hajtottunk végre. Ezen a mintán a férfi–nő arány 26–74%, a 17–25, 26–35, 36–50, 51–88 életkori övezetekbe tartozó személyek aránya pedig rendre 22, 23, 50, illetve 4% volt, 36,3 év életkori átlaggal (szórás = 11,3). A személyek rendre 6, 45, illetve 49%-a rendelkezett általános, középilletve felsőfokú végzettséggel.

A 6 fokú Likert-skálájú tételek normalitásának ellenőrzésére tételenként kiszámítottuk a ferdeségi és a csúcsossági együtthatót.[19] Ezek abszolút értékének maximuma 0,67, illetve 1,12 volt, mely szerint a normalitás nem sérült komoly mértékben (Curran et al., 1996; Mîndrilă, 2010).

---

18 Részletesen lásd egy későbbi fejezetben.

19 A ROPstat *Részletesebb mintastatisztikák* modulja segítségével.

A CFA segítségével szeretnénk megerősítést nyerni egyrészt arról, hogy az elméleti megfontolások alapján kialakított 4 skála a 17 tétel szerkezetileg stabil csoportosítása, másrészt arról, hogy létezik egy közös faktor, mely a 17 tételt összeköti. Vagy közvetett módon, egy másodrendű faktoron keresztül, vagy egy bifaktoros struktúrában közvetlenül.

Elsőrendű, négyfaktoros CFA

A 4 skála szerkezeti stabilitásának igazolásához elsőként egy elsőrendű CFA-t hajtunk végre négy faktorral, melyeket a faktorindexek segítségével az *5. ábr*án látható módon jelöltünk ki a ROP-R CFA-moduljának menüablakában. Célszerű a változókat skálánként külön-külön átküldeni, és egy tömbben (a jobb egérbillentyű segítségével) hozzájuk rendelni a megfelelő faktorindexet (1, 2 stb.).

Mivel a változók normalitásának sérülésére nem kaptunk adatokat, a 6 fokú tételeket folytonos változóknak tekintve az MLMV- és az MLR-módszert választjuk modellillesztésre, elsőként az alapértelmezett MLMV-t. A CFA-t a Futtat gombra kattintva indíthatjuk.

Indítás után a program elsőként rákérdez a kijelölt faktorok nevére. Ha az egyazon skálához/faktorhoz tartozó tételek neve tartalmaz közös részt (pl. az Érzelmi jóllét skála tételeinek nevében közös a név eleji GJ és a név végi É), akkor a program ezt a közös részt kínálja fel faktornévnek. A mi mintánk esetében az ily módon felajánlott faktornevek: GJÉ, GJP, GJSz és GJSp. Ha a nevekben nincs közös rész, akkor az ajánlott nevek: Faktor1, Faktor2 stb. Persze nem kötelező ezeket elfogadni, lehetséges bármilyen alfanumerikus,[20] maximum 8 karakter hosszú faktornév megadása.

Az eredménylistán az alapstatisztikák és a skálabesorolások listái után a kovarianciákra vonatkozó legnagyobb modifikációs indexek listáját láthatjuk (lásd közülük a 10 legnagyobbat a *2. táblázat*ban). Például a *2. táblázat* első sora arról tájékoztat, hogy ha megengedjük, hogy az egyazon (Érzelmi jóllét) skálába tartozó GJ01É (Alapvetően én egy boldog ember vagyok) és GJ02É (Az én mindennapjaimban legalább háromszor több az öröm, mint a bánat) tétel reziduálisa korreláljon egymással, akkor a modellt tesztelő χ^2-statisztika értéke 94,89-cel csökken, vagyis ilyen mértékben javul az illeszkedés. Ugyanitt a táblázat második sora arról tájékoztat, hogy ha megengedjük, hogy a Pszichológiai jóllét skálához tartozó GJ15P tétel (Szinte minden téren elégedett vagyok magammal) korreláljon az Érzelmi jóllét (GJÉ) skálát képviselő idegen látens faktorral, tehát engedélyezve ezt a kereszttöltést a faktorsúly-táblázatban, akkor a modellt tesztelő χ^2-statisztika értéke 93,55-tel csökken. Ezek alapján dönthetünk arról, hogy milyen kovarianciákat vonjunk be a modellbe.

[20] Betűvel kezdődő, de akár számot is tartalmazó.

2. táblázat. A kovarianciákra vonatkozó 10 legnagyobb modifikációs index

Változó1	Változó2	Modifikációs index	Kovariancia típusa
GJ01É	GJ02É	94,89	itemek egyazon faktorban
GJÉ	GJ15P	93,55	kapcsolat egy idegen faktorral
GJ05P	GJ07P	75,78	itemek egyazon faktorban
GJP	GJ08É	68,94	kapcsolat egy idegen faktorral
GJ11Sz	GJ13Sp	64,10	itemek különböző faktorokban
GJP	GJ11Sz	62,90	kapcsolat egy idegen faktorral
GJ09Sz	GJ12Sz	58,91	itemek egyazon faktorban
GJÉ	GJ11Sz	58,05	kapcsolat egy idegen faktorral
GJ05P	GJ15P	54,27	itemek egyazon faktorban
GJÉ	GJ07P	53,02	kapcsolat egy idegen faktorral

A kovarianciák modellbe való bevonásával érdemes csínján bánni, mert a túlzott engedékenység a szakmai értelmezhetőség rovására mehet. Meggyőzőbb a CFA eredménye, ha minimális számú bevont kovarianciával kapunk elfogadható illeszkedést. A jelen esetben jelentésük alapján a reziduálisok korrelálását GJ01É és GJ02É esetében még megengedhetjük, de GJ05P (Az életem tele van értelmes célokkal) és GJ07P (Évről évre folyamatosan fejlődöm szinte minden téren) esetében (lásd 3. sor) már kevésbé. A modell e szerint történő javításához a menüablak jobb alsó paneljén a modifikációs küszöböket kell megfelelően beállítani. Jelen esetben az egyetlen GJ01É-GJ02É kovariancia megengedéséhez a legfelső (faktorokon belüli) küszöb értékét csökkentjük 200-ról 80-ra, a másik kettőt pedig változatlanul hagyjuk. Ezután ismét futtatjuk a CFA-t, melynek eredménylistáján egyaránt megtaláljuk az eredeti és a javított modellre vonatkozó eredményeket. Ezek közül most csak a fontosabbakat emeljük ki.

A modifikációs indexek listájáról fentebb már szóltunk. Ez a javítás után annyiban módosul, hogy a listában a ROP-R „+” jellel meg is jelöli a javított modellbe bevont kovarianciák modifikációs indexét. E lista után következik a modellillesztés χ^2-próbás tesztelésére vonatkozó eredmények táblázata (lásd *3. táblázat*). Ebben megtaláljuk a modellt tesztelő χ^2-statisztika (Khi^2) értékét, annak szabadságfokát *(f),* valamint a szignifikancia mértékét jelző p-értéket. A szabadságfok alatti sor a χ^2/f takarékossági index értékét tartalmazza (vö. *1. táblázat*).

A táblázat négy oszlopa megfelel az eredeti modell sima ML- és robusztus MLMV-módszerrel, valamint a bevont kovarianciával javított ML- (jele: ML+) és MLMV- (jele: MLMV+) módszerrel történő tesztelésének.

3. táblázat. A kijelölt CFA faktormodell tesztelése

Mutató	**ML**	**MLMV**	**ML+**	**MLMV+**
Khi^2	948,08	599,05	857,31	542,08
f	113	113	112	112
χ^2/f	8,39	5,30	7,65	4,84
p-érték	< 0,001	< 0,001	< 0,001	< 0,001

Jelen esetben azt láthatjuk, hogy a χ^2-érték csökken ugyan a modell javításakor (pl. ML esetén 948,08-ról 857,31-re), de még mindig olyan magas, hogy a modell $p < 0{,}001$ szinten elvethető. Ez az erős szignifikancia a nagy elemszámnak ($N = 1000$) is köszönhető. Fontos ezért modellünk jóságát más mutatók segítségével is mérlegre tenni. A *3. táblázat*ban a χ^2/f takarékossági index (< 5) az MLMV+ modell esetében már elfogadható szintű illeszkedést jelez, az *1. táblázat* többi mutatójának értékét pedig egy külön táblázat tartalmazza a ROP-R outputján a *3. táblázat*ban feltüntetett négy modell x-módszer kombinációra (lásd *4. táblázat*).

4. táblázat. Illeszkedési mutatók a vizsgált modell-módszer kombinációkra

Mutató	**ML**	**MLMV**	**ML+**	**MLMV+**
AIC	48248,5	48248,5	48159,6	48159,6
BIC	48528,2	48528,2	48444,2	48444,2
RMSEA	0,086	0,066	0,082	0,062
C90	(0,081; 0,091)	(0,060; 0,071)	(0,077; 0,087)	(0,057; 0,067)
pClose	< 0,001	< 0,001	< 0,001	< 0,001
CFI	0,942	0,946	0,949	0,952
TLI	0,931	0,935	0,938	0,942
SRMR	0,034	0,034	0,033	0,033

A *4. táblázat*ból elsőként azt állapíthatjuk meg, hogy a sima ML-nél mindig jobb illeszkedést produkál a robusztus MLMV-módszer, és hogy a GJ01É-GJ02É kovariancia bevonásával mindig érezhetően javulnak az illeszkedési mutatók. Ennek következtében a legjobb mutatók az MLMV+ modellt jellemzik, CFI esetében kiváló, TLI és SRMR esetében jó, RMSEA esetében elfogadható szinten. Az elméleti RMSEA értékre vonatkozó C90 = (0,057; 0,067) konfidencia-intervallum egyértelműen a 0,080-as szint alatt van, ami jó, de a pClose *p*-érték sajnos erős, $p < 0{,}001$ szintű szignifikanciát mutat, vagyis RMSEA szignifikánsan nagyobb, mint az elméleti 0,050-es érték, amit ideális esetben elvárnánk.

Ugyanezt a CFA-elemzést az ML alapú MLR és az ULS alapú ULSMV robusztus módszerrel, mint alternatív lehetőséggel is elvégeztük, mindkét esetben úgy állítva be a modifikációs küszöböket, hogy csak a GJ01É-GJ02É kovarianciával javítsuk a modellt. Az összehasonlítást segíti a kapott eredményekből elkészített *5. táblázat*. Az eredmények nagyon hasonlóak, de SRMR kivételével minden mutató tekintetében kissé gyengébb illeszkedést tapasztalunk, mint MLMV esetében. MLR és ULSMV illeszkedését összevetve nem látunk szisztematikus különbséget.

5. táblázat. A GJ01É-GJ02É kovarianciával javított faktormodell illeszkedési mutatói három robusztus módszer esetén

Mutató	MLMV+	MLR+	ULSMV+
χ^2/f	4,84	5,44	5,13
RMSEA	0,062	0,067	0,064
C90	(0,057; 0,067)	(0,062; 0,071)	(0,059; 0,070)
pClose	< 0,001	< 0,001	< 0,001
CFI	0,952	0,950	0,939
TLI	0,942	0,939	0,926
SRMR	0,033	0,033	0,031

A CFA outputjának következő fontos eleme a standardizált faktorsúlyok és a kommunalitások táblázata. Ez a robusztus MLMV-módszer használata esetén az eredeti (MLMV) és a javított (MLMV+) modellre a *6. táblázat*ban látható. Ebből megállapíthatjuk, hogy a tételek a javított modell minden faktorában igen magas, 0,75 feletti faktorsúly és 0,55 feletti kommunalitás becsléssel jellemezhetők, ami a négyfaktoros struktúra konstrukciós validitását erősíti.

6. táblázat. Standardizált faktorsúlyok és a kommunalitások

Faktor/tétel	MLMV	MLMV	MLMV+	MLMV+
GJÉ	**Faktorsúly**	**Kommunalitás**	**Faktorsúly**	**Kommunalitás**
GJ01É	0,861	0,741	0,841	0,707
GJ02É	0,829	0,688	0,806	0,649
GJ03É	0,872	0,760	0,876	0,767
GJ04É	0,862	0,744	0,862	0,743
GJ08É	0,821	0,675	0,829	0,688
GJP	**Faktorsúly**	**Kommunalitás**	**Faktorsúly**	**Kommunalitás**
GJ05P	0,769	0,591	0,768	0,590
GJ06P	0,763	0,582	0,761	0,579
GJ07P	0,793	0,629	0,792	0,627
GJ15P	0,843	0,711	0,845	0,714

GJSz	Faktorsúly	Kommunalitás	Faktorsúly	Kommunalitás
GJ09Sz	0,788	0,620	0,788	0,620
GJ11Sz	0,829	0,687	0,829	0,687
GJ12Sz	0,763	0,583	0,764	0,583
GJ16Sz	0,846	0,716	0,846	0,716

GJSp	Faktorsúly	Kommunalitás	Faktorsúly	Kommunalitás
GJ10Sp	0,865	0,749	0,865	0,749
GJ13Sp	0,825	0,681	0,825	0,681
GJ14Sp	0,798	0,637	0,798	0,637
GJ17Sp	0,822	0,676	0,822	0,676

Az elsőrendű többfaktoros CFA modellje megengedi, hogy a látens faktorok korreláljanak egymással. Hogy milyen szoros kapcsolat van a faktorok között, arról az output súlymátrix alatti táblázata informál (lásd *7. táblázat*). Ebből megállapíthatjuk, hogy a robusztus MLMV becslést alkalmazva igen erős korrelációkat láthatunk a faktorok között mind az eredeti, mind a javított modell esetén. A leggyengébb korreláció (GJÉ és GJSp között) is 0,77 fölötti, a legnagyobb (GJÉ és GJP között) pedig 0,95 fölötti, nagyon magas érték, ami túlzott redundanciára utal GJÉ és GJP között.

7. táblázat. A látens faktorok páronkénti korreláció becslései

F(i)	F(j)	MLMV	MLMV+
GJÉ	GJP	0,960	0,970
GJÉ	GJSz	0,788	0,792
GJÉ	GJSp	0,776	0,778
GJP	GJSz	0,846	0,846
GJP	GJSp	0,828	0,828
GJSz	GJSp	0,936	0,936

A ROP-R minden CFA-elemzés során elkészíti a vizsgált faktormodell javítás előtti és utáni ábráját pdf-fájlban, a c:_vargha\ropstat\aktualis mappában (a javítás előtti modellhez tartozót pathplot1.pdf, a javított modellhez tartozót pedig – ha van javítás – pathplotR1.pdf névvel). A javítás előtti faktormodell ábrája a jelen elemzés esetén a *2. ábrá*n látható (vízszintes path-diagram irányítás beállításával). Ha a CFA menüablakban feltételes csoportosító változót is használunk (pl. a személy nemét vagy iskolázottsági szintjét), akkor a ROP-R minden csoportra végrehajtja ugyanazt az elemzést, és minden csoporthoz külön diagramok is készülnek (modelljavítás nélkül csak egy). A pdf-fájlok a nevükben feltüntetett csoportindex értéke (1, 2 stb.) alapján különböztethetők meg egymástól.

Összességében megállapíthatjuk, hogy az elsőrendű négyfaktoros modell illeszkedése ígéretes, mely jelzi az elsőrendű négyfaktoros struktúra elfogadhatóságát, de még nem olyan szintű, amit végső megoldásként el tudunk fogadni. Ez okból megvizsgáljuk a másodrendű és a bifaktoros modellt is, hátha valamelyik közülük elfogadható illeszkedésű lesz.

Másodrendű CFA

A ROP-R-ben a CFA-modul menüablakában a „Másodrendű CFA" opciót kell bejelölni ahhoz, hogy sima elsőrendű helyett egy másodrendű CFA-elemzést végezzünk.

Indítás után a program itt is rákérdez a kijelölt faktorok nevére, s ugyanúgy kínál fel nevet az elsőrendű faktoroknak, mint az elsőrendű CFA esetén. Ha ezek a faktornevek tartalmaznak közös részt, akkor a program ezt a közös részt kínálja fel a másodrendű faktornévnek. Ha az elsőrendű faktornevekben nincs közös rész, akkor az ajánlott másodrendű faktornév: F2rend. Persze mód van itt is bármilyen alfanumerikus faktornév megadására.

Másodrendű CFA esetén be lehet vonni kovarianciákat a modellbe, de csak az egyazon faktorba tartozó tételek között. Ugyanazokat a skálakijelöléseket és beállításokat használva, mint az előző alfejezetben az elsőrendű négyfaktoros elemzés során (MLMV-módszer) most is a GJ01É és a GJ02É tétel modifikációs indexe volt ebben a kategóriában a legmagasabb (84,56), s ezt egy 80-as küszöb beállításával vontuk be a modellbe.

A másodrendű CFA ROP-R-beli eredménylistája nagyon hasonló szerkezetű, mint az elsőrendűé. Erről leolvasható volt, hogy a másodrendű javított modell (MLMV+) illeszkedése mindenben rosszabb, mint az elsőrendűé (χ^2/f = 6,51 > 4,84; RMSEA = 0,074 > 0,062; CFI = 0,931 < 0,952; TLI = 0,917 < 0,942; SRMR = 0,048 > 0,033), így ezzel nem sikerült javítani az egyfaktoros modellt.

A másodrendű faktormodell standardizált regressziós együtthatóinak táblázata a tételekre vonatkozóan ugyanúgy néz ki, mint a *6. táblázat,* ahol a faktorsúlyértékek is közel ugyanakkorák (az abszolút eltérés mindenütt kisebb, mint 0,03). Egyetlen eltérés, hogy a táblázat alján megjelennek az elsőrendű faktorok standardizált faktorsúlyai is a másodrendű faktormodellben (lásd *8. táblázat*). Ebből azt láthatjuk, hogy az elsőrendű faktorok nagyon rásimulnak a globális jóllét GJ másodrendű faktorára. Mégpedig oly mértékben, hogy a túlzott redundancia felmerül problémaként.

8. táblázat. Az elsőrendű faktorok standardizált faktorsúlyai a másodrendű faktormodellben

Faktor	MLMV	MLMV+
GJÉ	0,930	0,940
GJP	0,980	0,986
GJSz	0,897	0,891
GJSp	0,880	0,873

Az eredménylistát a modifikációs indexek listája zárja a javított másodrendű faktormodellben, ahonnan a tíz legnagyobb modifikációs indexet a *9. táblázat*ban foglaltuk össze. Ez arról informál, hogy milyen mértékben javulna a modell, ha bizonyos korrelációkat megengednénk a modell bizonyos komponensei között. A legnagyobb javulást jelen esetben az eredményezné, ha megengednénk egyes elsőrendű faktorok (GJÉ és GJP, illetve GJSz és GJSp) reziduálisainak a korrelációját, ami ismét a faktorok túlzott redundanciájáról árulkodik.

9. táblázat. A 10 legnagyobb modifikációs index a javított másodrendű faktormodellben

Változó1	**Változó2**	**Modifikációs index**
GJÉ	GJP	317,57
GJSz	GJSp	317,57
GJÉ	GJ15P	213,13
GJ	GJ15P	169,65
GJ11Sz	GJ13Sp	83,85
GJÉ	GJSz	71,63
GJP	GJSp	71,63
GJÉ	GJSp	65,38
GJP	GJSz	65,37
GJ03É	GJ15P	62,97

A ROP-R ez esetben is elkészíti a vizsgált faktormodell javítás előtti és utáni ábráját pdf-fájlban és a c:_vargha\ropstat\aktualis mappában helyezi el őket (a javítás előtti modellhez tartozót pathplot1.pdf, a javított modellhez tartozót pathplotR1.pdf névvel). A javítás előtti faktormodell ábrája a jelen elemzés esetén a *3. ábrá*n látható (függőleges path-diagram irányítás beállításával).

Összességében megállapíthatjuk, hogy a másodrendű faktormodell illeszkedése gyengébb, mint az elsőrendűé, és az elfogadhatóság határán billeg. Jó jel, hogy a tételek itt is magas faktorsúlyokkal illeszkednek az elsőrendű faktorokra, ami a teszt négy skálájának megbízhatóságát erősíti, de a másodrendű faktorok egyöntetűen igen magas, 1-hez közeli faktorsúlyai (lásd *8. táblázat*) azt mutatják, hogy a négy skála nem különül el kellő mértékben egymástól. Különösen erős összefüggés van az érzelmi és a pszichológiai jóllét (GJÉ és GJP), illetve a szociális és a spirituális jóllét (GJSz és GJSp) skálája között (lásd a *9. táblázat* első két sorát). Ez okból megvizsgáljuk a bifaktoros modellt is, hátha az jobb illeszkedést mutat.

Bifaktoros CFA

A ROP-R segítségével bifaktoros CFA-elemzést is végezhetünk, bár kissé komplikáltabb módon, mint a korábban bemutatott egyszerűbb modellek esetén. A lényeg az, hogy ha a CFA-modul menüablakában a „Másodrendű CFA" opciót bejelöljük, akkor a ROP-R amellett, hogy végrehajt egy másodrendű CFA-elemzést a fentebbi alfejezetben leírt módon, elkészíti és fájlba menti a bifaktoros CFA-hoz szükséges R-scriptet is a c:_vargha\ropstat\aktualis mappában, CFAbif.r, illetve CFAbifR.r néven. Utóbbi az egyazon faktorba tartozó egyes tételek reziduálisai közötti korrelációkat is megengedő javított modell CFA-elemzését teszi lehetővé, és csak akkor jön létre, ha van olyan modifikációs index, amely a kijelölt küszöböt meghaladja. Ha az R szoftver RGui keretprogramját elindítjuk, és ezeket a scripteket (a fájlokat pl. a Windows jegyzettömb [notepad] programja segítségével megnyitva) simán bemásoljuk az RGui konzoljába, majd Enterrel futtatjuk, akkor a bifaktoros elemzés eredménylistája a c:_vargha\ropstat\aktualis mappában, a CFAbif1.txt, illetve a CFAbifR1.txt[21] szövegfájlban tekinthető meg, a modellhez tartozó faktordiagram pedig a bifplot1.pdf, illetve a bifplotR1.pdf nevű pdf-fájlban (utóbbi a javított modellre vonatkozó faktorábrát tartalmazza). Megjegyezzük, hogy a ROP-R legfrissebb változata már úgy működik, hogy ha a bejelölt „Másodrendű CFA" opció alatt a „Bifaktoros CFA végrehajtása" opciót is bejelöljük, akkor ROP-R maga is lefuttatja a létrehozott bifaktoros scriptet.

Szemléltetésül az előző alfejezetben is használt boldogságkutatás adatait elemezzük, ugyanazokkal a beállításokkal, mint a másodrendű CFA esetén. A GJ01É és a GJ02É tétel reziduálisának kovarianciáját is megengedve ROP-R a *10. táblázat*ban látható R-scriptet készíti el, melyet az RGui konzoljába bemásolva az Enter billentyű lenyomásával futtatunk.

10. táblázat. A bifaktoros CFA R-scriptje

```
library(lavaan)
library(lavaanPlot)
dat<-read.delim(‚c:\_vargha\ropstat\aktualis/tmpdat1.txt', header=T)
model <- ‚
# Specific factors
GJÉ=~ GJ01É+GJ02É+GJ03É+GJ04É+GJ08É
GJP=~ GJ05P+GJ06P+GJ07P+GJ15P
GJSz=~ GJ09Sz+GJ11Sz+GJ12Sz+GJ16Sz
GJSp=~ GJ10Sp+GJ13Sp+GJ14Sp+GJ17Sp
# General factor
ÁltFak=~ GJ01É+GJ02É+GJ03É+GJ04É+GJ05P+GJ06P+GJ07P+
GJ08É+GJ09Sz+GJ10Sp+GJ11Sz+GJ12Sz+GJ13Sp+GJ14Sp+
GJ15P+GJ16Sz+GJ17Sp
# Set zero correlations between the general factor and each specific factor
ÁltFak~~ 0*GJÉ
ÁltFak~~ 0*GJP
ÁltFak~~ 0*GJSz
ÁltFak~~ 0*GJSp
# Set zero correlations among specific factors
GJÉ~~ 0*GJP      # omit „0*" for allowing correlating factors
GJÉ~~ 0*GJSz     # omit „0*" for allowing correlating factors
GJÉ~~ 0*GJSp     # omit „0*" for allowing correlating factors
GJP~~ 0*GJSz     # omit „0*" for allowing correlating factors
GJP~~ 0*GJSp     # omit „0*" for allowing correlating factors
GJSz~~ 0*GJSp    # omit „0*" for allowing correlating factors
```

[21] E nevekben az 1-es szám a csoportindex. Ha feltételes csoportosító változót jelölünk ki, akkor minden csoport kap egy-egy indexet, mely megjelenik ezekben a fájlnevekben. Például három csoport esetén a 2. csoporthoz létrejön a CFAbif2.txt és a CFAbifR2.txt, a 3. csoporthoz pedig a CFAbif3.txt és a CFAbifR3.txt fájl.

```
#Residual covariances
GJ01É ~~ GJ02É
,
sink(‚c:\_vargha\ropstat\aktualis/CFAbifR1.txt')
fit <- cfa(model, data=dat, likelihood="wishart", mimic="Mplus", estimator="MLMV")
summary(fit, fit.measures = T, rsq = T, standardized = T)
modindices(fit, sort = TRUE, maximum.number = 50)
pl <- lavaanPlot(model = fit, graph_options = list(rankdir = „LR", fontsize = „10"),
node_options = list(shape = „box", fontname = „Helvetica"),
edge_options = list(color="black"), coefs=T, stand=T, covs=T, digits=2)
embed_plot_pdf(pl,'c:\_vargha\ropstat\aktualis/bifplotR1.pdf')
sink()
```

Erre a javított modellre vonatkozóan az eredmények a CFAbifR1.txt szövegfájlba íródtak. Az első, amit érdemes megnézni, a modell tesztelése, mely a „Model Test User Model" című rovatnév után található mind a standard ML-, mind a robusztus MLMV-módszerre vonatkozóan. A χ^2-statisztika értéke a „Test Statistic" sorban olvasható (ML esetén 916,904, MLMV esetén 584,804). E sor alatt láthatók a szabadságfokok (mindkét módszer esetén 101), amelyek segítségével a takarékossági index is kiszámítható. Például MLMV esetén

$$\chi^2/f = 584{,}804/101 = 5{,}79,$$

ami picit jobb, mint a másodrendű javított modellé (6,51), de még mindig határozottan gyengébb, mint az elsőrendű négyfaktoros javított modellé (4,84).

Ugyanezt a viszonyt láthatjuk a CFI (0,947), a TLI (0,928), az RMSEA (0,069), az SRMR (0,049) illeszkedési mutatók esetén is, melyek értékei értelemszerűen rendre a „Comparative Fit Index (CFI)", a „Tucker-Lewis Index (TLI)", az „RMSEA", illetve az „SRMR" kezdetű sorokból olvashatók ki.

Igen fontos a „Latent Variables" rovatnév után található táblázat, mely a bifaktoros modell paramétereinek a regressziós becsléseit tartalmazza. E táblázatnak a specifikus faktorokhoz tartozó részét a *11. táblázat*ban mutatjuk be. Ennek fejlécében „Estimate" a nyers becslés, „Std.Err" a standard hiba, „z-value" a becslés szignifikanciáját tesztelő z-érték, „P(>|z|)" e teszteléshez tartozó p-érték, „Std.lv" a becslés standardizált látens faktorok esetén, „Std.all" pedig a becslés, ha minden változót standardizálunk, vagyis a standardizált faktorsúlybecslés.

11. táblázat. A javított bifaktoros faktormodell specifikus faktorokhoz tartozó paramétereinek regressziós becslései (a fejléc rövidítéseinek jelentését lásd a szövegben)

Variable	**Estimate**	**Std.Err**	**z-value**	**P(>|z|)**	**Std.lv**	**Std.all**
GJÉ	=~					
GJ01É	1				0,348	0,243
GJ02É	1,195	0,160	7,456	0,000	0,415	0,268
GJ03É	1,268	0,218	5,827	0,000	0,441	0,304
GJ04É	1,452	0,243	5,979	0,000	0,505	0,359
GJ08É	0,272	0,149	1,832	0,067	0,095	0,063

Variable	Estimate	Std.Err	z-value	P(>|z|)	Std.lv	Std.all
GJP	=~					
GJ05P	1				0,513	0,347
GJ06P	0,631	0,114	5,552	0,000	0,324	0,255
GJ07P	0,807	0,138	5,836	0,000	0,414	0,302
GJ15P	-0,543	0,164	-3,320	0,001	-0,279	-0,189
GJSz	=~					
GJ09Sz	1				0,602	0,387
GJ11Sz	1,443	0,149	9,679	0,000	0,868	0,536
GJ12Sz	0,890	0,084	10,643	0,000	0,536	0,366
GJ16Sz	0,865	0,074	11,662	0,000	0,520	0,345
GJSp	=~					
GJ10Sp	1				0,762	0,478
GJ13Sp	0,911	0,076	11,942	0,000	0,694	0,454
GJ14Sp	1,007	0,070	14,377	0,000	0,767	0,476
GJ17Sp	0,724	0,061	11,792	0,000	0,551	0,345

A *11. táblázat* arról informál, hogy milyen az illeszkedése a tételeknek a specifikus faktorokra. Az illeszkedés akkor jó, ha a standardizált faktorsúlyok (lásd az utolsó, Std. all című oszlopot) mind nagyobbak 0,50-nél. Ez távolról sem teljesül, mert csak egyetlen faktorsúly nagyobb 0,50-nél (a GJ11Sz tételé). A minimális elvárás, hogy a faktorsúlyok mind pozitívak legyenek (ha nincs átfordítandó tétel), és szignifikánsan különbözzenek a 0-tól. A jelen esetben ez a minimális elvárás sem teljesül egy negatív (GJ15P esetén -0,189) és egy nem szignifikáns (GJ08É esetén p = 0,067) faktorsúly miatt.

A „Latent Variables" rovatnév után található táblázat alsó része a 17 teszttételnek az általános faktorra (neve az outputon: ÁltFak) való illeszkedését mutatja. Tekintve, hogy a standardizált faktorsúlyok minden tétel esetén nagyobbak 0,65-nél, az illeszkedés jónak mondható. Ezt is figyelembe véve a kapott eredményeket úgy értelmezhetjük, hogy az általános faktor létezése igazoltnak tűnik, azt viszont nem jelenthetjük ki, hogy a tételek az általuk tartalmazott ezen közös konstruktumon túl rendelkeznének még olyan specifikus információval is, ami létjogosultságot adna a négy tesztskálának megfelelő négy specifikus faktornak.

A bifaktoros modellt valószínűsíti jelen esetben, hogy a teszttételek FKA-elemzése során az első főkomponens sajátértéke (10,54) jóval több mint a háromszorosa az utána következőnek (1,33), viszont ellene szól, hogy az EFA-elemzés (ML-módszer, oblimin ferde forgatás) olyan kétfaktoros struktúrára vezet, amelynek az első faktorát az érzelmi és a pszichológiai jóllét tételei, második faktorát pedig a szociális és a spirituális jóllét tételei feszítik ki, mind 0,55 fölötti faktorsúlyokkal. A négy skálát tehát az EFA sem tudja alátámasztani.

Az R segítségével végrehajtott bifaktoros CFA eredménylistáján még megemlítendő az output végén található modifikációs indexek listája, melynek fejléce az „lhs op rhs" szöveggel kezdődik. Itt a modifikációs indexek (jele: mi) az első adatoszlopban láthatók. Konkrét példánk esetében ez a GJSz és a GJSp közti kovarianciára vonatkozik, és értéke 354,887, mely szintén azt erősíti meg, hogy ez a két specifikus faktor olyan szoros kapcsolatban van, hogy a modellben meg kell engednünk, hogy korreláljanak egymással. Ehhez mindössze annyit kell tennünk, hogy a *10. táblázat*ban látható script „GJSz ~~ 0*GJSp" utasításában (ez most a „#Residual covariances" sor fölött található) töröljük a GJSp faktornév előtti „0*" részt, mely a két faktor közötti 0 korrelációt fixálja. Ha ezt az elemzést is elvégezzük, akkor az illeszkedési mutatók tekintetében kiváló faktormodellt láthatunk (RMSEA = 0,046, C90 = [0,041; 0,052], pClose = 0,851, CFI = 0,976, TLI = 0,968, SRMR = 0,023). Ennek ellenére továbbra is negatív (-0,236) lesz GJ15P faktorsúlya a GJP specifikus faktorban, és nem szignifikáns a GJ08É tételé a GJÉ specifikus faktorban (p = 0,643). És ez a helyzet akkor sem változik, ha még a GJÉ és a GJP specifikus faktor közti kovarianciát is beépítjük a modellbe.

A Globális Jóllét Kérdőívvel kapcsolatban végrehajtott CFA-elemzések eredményeit összegezve azt mondhatjuk, hogy a teszt 17 tétele kellően homogén, jól illeszkedik egy közös konstruktumra, ami hitelesíti, hogy a 17 tétel átlagával vagy akár a négy skála átlagával mérjük ezt a globális jóllétnek nevezhető konstruktumot mint látens faktort. Ami a kérdőívben problémás: a definiált négy skála nem különül el jól egymástól (legkevésbé az érzelmi és a pszichológiai, illetve a szociális és a spirituális jóllét skálája). Ezt a másodrendű CFA-ban az elsőrendű faktorok nagyon magas, 1-hez közeli faktorsúlyai mutatják a másodrendű faktorra való illeszkedés során, a bifaktoros CFA-ban pedig a tételek alacsony (1-1 esetben nem szignifikáns, illetve negatív) faktorsúlyai a specifikus faktorokra való illeszkedés során. Ezen a helyzeten nem segít, ha a legnagyobb elsőrendű faktorok közti kovarianciákat (GJSz és GJSp, illetve GJÉ és GJP között) beépítjük a modellbe. Tanulságos, hogy e kovarianciák beépítése például a bifaktoros CFA esetén formálisan jó illeszkedésű modellt jelez, a faktorsúlyok egy része azonban nem felel meg az elfogadhatóság kritériumának, vagyis a faktorstruktúra így, ahogy van, nem jó. Mi lehet a megoldás?

Szakmailag elfogadható az a másodrendű faktormodell, melyben a kérdőív 17 tétele négy látens faktor közvetítésével kapcsolódik a globális jóllét másodrendű konstruktumához. De a GJSz és GJSp, illetve a GJÉ és GJP közötti igen erős összefüggés arra utal, hogy ez a másodrendű konstruktum nem simán egydimenziós, hanem maga is két pilléren nyugszik, az érzelmi és pszichológiai összetevőből építkező belső (internális), illetve a társas és spirituális összetevőből építkező külső (externális) jólléten. Ennek tesztelését a bifaktoros CFA logikáját követve olyan R-script segítségével végezhetjük el, melyet a ROP-R-ben végrehajtott másodrendű elemzés javított modellje során kapunk (a GJ01É és a GJ02É közti kovariancia beépítésével). A ROP-R ezt a c:_vargha\ropstat\aktualis mappában menti el, egy CFA2.r nevű fájlban. Ez a script nagyon hasonló a bifaktoros CFA *10. táblázat*ban látható scriptjéhez, s ennek egy

GJ=~ GJÉ+GJP+GJSz+GJSp

sora definiálja az egydimenziós másodrendű faktort. Ha ezt úgy módosítjuk, hogy a fenti egy helyett két másodrendű faktort definiálunk a

GJint=~ GJÉ+GJP
GJext=~ GJSz+GJSp

utasításokkal, akkor máris tesztelhetjük a kétfaktoros másodrendű modellt. Az eredmények ez esetben egy o2.txt fájlban találhatók (ugyanabban a mappában, mint CFA2.r), mely nagyon hasonló szerkezetű a bifaktoros CFA esetében kapott eredményfájlhoz. Ebből kiolvasható, hogy a robusztus MLMV becslési módszerre vonatkozóan RMSEA = 0,054, C90 = [0,049; 0,059], pClose = 0,104, CFI = 0,964, TLI = 0,956, SRMR = 0,031, ami már kiváló illeszkedésre utal. Egyetlen probléma, hogy a GJ15P (Szinte minden téren elégedett vagyok magammal) tétel negatív (-0,788) súllyal illeszkedik saját GJP faktorához. Ha most ezt a rosszul viselkedő tételt kihagyjuk az inputváltozók közül (3 tételesre csökkentve ezzel a pszichológiai jóllét skáláját), akkor egy minden tekintetben elfogadható illeszkedéshez jutunk, melyet az alábbi eredmények támasztanak alá:

1. Jó illeszkedési mutatók: χ^2/f = 3,95, RMSEA = 0,054, C90 = [0,049; 0,060], pClose = 0,100, CFI = 0,966, TLI = 0,958, SRMR = 0,030.
2. Minden teszttétel 0,76 feletti pozitív faktorsúllyal illeszkedik saját faktorára.
3. GJÉ és GJP 0,94 feletti súllyal illeszkedik saját GJint, s ugyanígy GJSz és GJSp 0,95 feletti súllyal saját GJext másodrendű faktorára.
4. A GJint és a GJext másodrendű látens faktor becsült korrelációja 0,862, ami azt jelzi, hogy a két faktor tartalmilag markáns (74,3%-os) átfedést mutat (ez a közös globális jóllét konstrukciója), van azonban olyan mértékű (100 – 74,3 = 25,7%) egyedi részük is, mely jogossá teszi külön skálaként való mérésüket.

Mindezen CFA-elemzések alapján a Globális Jóllét kérdőívének parányi módosításával (a GJ15P tétel elhagyásával) a teszt alkalmasnak tűnik a szokásosan használt négy önálló elsőrendű skála (GJÉ, GJP, GJSz, GJSp) és a globális jóllétet mérő összpontszám (GJ) definiálására. Ha a teszt elméleti konstrukciójával szakmailag összeegyeztethető, akkor szóba jöhet emellett két másodrendű skála (GJint és GJext) definiálása is. Természetesen egy ilyen modell stabilitását szükséges megerősíteni más független adatmintákon végrehajtott másodrendű kétfaktoros CFA-elemzések segítségével is.

Speciális lehetőségek az R szoftver segítségével

A ROP-R nemcsak a bifaktoros CFA futtatásához szükséges scriptet készíti el, de minden futás során létrejön a c:_vargha\ropstat\aktualis mappában a teljes elemzés R-beli futtatását lehetővé tevő script egy vagy több szövegfájlban. Ez egyrészt segítség lehet az R-rel ismerkedőknek a szintaktikailag helyes utasításcsomagok elkészítéséhez, másrészt arra is jó, hogy ennek segítségével a felhasználó néhány olyan opciót is kipróbálhasson, ami a ROP-R menüjéből nem elérhető. Az alábbiakban ehhez adunk egy kis eligazítást. Az első, amit a gördülékeny R-beli futás érdekében fontos megjegyezni, hogy ezeket a scripteket csak a ROP-R-ből – és a ROP-R által készített fájlokból – való kilépés után szabad futtatni. Ha ugyanis például a CFA faktordiagramját tartalmazó pdf-fájl nyitva van, ugyanolyan névvel a script nem tudja elmenteni a létrehozandó pdf-fájlt (ha nincs nyitva, akkor gond nélkül felülírja). Az alábbiakban néhány további fontos kérdést tisztázunk.

Milyen R-scriptek jönnek létre a ROP-R CFA-moduljában?

A bifaktoros CFA R-beli futtatására alkalmas scriptekről fentebb már szóltunk. Ettől függetlenül a CFA-modul futtatásakor mindig létrejön a c:_vargha\ropstat\aktualis mappában egy „CFA.r” nevű szövegfájl, mely a kovarianciák nélküli alapelemzés R-beli futtatásához tartozó scriptet tartalmazza. Ez csak abban különbözik a bifaktoros CFA futtatására alkalmas scripttől (lásd *10. táblázat*), hogy a „model” kezdetű és az első „sink” kezdetű sor közti rész, amely a CFA aktuális modelljét jelöli ki a *lavaan* R-package-ben, kisebb-nagyobb mértékben eltérő lesz.

Ha az első CFA-elemzés során van olyan modifikációs index, amelynek értéke a rá vonatkozó megadott küszöb értékét meghaladja, akkor létrejön a javított CFA-modell R-beli futtatására alkalmas script is, a CFA2.r nevű szövegfájlban.

Milyen adatállományon végeznek CFA-elemzéseket a létrehozott R-scriptek?

A ROP-R minden CFA-elemzés során szövegfájlba menti (a c:_vargha\ropstat\aktualis mappában) a kiválasztott változókra vonatkozó adatokat. Ennek neve tmpdat1.txt, ha nincs kijelölve csoportosító változó a CFA-modul menüablakában. Ha van, akkor ez a fájl az 1. csoport adatait tartalmazza, tmpdat2.txt a 2. csoportét, tmpdat3.txt a 3. csoportét stb. A CFA.r, CFA2.r fájlok mindig az utolsó (legnagyobb indexű) csoportra vonatkozó fájlnevet tartalmazzák, vagyis például négy csoport esetén a script 3. sora mindig így fog kinézni:

dat<-read.delim(‚c:_vargha\ropstat\aktualis/tmpdat4.txt’, header=T)

Ebből következően, ha több csoportunk van, és az R-elemzést nem a legnagyobb indexű csoporttal akarjuk elvégezni, akkor ebben az adatbeolvasó „dat” utasításban a fájlnevet ennek megfelelően módosítani kell. Például négy csoport esetén a 2. csoport futtatásához tmpdat4.txt helyett a tmpdat2.txt fájlnév megadása szükséges.

Milyen módosításokat érdemes végrehajtani a ROP-R által létrehozott R-scripteken?

Négy ilyen példát említünk.
1. Ha valamilyen okból olyan CFA becslési módszert szeretnénk választani, ami nem található meg a ROP-R által felkínált listában (pl. MLMV vagy MLR helyett a szintén ML-típusú robusztus MLM-módszert), akkor az első „sink” kezdetű sor alatt található „fit <- cfa” kezdetű utasításban az „estimator” paraméter értékét ennek megfelelően módosítani kell (pl. estimator=”MLMV” helyett estimator=”MLM” írandó). Fontos, hogy ha egy új ML-típusú módszert akarunk kipróbálni (pl. az MLM- vagy az MLF-módszert), akkor egy szintén ML-típusú (MLMV vagy MLR) módszer futtatása utáni scriptet módosítsunk, ha pedig egy új nem ML-típusú módszert (pl. a GLS-, WLS-, DLS-módszerek valamelyikét), akkor egy szintén nem ML-típusú (ULS vagy DWLS) módszer futtatása utáni scriptet módosítsunk. A különböző lehetséges módszerekkel kapcsolatban lásd https://lavaan.ugent.be/tutorial/est.html.

2. Ha a faktordiagramot nem a robusztus, hanem az alapmódszerre vonatkozóan

szeretnénk elkészíteni, akkor a „fit <- cfa" kezdetű utasításban estimator = „MLMV" helyett írjunk estimator = „ML"-t, estimator = „ULSMV" helyett estimator = „ULS"-t, estimator = „WLSM" helyett estimator = „DWLS"-t stb.

3. ROP-R becsléseinél *lavaan* az Mplus szoftver (Muthén & Muthén, 1998–2017) CFA-elemzésének eredményére legjobban hasonlító algoritmusokat használja. Ha Mplus helyett az EQS szoftvert (Bentler & Wu, 2005) preferálnánk referenciaként, akkor a „fit <- cfa" kezdetű utasításban mimic="Mplus" helyett mimic="EQS" írandó (vö. Rosseel, 2012), ha pedig a *lavaan* saját rendszerét, akkor töröljük az utasításból a mimic="Mplus" szöveget (az utána lévő vesszővel).

4. Ha a modifikációs indexküszöbök segítségével nem lehet úgy megadni a modellbe bevonandó kovarianciákat, ahogy szeretnénk (pl. azért, mert szakmai okokból egy nagyobb modifikációs indexű kovarianciát nem akarunk bevonni a modellbe, míg egy vagy több nála kisebbet igen), akkor manuálisan lehet a scriptben az ehhez szükséges utasításokat megadni a *10. táblázat*ban látható formai szabályokat követve. Egy sorban egy ilyen utasítást helyezzünk el! Ha egyenrangú elemek (pl. tételek) közti kovarianciát akarunk beépíteni, akkor a két érintett elem neve közé írjunk egy „~~" jelsorozatot (lásd pl. a *10. táblázat*ban a „#Residual covariances" sor alatti utasítást), ha pedig azt akarjuk megengedni, hogy egy tétel korreláljon egy idegen faktorral (kereszttöltés megengedése), akkor a sor elejére írjuk be a szóban forgó faktor nevét, utána a „=~" jelsorozatot, végül az érintett tétel nevét (pl. így: GJÉ =~ GJ15P).

Hol találjuk meg az R-script futtatásának eredményét?

Bármely ROP-R-beli CFA-elemzés végrehajtása után létrejövő script fájlban (ilyen pl. CFA.r és CFA2.r) az első „sink" kezdetű sor utasításában van megadva, hogy a script futásához tartozó részletes eredménylista milyen nevű szövegfájlba íródjék. Például a bifaktoros CFA futtatása esetén ez a fájl CFAbifR1.txt (a c:_vargha\ropstat\aktualis mappában; vö. *10. táblázat*). Ha ezen módosítunk, más fájlban lesz elmentve az eredménylista. Ezek minden CFA-elemzés során úgy olvasandók, ahogy azt a bifaktoros CFA R-beli futtatásának ismertetése során leírtuk.

A CFA-val kapcsolatos lehetőségek a jamovi és a JASP programban

Ebben a fejezetben a jamovi (The jamovi project, 2022; Şahin & Aybek, 2019) és a JASP (JASP Team, 2023) ingyenes szoftver lehetőségeit tekintjük át a CFA-val kapcsolatban. Mindenekelőtt megjegyezzük, hogy a ROP-R-ből úgy lehet egy adatállományt a legegyszerűbben áttenni jamoviba és JASP-ba, hogy a ROP-R-ben az aktuális msw-fájlt a *Fájl* menüpontban a Mentés tabulált szövegfájlformátumban opcióval mentjük el, majd ezt a szövegfájlt olvassuk be az említett két programban. A JASP esetén fontos odafigyelni arra, hogy míg a ROP-R-ben szabadon lehet megadni a változóneveket, a JASP a beolvasás során minden magyar ékezetes és nem alfanumerikus karakter (pl. kötőjel, kérdőjel, + jel stb.) helyett egy pont (.) karaktert tesz be. Ha ez nekünk nem felel meg, érdemes a ROP-R-ben (vagy a létrehozott szövegfájlban a változóneveket tartalmazó első sorban) az ilyen neveket számunkra megfelelő módon módosítani.

A CFA-elemzés mindkét programban a Factor-elemzéscsoport *Confirmatory Factor Analysis* modulja segítségével végezhető, melyek ugyanúgy a *lavaan* R-package-en alapulnak, ahogy a ROP-R CFA-elemzései is. A két program CFA-val kapcsolatos főbb jellemzőit a *12. táblázat*ban foglaljuk össze. Megjegyezzük, hogy az Mplus fizetős szoftver (Muthén & Muthén, 1998–2017) – néhány speciális illeszkedési mutató kiszámításától eltekintve – mindezen lehetőségekre alkalmas (és még sok másra is).

12. táblázat. A ROP-R, a jamovi és a JASP CFA-val kapcsolatos főbb jellemzői

Jellegzetesség	**ROP-R**	**jamovi**	**JASP**
Alkalmas a hiányzó adatok pótlására	csak az MLR becslési módszer esetén	igen	igen, de néha váratlan hibaüzenet megjelenik
Alkalmazható becslési módszer	ML, ULS és DWLS	ML	ML, GLS, WLS, ULS és DWLS
Standard hibák robusztus becslése	lehetséges	nem lehetséges	lehetséges
Mplus szoftverhez igazított output	lehetséges (fix beállítás)	lehetséges (fix beállítás)	lehetséges
EQS szoftverhez igazított output	nem lehetséges	nem lehetséges	lehetséges
Kiszámítható illeszkedési mutatók	χ^2/f, AIC, BIC, RMSEA, CI_{90}, pClose, CFI, TLI, SRMR	ugyanazok, mint a ROP-R-ben, kivéve χ^2/f és pClose	ugyanazok, mint a ROP-R-ben, valamint NNFI, NFI, PNFI, RFI, IFI, RNI, GFI, MFI, EVCI, SSABIC, Hoelter-féle kritikus N
Reziduális kovarianciák modellbe való beépítése	modifikációs index-küszöbök kijelölésével lehetséges	tételpárok kijelölésével manuálisan lehetséges	tételpárok kijelölésével manuálisan lehetséges
Faktorok és idegen tételek közti kovarianciák modellbe való beépítése	modifikációs index-küszöb kijelölésével lehetséges	nem lehetséges	nem lehetséges
Másodrendű modell vizsgálata	lehetséges	nem lehetséges	lehetséges
Bifaktoros modell vizsgálata	lehetséges	nem lehetséges	nem lehetséges
Mérésiinvariancia-elemzés	nem lehetséges	nem lehetséges	lehetséges

Megbeszélés

A pszichológiai tesztekkel kapcsolatos kutatásokban kiemelt fontosságú a tesztek hátterében álló faktorstruktúrák feltárása, valamint érvényességük vizsgálata. Ezzel kapcsolatban ismertettük cikkünkben a faktoranalízis, majd ezen belül a konfirmatív (megerősítő, modelltesztelő) faktoranalízis, a CFA alapfogalmait.

A CFA ismertetése során részletesen taglaltuk a különböző faktormodelleket, úgymint az elsőrendű egy- és többfaktoros modellt, valamint a másodrendű és a bifaktoros modellt. Ezután rátértünk a ROP-R CFA-moduljának bemutatására, ahol egy konkrét pszichológiai kutatás adatainak felhasználásával szemléltettük a CFA végrehajtásának módját különböző faktormodellek esetén.

Kitértünk arra is, hogy a ROP-R által létrehozott scriptek segítségével hogy lehet kétfaktoros másodrendű és bifaktoros CFA-elemzéseket végrehajtani, illetve olyan speciális beállításokat alkalmazni a CFA-ban, amelyek menüből nem elérhetők.

Végezetül áttekintettük két további ingyenes statisztikai szoftver: a jamovi és a JASP *Confirmatory Factor Analysis* moduljának lehetőségeit a CFA-elemzés végrehajtására. Mind a ROP-R, mind a jamovi és a JASP a *lavaan* R-package-en alapuló elemzéseket végez felhasználóbarát keretben. Néhány észrevételt a három szoftver összehasonlítását illetően az alábbiakban fogalmazunk meg.

1. A ROP-R és a JASP jó tulajdonsága, hogy többféle becslési módszer közül lehet választani. Ugyanakkor a jamoviban csak a standard ML-módszer használható.
2. Míg a jamoviban csak egyfaktoros modell tesztelhető, a ROP-R-ben és a JASP-ban másodrendű is. Ezen kívül a ROP-R lehetővé teszi a bifaktoros CFA végrehajtását is (alkalmas R-script elkészítésével).
3. Mindhárom szoftverben lehetséges az indokolt reziduális kovarianciák modellbe való beépítése. A ROP-R-ben a modifikációs küszöbök alkalmas beállításával, a jamovi és a JASP esetén pedig manuálisan. Az utóbbi két szoftver esetén csak tételek közti kovarianciák építhetők be, a ROP-R esetén viszont tételek és idegen faktor közti kovarianciák is.
4. A JASP egyedi vonása, hogy képes mérésiinvariancia-elemzésre is, különböző csoportokon végzett CFA-elemzések eredményeinek összehasonlításával.

KÖSZÖNETNYILVÁNÍTÁS

A jelen cikk elkészítését a Károli Gáspár Református Egyetem Bölcsészet- és Társadalomtudományi Karának *Pszichológiai kutatások módszertani platformja* című kutatói pályázata támogatta (témaszám: 20754B800/2022). Köszönettel tartozom Bánsági Péternek, hogy szerzőtársként nagy lelkesedéssel és hozzáértéssel vett részt a ROP-R CFA-moduljának programozásában. Itt szeretnék köszönetet mondani Jakab Zoltánnak, valamint a kéziratot bíráló anonim lektornak is, hogy alaposan elmélyedtek a kézirat első változatában, és számos hasznos észrevétellel és javaslattal járultak hozzá a cikk színvonalának emeléséhez. Köszönöm végül Oláh Attilának és a Jobb Veled a Világ Alapítványnak, hogy a Magyarország 2022-es Boldogságtérképe kutatás adatait a cikk szemléltető példáiban felhasználhattam.

Summary

Confirmatory factor analysis with ROP-R

Confirmatory factor analysis, commonly abbreviated as CFA, is a method of great importance in the study of the structural validity of psychological tests. Not so long ago, this method could only be carried out with the help of expensive software. Today, fortunately, several free software programs, available to everyone, include CFA modules. This paper presents the theoretical background of CFA and the use of the latest such software, ROP-R, to perform CFA. The paper also discusses the options available for the also free jamovi and JASP software. To help the reader to understand the essence of CFA analysis, the paper provides an illustrative example from real psychological research.

Keywords: confirmatory factor analysis, CFA, ROP-R, JASP, jamovi.

Irodalom

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705

Babakus, E., Ferguson, C. E., & Jöreskog, K. G. (1987). The sensitivity of confirmatory maximum likelihood factor analysis to violations of measurement scale and distributional assumptions. *Journal of Marketing Research, 24*(2), 222–228. https://doi.org/10.1177/002224378702400209

Bentler, P. M., & Wu, E. J. C. (2005). *EQS 6.1 for Windows. Structural equations program manual.* Multivariate Software.

Caprara, G. V., Barbaranelli, C., Borgogni, L., & Perugini, M. (1993). The "Big Five Questionnaire": A new questionnaire to assess the five factor model. *Personality and individual Differences, 15*(3), 281–288. https://doi.org/10.1016/0191-8869(93)90218-R

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling: A Multidisciplinary Journal, 14*(3), 464–504. https://doi.org/10.1080/10705510701301834

Collins, L. M., Schafer, J. L., & Kam, C. M. (2001). A comparison of inclusive and restrictive strategies in modern missing data procedures. *Psychological Methods, 6*(4), 330–351. https://doi.org/10.1037/1082-989X.6.4.330

Curran, P. J., West, S. G., & Finch, J. F. (1996). The robustness of test statistics to nonnormality and specification error in confirmatory factor analysis. *Psychological Methods, 1*(1), 16–29. https://doi.org/10.1037/1082-989X.1.1.16

Dolan, C. V. (1994). Factor analysis of variables with 2, 3, 5 and 7 response categories: A comparison of categorical variable estimators using simulated data. *British Journal of Mathematical and Statistical Psychology, 47*(2), 309–326. https://doi.org/10.1111/j.2044-8317.1994.tb01039.x

Dunn, K. J., & McCray, G. (2020). The place of the bifactor model in confirmatory factor analysis investigations into construct dimensionality in language testing. *Frontiers in Psychology, 11,* 1357. https://doi.org/10.3389/fpsyg.2020.01357

Gao, C., Shi, D., & Maydeu-Olivares, A. (2019). Estimating the maximum likelihood root mean square error of approximation (RMSEA) with non-normal data: A Monte-Carlo study. *Structural Equation Modeling: A Multidisciplinary Journal, 27*(2), 192–201. https://doi.org/10.1080/10705511.2019.1637741

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A Comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research, 51*(5), 661–680. https://doi.org/10.1080/00273171.2016.1208074

Hoyle, R. H. (2023) (Ed.). *Handbook of Structural Equation Modeling, Second Edition.* Guilford Press.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling, 6*(1), 1–55. https://doi.org/10.1080/10705519909540118

Hutchinson, S. R., & Olmos, A. (1998). Behavior of descriptive fit indexes in confirmatory factor analysis using ordered categorical data. *Structural Equation Modeling: A Multidisciplinary Journal, 5*(4), 344–364. https://doi.org/10.1080/10705519809540111

JASP Team (2023). *JASP (Version 0.17.2) [Computer software]*. https://jasp-stats.org

John, O. P., & Srivastava, S. (1999). The Big-Five trait taxonomy: History, measurement, and theoretical perspectives. In L. A. Pervin & O. P. John (Eds.), *Handbook of personality: Theory and research* (Vol. 2, pp. 102–138). Guilford Press.

Johnson, D. R., & Creech, J. C. (1983). Ordinal measures in multiple indicator models: A simulation study of categorization error. *American Sociological Review, 48*(3), 398–407. https://doi.org/10.2307/2095231

Kenny, D. A. (2020). Measuring Model Fit. http://davidakenny.net/cm/fit.htm

Kim, E. S., & Yoon, M. (2011). Testing measurement invariance: A comparison of multiple-group categorical CFA and IRT. *Structural Equation Modeling, 18*(2), 212–228. https://doi.org/10.1080/10705511.2011.557337

Kline, R. B. (2010). *Principles and practice of structural equation modeling* (3rd Ed.). Guilford Press.

Kuha, J. (2004). AIC and BIC: Comparisons of assumptions and performance. *Sociological methods & research, 33*(2), 188–229. https://doi.org/10.1177/0049124103262065

Kyriazos, T. A., & Poga-Kyriazou, M. (2023). Applied Psychometrics: Estimator Considerations in Commonly Encountered Conditions in CFA, SEM, and EFA Practice. *Psychology, 14*(5), 799–828. https://doi.org/10.4236/psych.2023.145043

Lee, S. Y., Poon, W. Y., & Bentler, P. M. (1995). A two-stage estimation of structural equation models with continuous and polytomous variables. *British Journal of Mathematical and Statistical Psychology, 48*(2), 339–358. https://doi.org/10.1111/j.2044-8317.1995.tb01067.x

Li, Y. (2014). *Confirmatory factor analysis with continuous and ordinal data: An empirical study of stress level. Doctoral thesis.* Uppsala University.

Li, C-H. (2016). The performance of ML, DWLS, and ULS estimation with robust corrections in structural equation models with ordinal variables. *Psychological Methods, 21*(3), 369–387. https://doi.org/10.1037/met0000093

Li, C-H. (2021). Statistical estimation of structural equation models with a mixture of continuous and categorical observed variables. *Behavior Research Methods, 53*(5), 2191–2213. https://doi.org/10.3758/s13428-021-01547-z

Lishinski, A. (2021). *lavaanPlot: Path Diagrams for ‚Lavaan' Models via ‚DiagrammeR'.* R package version 0.6.2. https://CRAN.R-project.org/package=lavaanPlot

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: the problem of capitalization on chance. *Psychological Bulletin, 111*(3), 490–504. https://doi.org/10.1037/0033-2909.111.3.490

Mîndrilă, D. (2010). Maximum likelihood (ml) and diagonally weighted least squares (dwls) estimation procedures: a comparison of estimation bias with ordinal and multivariate nonnormal data. *International Journal of Digital Society, 1*(1), 60–66. https://www.researchgate.net/publication/291006787

Muthén, B. O., du Toit, S. H. C. & Spisic, D. (1997). Robust inference using weighted least squares and quadratic estimating equations in latent variable modeling with categorical and continuous outcomes. *Unpublished technical report.* https://www.statmodel.com/papers_date.shtml

Muthén, B. O. (1994). Multilevel Covariance Structure Analysis. *Sociological Methods & Research, 22*(3), 376–398. https://doi.org/10.1177/0049124194022003006

Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide.* (8th Ed.) Muthén & Muthén.

Navruz, B. (2016). *The behaviors of robust weighted least squares estimation techniques for categorical/ordinal data in multilevel CFA models. PhD Dissertation.* Texas A&M University.

Oláh A. (2021). A Globális Jól-lét Modell kidolgozása és empirikus validitásának igazolása a személyiségtényezők figyelembe vételével – A K1A: 116965. számú NKFI kutatás zárójelentése. https://www.otka-palyazat.hu/download.php?type=zarobeszamolo&projektid=116965

Osborne, J. W. (2014). *Best practices in exploratory factor analysis.* CreateSpace Independent Publishing.

R Core Team (2021). *R: A language and environment for statistical computing.* R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/

Rosseel, Y. (2012). lavaan: An R package for structural equation modeling. *Journal of Statistical Software, 48*(2), 1–36. https://doi.org/10.18637/jss.v048.i02

Şahin, M. D., & Aybek, E. C. (2019). Jamovi: an easy to use statistical software for the social scientists. *International Journal of Assessment Tools in Education, 6*(4), 670–692. https://doi.org/10.21449/ijate.661803

Savalei, V., & Rhemtulla, M. (2013). The performance of robust test statistics with categorical data. *British Journal of Mathematical and Statistical Psychology, 66*(2), 201–223. https://doi.org/10.1111/j.2044-8317.2012.02049.x

Schermelleh-Engel, K., Moosbrugger, H., & Müller, H. (2003). Evaluating the fit of structural equation models: Tests of significance and descriptive goodness-of-fit measures. *Methods of Psychological Research Online, 8*(2), 23–74. http://www.mpr-online.de

Schivinski, B. (2013). Implementing second-order CFA model for the factorial validity of brand equity. *PhD Interdisciplinary Journal, 3*(1), 53–59. https://api.semanticscholar.org/CorpusID:51762909

Schreiber, J. B., Nora, A., Stage, F. K., Barlow, E. A., & King, J. (2006). Reporting Structural Equation Modeling and Confirmatory Factor Analysis Results: A Review. *The Journal of Educational Research, 99*(6)*,* 323–338. https://doi.org/10.3200/JOER.99.6.323-338

Schwarz, G. E. (1978). Estimating the dimension of a model. *Annals of Statistics, 6(2),* 461–464. https://doi.org/10.1214/aos/1176344136

Shaharudin, S. M., Ahmad, N., Zainuddin, N. H., & Mohamed, N. S. (2018). Identification of rainfall patterns on hydrological simulation using robust principal component analysis. *Indonesian Journal of Electrical Engineering and Computer Science, 11*(3), 1162–1167. https://doi.org/10.11591/ijeecs.v11.i3.pp1162-1167

Shi, D., & Maydeu-Olivares, A. (2020). The effect of estimation methods on SEM fit indices. *Educational and Psychological Measurement, 80*(3), 421–445. https://doi.org/10.1177/0013164419885164

Tabachnick, B. G., & Fidell, L. S. (2013). *Using multivariate statistics* (6th Ed.). Pearson.

The jamovi project (2022). jamovi (Version 2.3) [Computer Software]. https://www.jamovi.org

Tucker, L. R., & Lewis, C. (1973). A reliability coefficient for maximum likelihood factor analysis. *Psychometrika, 38*(1), 1–10. https://doi.org/10.1007/BF02291170

Vandenberg, R. J., & Lance, C. E. (2000). A Review and Synthesis of the Measurement Invariance Literature: Suggestions, Practices, and Recommendations for Organizational Research. *Organizational Research Methods, 3*(1), 4–70. https://doi.org/10.1177/109442810031002

Vargha, A. (2016). A ROPstat statisztikai programcsomag. *Statisztikai Szemle, 94*(11–12), 1165–1192. https://doi.org/10.20311/stat2016.11-12.hu1165

Vargha, A. (2019). *Többváltozós statisztika dióhéjban: változó-orientált módszerek.* Pólya Kiadó.

Vargha A. (2023). *Többváltozós statisztikai elemzések pszichológiai kutatásokban ROP-R-rel.* Pólya Kiadó. https://www.bansagi.hu/r/vargha_tv_elemzesek_rop-r_rel_2024.pdf

Vargha, A. & Bánsági, P. (2022). ROP-R: a free multivariate statistical software that runs R packages in a ROPstat framework. *Hungarian Statistical Review, 5*(2), 3–29. https://doi.org/10.35618/HSR2022.02.en003

Vargha, A., Bánsági, P., & Jantek, Gy. (2024). Statisztikai elemzések a ROP-R szoftver segítségével és szemléltetésük egy kötődéskutatás adataival. *Mentálhigiéné és Pszichoszomatika, 25*(1), 36–55. https://doi.org/10.1556/0406.2024.00028

Vargha, A., Zábó, V., Török, R., & Oláh, A. (2020). A jóllét és a mentális egészség mérése: a Mentális Egészség Teszt. *Mentálhigiéné és Pszichoszomatika, 21*(3), 281–322. https://doi.org/10.1556/0406.21.2020.014

Melléklet

1. melléklet. A faktoranalízissel kapcsolatos fogalmak angol elnevezéssel és jelentéssel

Magyar elnevezés	**Angol elnevezés**	**Jelentés (magyarázat)**
χ^2/f (takarékossági index)	χ^2/df (parsimony index, CMIN/DF)	A CFA-ban a modellt tesztelő χ^2-statisztika és szabadságfokának hányadosa.
AIC információs kritérium	AIC (Akaike Information Criterion)	A CFA-ban egy likelihoodon alapuló illeszkedési mutató, mellyel különböző CFA-modellek hasonlíthatók össze.
alapmodell	baseline model	A CFA-ban a lehető legrosszabb modell, ahol a változók és a látens faktorok együttesében minden korreláció 0, vagyis ahol semmi nem függ össze semmivel.
alfaktor (facet)	facet	Lásd másodrendű faktormodell.
általános faktor	general factor	Lásd bifaktoros modell.
BIC információs kritérium	BIC (Bayesian Information Criterion)	A CFA-ban egy likelihoodon alapuló illeszkedési mutató, mellyel különböző CFA-modellek hasonlíthatók össze.
bifaktoros modell	bifactor model	A CFA-ban olyan modell, amelynél az elemzett teszttételeket egyidejűleg egy elsőrendű általános látens faktor, valamint ettől függetlenül több elsőrendű specifikus faktor magyarázza.
CFI	CFI (Comparative Fit Index)	Egy relatív illeszkedési mutató a CFA-ban (mennyivel jobb a kijelölt modell χ^2-értéke, mint az alapmodellé).
DWLS	DWLS (diagonally weighted least squares)	A CFA-ban a diagonálisan súlyozott legkisebb négyzetek módszere.
egyediség (unicitás, reziduális variancia-arány)	uniqueness (residual variance ratio)	Az EFA-ban és a CFA-ban az egyes inputváltozók varianciájának az a része, amelyet a látens faktorok nem magyaráznak meg (a kommunalitást 1-re kiegészítő érték).
elsődleges faktorok	primary factors	Az EFA-ban elsődlegesen feltárt látens faktorstruktúra faktorai.
elsődleges faktorsúly-mátrix	primary factor loading matrix	Standardizált regressziós együtthatók, amelyeket az EFA-ban az elsődlegesen feltárt látens faktorok együttese mint független változók segítségével az elemzésbe bevont változók regressziós előrejelzésére kapunk (egyben a változók és a látens faktorok közti becsült korrelációk).
elsőrendű faktormodell	first order factor model	A CFA-ban olyan modell, amelynél az elemzett teszttételeket egy vagy több egymással egyenrangú látens faktor feszíti ki (magyarázza).

Magyar elnevezés	Angol elnevezés	Jelentés (magyarázat)
elsőrendű többfaktoros faktormodell	first order multifactor model	A CFA-ban olyan elsőrendű faktormodell, amelynél az elemzett teszttételeket egynél több egymással egyenrangú látens faktor feszíti ki.
facet (alfaktor)	facet	Lásd másodrendű faktormodell.
faktorkorrelációk	factor correlations	Az EFA-ban (ferde forgatás esetén) vagy a CFA-ban a látens faktorok egymás közti becsült korrelációi.
faktordiagram (faktorábra)	path diagram	A faktormodell ábrázolása, ahol a diagramon a látens faktorokat ellipszis vagy kör, a manifeszt változókat téglalap alakú keretbe helyezik. A diagramon az egymással kapcsolatban lévő elemeket irányított nyilak kötik össze, ezeken regressziós együtthatók láthatók.
feltáró, exploratív faktoranalízis (EFA)	exploratory factor analysis (EFA)	Egy statisztikai módszer, amelyet viszonylag nagyszámú inputváltozó mögöttes (látens) struktúrájának feltárására használnak.
FIML	FIML (Full Information Maximum Likelihood)	Olyan adatkezelési módszer, mely hiányzó adatok esetén az összes rendelkezésre álló információt felhasználja a modell becsléséhez, és olyan faktormodellt konstruál, amely mellett a legnagyobb az adott minta bekövetkezésének a valószínűsége.
forgatás (rotáció)	rotation	Az FKA-ban a főkomponensek, az EFA-ban az elsődlegesen feltárt faktorok olyan lineáris transzformációja, amelynek hatására javul az illeszkedés az elemzett változókra, ami segít a forgatott komponensek, illetve faktorok értelmezésében.
forgatott faktorsúlymátrix (mintázatmátrix)	rotated factor loading matrix (pattern matrix)	Az EFA-ban a forgatott faktorok súlymátrixa.
forgatott komponensek korrelációs mátrixa	correlation matrix of rotated components	Az FKA-ban ferde forgatás esetén a forgatott komponensek korrelációs mátrixa.
forgatott komponenssúly-mátrix (mintázatmátrix)	rotated component loading matrix (pattern matrix)	Az FKA-ban a forgatott komponensek súlymátrixa.
főfaktorelemzés (PAF)	principal axis factoring (PAF)	Módszer az EFA-ban az elsődleges faktorok feltárására.
főkomponens	principal component	Lásd főkomponens-analízis.
főkomponens-analízis (FKA)	principal component analysis (PCA)	Az inputváltozók súlyozott összegeiként kevés számú korrelálatlan összetevőt, ún. főkomponenst hozunk létre, amelyek az eredeti változók teljes varianciájának lehető legnagyobb részét megmagyarázzák.

Magyar elnevezés	Angol elnevezés	Jelentés (magyarázat)
hierarchikus faktor-modell	hierarchical factor model	Lásd másodrendű faktormodell.
Kaiser-Meyer-Olkin (KMO) mutató	Kaiser-Meyer-Olkin criterion	Az EFA-ban az inputváltozók átfedésének (közös részének) mérésével az EFA alkalmasságát jelző mutató.
kommunalitás	communality	Az FKA-ban, az EFA-ban vagy a CFA-ban a komponensek vagy faktorok által együtt az egyes inputváltozók varianciájából megmagyarázott varianciaarány.
konfirmatív faktoranalízis (CFA)	confirmatory factor analysis (CFA)	Megerősítő, modelltesztelő faktoranalízis.
látens faktorok	latent factors	Az EFA-ban a feltárt látens faktorstruktúra elemei.
másodrendű faktor-modell (hierarchikus faktormodell)	second order factor model (hierarchical factor model)	A CFA-ban olyan modell, melynél az elemzett teszttételeket közvetlenül egy vagy több elsőrendű látens faktor magyarázza úgy, hogy közben ezek az elsőrendű faktor (alfaktor, facet) komponensek együtt egy vagy több másodrendű faktorra illeszkednek.
maximum likelihood (ML)	maximum likelihood (ML)	Becslési módszer a statisztikában, egyben egy speciális módszer az EFA-ban az elsődleges faktorok feltárására és a CFA-ban a faktormodell becslésére.
mérési invariancia	measurement invariance	A CFA-ban olyan elemzés, mely azt nézi, hogy egy faktormodell különböző csoportokban ugyanolyan szintű és mintázatú illeszkedést mutat-e. Ha igen, akkor fennáll a mérési invariancia.
minimum reziduális módszer (MinRes)	minimum residual method (MinRes)	Módszer az EFA-ban az elsődleges faktorok feltárására.
mintázatmátrix	pattern matrix	Az FKA-ban vagy az EFA-ban a forgatott komponensek, illetve faktorok súlymátrixa.
MLMV	MLMV (ML estimation with robust SEs and a mean- and variance adjusted test statistic)	ML-becslés a CFA-ban robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával.
MLR	MLR	ML-becslés a CFA-ban Huber-White robusztus standard hibákkal, valamint egy Yuan-Bentler-féle statisztikával aszimptotikusan megegyező tesztstatisztikával.
modifikációs index	modification index	A CFA-ban megadja, hogy milyen mértékben javulna a modell (milyen mértékben csökkenne a modellilleszkedést mérő χ^2-érték), ha egy kovarianciára vonatkozó korlátozást megszüntetnénk.

Magyar elnevezés	Angol elnevezés	Jelentés (magyarázat)
Oblimin	Oblimin	Egy ferde forgatási módszer az FKA-ban és az EFA-ban.
pClose	pClose (probability of a close fit)	Annak a próbának a p-értéke, mely azt teszteli, hogy az elméleti RMSEA kisebb-e 0,05-nél.
polikorikus korreláció	polichoric correlation	Feltételezve, hogy két diszkrét változónk értékei valójában két folytonos normális eloszlású változó kerekített értékei, a polikorikus korreláció a reprodukált normális eloszlású változók között számított Pearson-korreláció.
Promax	Promax	Egy ferde forgatási módszer az FKA-ban és az EFA-ban.
reziduális variancia-arány	residual variance ratio	Lásd egyediség.
RMSEA	RMSEA (Root Mean Square Error of Approximation)	Egy abszolút illeszkedési mutató a CFA-ban (a négyzetre emelt reziduálisok átlagának a gyöke).
rotáció		Lásd forgatás.
sajátérték	eigenvalue	Az FKA-ban egy-egy főkomponens esetén az inputváltozókkal mért korrelációk négyzetösszege (ez mutatja, hogy a főkomponens hány változónyi varianciát fed le, azaz magyaráz).
specifikus faktor	specific factor	Lásd bifaktoros modell.
SRMR	SRMR (Standardized Root Mean Square Residual)	Egy abszolút illeszkedési mutató a CFA-ban (a hipotetikus modell és a megfigyelt modell korrelációs mátrixa közötti átlagos négyzetes eltérés gyöke).
struktúramátrix	structure matrix	Az FKA-ban vagy az EFA-ban ferde forgatás esetén az elemzésbe bevont változók és a forgatott komponensek, illetve faktorok korrelációs mátrixa.
takarékossági index	parsimony index	Lásd χ^2/f.
TLI	TLI (Tucker-Lewis Index)	Egy relatív illeszkedési mutató a CFA-ban, a CFI korrigált változata.
ULS	ULS (unweighted least squares)	A CFA-ban a súlyozatlan legkisebb négyzetek becslési módszere.
ULSMV	ULSMV	A CFA-ban az ULS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával.
unicitás	uniqueness	Lásd egyediség.
Varimax	Varimax	Egy ortogonális (a korrelálatlanságot megőrző) forgatási módszer az FKA-ban és az EFA-ban.
WLS	WLS (weighted least squares)	A CFA-ban a súlyozott legkisebb négyzetek becslési módszere.

Magyar elnevezés	Angol elnevezés	Jelentés (magyarázat)
WLSM	WLSM	A CFA-ban a DWLS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot alkalmazó próbastatisztikával.
WLSMV	WLSMV	A CFA-ban a DWLS-becslés robusztus variánsa robusztus standard hibákkal, valamint korrigált átlagot és varianciát alkalmazó próbastatisztikával.