Зборник Института за педагошка истраживања УДК 303.725.3:519
Година 40 • Број 1 • Јун 2008 • 101-121 Прегледни чланак
ISSN 0579-6431 DOI: 10.2298/ZIPI0801101L

ПРИМЕНА ИНДЕКСА ПОДЕСНОСТИ У ТЕСТИРАЊУ ТЕОРИЈСКИХ МОДЕЛА У ПСИХОЛОГИЈИ: МОГУЋНОСТИ И ОГРАНИЧЕЊА

*Љиљана Лазаревић**
Београд

Апстракт. Рад је посвећен индексима подесности који се користе у моделовању структуралним једначинама (SEM) за тестирање теоријских модела и тешкоћама које се могу јавити приликом тестирања теоријских модела у различитим областима психологије. Разматране су основне поставке SEM-а и представљени су индекси за процену подесности теоријских модела. У раду су приказани и поступци за рачунање основног статистика за процену подесности модела (χ^2), као и најчешће примењиваних индекса подесности, како би се стекао бољи увид у предности и потенцијалне тешкоће које се могу јавити приликом њихове употребе. Наводе се тешкоће у процени подесности модела на основу χ^2 и разматраних индекса подесности које произлазе из величине узорка, дистрибуције података и метода процене, погрешне спецификације модела и нарушавања нормалности и независности латентних варијабли, као и начини на које се ове тешкоће могу превазићи. У раду се даје предлог приступа у представљању индекса подесности у извештајима о истраживањима у којима су теоријски модели тестирани помоћу SEM-а.
Кључне речи: моделовање структуралним једначинама, подесност теоријских модела, индекси подесности, психологија.

У психологији, као и у другим друштвеним наукама се још од седамдесетих година XX века веома много ради на развијању математичко-статистичке каузалне анализе која би омогућила тестирање и доказивање каузалних хипотеза. У те сврхе често је присутно моделовање структуралним једначинама (structural equation modelling – SEM) (Anderson & Gerbing, 1988; DiLalla, 2000; Fajgelj, 2004). У актуелним истраживањима у различитим областима психологије, као што су психологија индивидуалних разлика, психологија образовања, психологија спорта и физичког вежбања, клиничка психологија, итд., моделовање структуралним једначинама представља саставни део истраживачке методологије (Fassinger, 1987; Lazarević, 2007a, 2007b; Novović i sar, 2006; Schutz, 1998; Smederevac i sar, 2006; Wright, 2006).

* E-mail: lily@ptt.yu

Моделовање структуралним једначинама (SEM) представља низ хипотеза о томе на који начин су варијабле у анализи генерисане и у ком су међусобном односу (Hu & Bentler, 1999). SEM укључује анализу латентних варијабли и анализу пута за тестирање претпостављених модела и откривање односа манифестних и латентних варијабли (DiLalla, 2000).

На основу SEM-а, истраживачи су у могућности да тестирају да ли варијансе и коваријансе у матрици коваријанси одговарају специфичној структури (Bentler & Bonett, 1980; Bentler & Mooijaart, 1989). Теорије које се тестирају у оваквим анализама јесу оне које се могу представити као систем једначина који описује једносмерне или двосмерне утицаје манифестних и латентних варијабли међусобно (Bentler & Bonett, 1980).

У SEM-у претпостављамо да између сета неопсервабилних конструката које меримо помоћу одређених индикатора, постоји одређена каузална структура, чију ћемо подесност накнадно тестирати у одређеној популацији (Fassinger, 1987). Основни циљеви SEM-а су процена подесности, тј. фитовања модела и процена параметара (индекса подесности, тј. фит индекса) постављеног, односно дефинисаног модела (Hu & Bentler, 1999).

Међутим и поред учестале примене SEM-а у истраживањима, постоји потреба за даљим усавршавањем овог статистичког поступка (Vernon & Eysenck, 2007). Процена да ли одређени теоријски модел фитује емпиријским подацима, доноси се на основу више параметара (Hu & Bentler, 1999; Thompson, 2005). Међу ауторима постоји несагласност у погледу примењивости одређених индекса подесности. Пре свега, тешкоће у избору одређених индекса подесности зависе од одређених карактеристика истраживања, као што су величина узорка или нарушавање претпоставки које стоје у основи теста. Такође, компјутерски софтвери нуде велики број индекса што доводи до неколико проблема. Наиме, у извештајима о истраживањима се од обиља индекса подесности које нуде статистички софтвери појављују само неки, а може се десити да рецензенти истог чланка захтевају навођење различитих индекса (Kline, 2005). Оваква ситуација доводи истраживача у недоумицу које индексе би требало да наведе у свом извештају. Такође, може се десити да истраживач, у жељи да представи свој модел као подесан, извести само о оним индексима који иду у прилог његовим закључцима, или да у случају доброг фита модела превиди анализу смислености индекса подесности (Kline, 2005).

Хи-квадрат, основни статистик за процену подесности модела: карактеристике и тешкоће у примени

Основни статистик за процену фитовања теоријског модела емпиријским подацима је хи-квадрат којим проверавамо величину разлике између матрица коваријанси које предвиђа модел и матрица емпиријских коваријанси (Barrett, 2007; Hu & Bentler, 1999). Хи-квадрат спада у апсолутне индексе подесности, што значи да се као основа за поређење не користи алтернативни модел (Hoyle, 2000).

Уколико је статистик фитовања Т, или како га називају многи истраживачи – χ^2, мањи од очекиване вредности са вероватноћом већом од 0.05, процењујемо да модел одговара подацима, односно да нулту хипотезу »да нема разлика између популацијских коваријанси које предвиђа модел и емпиријских коваријанси узорка«, не одбацујемо (Barrett, 2007: 816). Т статистик се процењује на основу различитих метода процене, који варијају у степену осетљивости на претпоставке дистрибуције. Неки аутори наводе да је Т статистик, процењен на основу метода *максималне веродостојности* (Maximum likelihood – ML) под претпоставком мултиваријатне нормалне дистрибуције, највише коришћен статистик за процену адекватности структурног модела (Hu & Bentler, 1998)

Поред претходно наведеног начина одређивања значајности χ^2, неки аутори (Byrne, 1989; Kline, 2005) наводе да се у ту сврху може користити нормирани хи-квадрат, односно количник хи-квадрата и степени слободе (df) (CMIN/df ili χ^2/df). Овакав поступак се спроводи да би се смањила осетљивост χ^2 на величину узорка (Kline, 2005). Уколико је тај количник мањи од 2, сматра се да је то прихватљиво и да по том параметру модел фитује подацима. Ипак, неки аутори (Barrett, 2007; Hu & Bentler, 1998; McIntoch, 2007) наводе извесне изворе тешкоћа у коришћењу хи квадрата:

• *Величина узорка*. Многи аутори (Barrett, 2007; Bentler, 2007; Bentler & Bonett, 1980; Hu & Bentler, 1999; Miles & Shevlin, 2007; Thompson, 2005) наводе да, пошто је један од параметара за рачунање χ^2 статистика и величина узорка, већина тестираних модела са узорком већим од 200 субјеката, не одговара подацима. У великим узорцима, добија се велики χ^2 који има већу шансу да буде статистички значајан, чиме се повећава могућност јављања грешке типа I, то јест могућност да одбацимо тачну нулту хипотезу. Другим речима, величина узорка мултипликује функцију дискрепанце, због чега се у великим узорцима мале дискрепанце између опсервираних тест статистика и њихових очекиваних вредности под претпоставком нулте хипотезе повећавају, а такви резултати се сматрају последицом лошег фита. У малим узорцима већа је вероватноћа да прихватимо погрешан модел, чиме се повећава могућност јављања грешке типа II, то јест прихватања погрешне нулте хипотезе.

• *Величина корелација*. Веће корелације доводе до већих вредности χ^2, делимично због тога што веће корелације дозвољавају могућност већих разлика између опсервираних и моделом претпостављених корелација (Kline, 2005).

• *Комплексност модела*. Модели који имају већи број варијабли имају веће хи-квадрате. Рајков и Маркулидес (Raykov & Marcoulides, 2000; према: Kline, 2005: 136) наводе да »сваки степен слободе за хи-квадрат тестираног модела може бити димензија дуж које модел може бити одбачен«. Другим речима, ако узмемо у обзир два модела са различитим бројем степени слободе, онај модел који има више степени слободе (тј. мање параметара) је пожељнији јер има већу снагу да не буде одбачен. Ако узмемо у обзир два модела са истом експланаторном моћи за исте податке, преферираћемо онај модел који је једноставнији, јер је циљ тежити парсимонији (Kline, 2005).

• *Дистрибуција варијабли*. Неодговарајућа вредност χ^2 може бити последица озбиљног нарушавања нормалности расподеле података, чак и ако је модел добро специфициран (McIntosh, 2007; Mulaik *et al.*, 1989). Барет (Barrett, 2007) сугерише да је у случају одбацивања модела на основу χ^2, најбоље прво испитати мултинормалност дистрибуције и уколико је потребно, извршити рескалирање, трансформацију или чак избацивање појединих варијабли из анализе. Степен одступања од мултинормалне дистрибуције може se проценити и на основу мултиваријатног скјуниса и куртозиса (MacIntosh, 2007). Неки аутори наводе (McIntosh, 2007) да је у случају нарушавања дистрибуције могуће спровести и мултиваријатне трансформације како би се нормализовали подаци. Такође, могуће је користити и другачије методе процене као што су ADF (енг. asymptotic distribution free) или GLS (енг. generalised least squares) (McIntosh, 2007).

МекИнтош (McIntosh, 2007) наводи да се може показати да модел не фитује подацима и због погрешних рестрикција на параметрима (нпр. неукључивање неопходних путева) што би могао да буде индикатор фундаменталних грешака у полазној теорији.

Неки аутори сматрају (Markland, 2007) да вредност χ^2 не треба посматрати на дихотоман начин (прихвата се/одбија), већ као индикатор исправности модела, а да као помоћ у интерпретацији података треба користити друге параметре. С обзиром на то да је опште прихваћено (Hu & Bentler, 1998; Jöreskog, 1978) да су модели апроксимације реалности, употреба χ^2 за тестирање хипотезе да се матрица коваријанси у популацији поклапа са моделом претпостављеном матрицом, може бити превише строга. По мишљењу Јорескога (Jöreskog, 1978: 448), приликом тестирања модела, »модел би требало да представља прихватљиву апроксимацију популацијске матрице коваријанси«. Стога, употреба χ^2 теста некад није довољно добра за процену адекватности модела.

Индекси фитовања: индекси за процену подесности модела

Да би се превазишле неке од тешкоћа везане за χ^2 тест (величина узорка, мултинормалност расподеле, нарушавање претпоставки теста итд.), користе се и индекси фитовања, који су препоручене мере фитовања модела (Hu & Bentler, 1998). За разлику од χ^2 теста, који нуди дихотоман поглед на адекватност модела, индекси фитовања нуде могућност квантификације степена фитовања дуж континуума, односно квантификације степена до ког се варијације и коваријације података могу објаснити моделом (Hu & Bentler, 1998). Опште прихваћена подела индекса фитовања је на апсолутне и инкременталне, тј. компаративне (Hu & Bentler, 1998; Hu & Bentler, 1999).

Апсолутни индекси фитовања процењују колико добро *a priori* модел репродукује податке из узорка. У случају апсолутних индекса фитовања нема референтног модела помоћу кога бисмо могли да проценимо степен инкремента у фитовању модела, али се може вршити поређење са сатурисаним моделом који тачно репродукује матрицу коваријанси посматраног узорка (Hu & Bentler, 1998).

У апсолутне индексе фитовања спадају »Индекс подесности модела« (Goodness-of-Fit – GFI), »Кориговани индекс подесности модела« (Adjusted Goodness-of-Fit – AGFI), »Акајкеов критеријум информативности« (Akaike information criterion – AIC), χ^2/df количник, »Бајесов критеријум информативности« (Bayesian information criterion – BIC), »Браун-Кудеков критеријум« (Brown-Cudeck – BCC), »Холтеров критични Н« (Hoelter's critical N - CN или HOELTER), »Индекс очекиване кросвалидације« (енг. »Expected Cross-validation Index - ECVI)«, Стандардизовани квадратни корен просечног квадрата резидуала (Standardized root mean squared residual – SRMR), »Квадратни корен просечне квадриране грешке апроксимације« (Root-mean-square error of approximation – RMSEA) и други (Barrett, 2007; Hu & Bentler, 1999). Код већине ових индекса, осим код SRMR, могу се јавити исти проблеми као и код хи-квадрата, јер сви представљају неку од варијација хи-квадрата.

Инкрементални индекси подесности мере пропорцију побољшавања у фитовању поредећи тестирани модел са рестриктивнијим, основним моделом. Типични основни модел је нулти модел у којем су све опсервиране варијабле некорелисане и који увек има веома велики χ^2, што означава веома лош фит модела (Hu & Bentler, 1998; Hu & Bentler, 1999). Другим речима, инкрементални индекси нам говоре »колико је добар мој модел у односу на најгори модел који постоји« (Miles & Shevlin, 2007: 870). Предност инкременталних индекса је у томе што без обзира на повећање величине узорка, њихова вредност остаје непромењена (Miles & Shavlin, 2007).

У инкременталне фит индексе спадају: »Нормирани фит индекс« (Normed fit index – NFI), »Ненормирани фит индекс (Non-normed fit index – NNFI) Бентлера и Бонета, који се још зове и »Такер-Луисов индекс« (Tucker-Lewis Index – TLI), »Индекс инкременталног фитовања« (Bollen's incremental index – BL 89 или IFI), »Индекс компаративног фитовања« (Bentler's Comparative fit index – CFI), »Релативни нецентрализовани индекс« (Relative noncentrality index – RNI), »Индекс релативног фитовања« (Bollen's incremental index (BL 86 ili RFI)), и други (Hu & Bentler, 1999; Miles & Shevlin, 2007). Ху и Бентлер (Hu & Bentler, 1998) су инкременталне фит индексе поделили у три групе:

• *Индекси у којима основни T статистик („goodness-of-fit“ статистик) нема претпостављену дистрибуцију.* Општа формула оваквог индекса је:

$|T_B - df_T|/T_B$, где T_B представља T статистик за основни модел, а df_T број степени слободе за тестирани модел. Ови индекси говоре колики проценат информација о везама између варијабли се може објаснити претпостављеним моделом (Mulaik *et al.*, 1989). У ове индеске спадају NFI i RFI.

• *Индекси у којима статистик има претпостављену асимптотску χ^2 дистрибуцију, са аритметичком средином df (где је df број степени слободе за процењивани модел).* У овом случају, основни статистик фитовања T_B пореди се са df_T, а именилац уместо T_B постаје (T_B - df_T). Општа формула оваквих индекса је:

$|T_B - T_T|/(T_B - df_T)$, где T_B представља T statистик за основни модел, T_T представља T статистик за тестирани модел, а df_T број степени слободе за тестирани модел. Ови индекси се зову и ненормирани индекси фитовања, јер немају опсег од 0 до 1 чак и када је $T_B \geq T_T$. У ове индексе спадају NNFI ili TLI i IFI ili BL89.

Индекси који користе информације добијене на основу индекса прве групе као и додатне информације које добијамо из очекиваних вредности T_B и/или T_T уз нецентрализовану хи-квадрат дистрибуцију. Нецентрализовани индекси фитовања укључују најпре дефинисање популацијског индекса фитовања, а затим користе процене ових параметара да би се дефинисао узорачки индекс фитовања. У ове индексе спадају RNI i CFI.

Уколико T статистик нема претпостављену дистрибуцију, дешава се да информација о подесности модела не буде адекватна. Ипак, Ху и Бентлер (Hu & Bentler, 1998) сматрају да када су претпостављене дистрибуције исправне, индекси друге и треће групе дају прецизније процене него индекси из прве групације.

Неки аутори сматрају (Mulaik *et al*., 1989) да је приликом разматрања подесности модела, поред индекса фитовања потребно узети у обзир и индексе парсимоније, који чине посебну категорију фит индекса. Ови индекси спадају у релативне индексе фитовања (индексе код којих се χ^2 тестираног модела пореди са χ^2 независног модела). Помоћу ових индекса може се олакшати одбацивање модела који су мање парсимонични (како би се фаворизовале једноставније теорије у односу на комплексне). У индексе парсимоније спадају PGFI (који је заснован на GFI), PNFI (заснован на NFI), PNFI2 (заснован на IFI), PCFI (заснован на CFI). Индекси парсимоније могу бити корисни када тестирамо алтернативни модел, како бисмо утврдили који модел ће дати једноставније решење. Неки аутори наводе (Hu & Bentler, 1999) да индекси парсимоније могу бити корисни у неким решењима, али да често могу навести истраживача на погрешно решење, па их због тога треба користити само као »оријентационо правило« (rule of thumb). Уколико се истраживач одлучи да за процену пожељнијег модела користи индексе парсимоније, Ху и Бентлер (Hu & Bentler, 1999) сугеришу да се вредности индекса парсимоније које су једнаке 0.80 или веће од 0.80 сматрају прихватљивим.

Тешкоће у употреби индекса фитовања за процену подесности модела

Међу ауторима (Hu & Bentler, 1998) постоји схватање да добар фит индекс мора да испуни неколико услова: да буде осетљив на погрешну спецификацију модела, стабилан у различитим методима оцене, на узорцима различите величине и у подацима различитих дистрибуција. У истом раду, Ху и Бентлер (Hu & Bentler, 1998) наводе да постоје четири основна проблема који могу утицати на процену фитовања модела на основу индекса фитовања:

• Осетљивост индекса фитовања на погрешну спецификацију модела.

Исправна спецификација модела подразумева да популација потпуно тачно одговара хипотетичном моделу и да параметри процењени у узорку одражавају ту структуру. Другим речима, модел је погрешно спецификован (Hu & Bentler, 1998):

а) када је вредност једног или више параметара у популацији процењена да буде нула – препараметризовани модел (енг. overparameterized model),

б) када је вредност једног или више параметара фиксирана на нулу, а популацијске вредности нису нула – недовољно параметризовани модел (енг. underparameterized model), и

в) када постоји комбинација и једне и друге могућности.

Ху и Бентлер (Hu & Bentler, 1998) извештавају да су истраживања показала да постоји утицај исправности спецификације модела на фит индексе. Истраживање Марша и сарадника је показало да величина узорка утиче на неколико индекса фитовања и код исправних и код погрешних модела (Marsh, Balla & McDonald, 1988). Такође, вредности већине апсолутних и индекса фитовања из друге групе Ху и Бентлера у исправним моделима имају значајно веће вредности него индекси добијени у погрешним моделима.

Ла Ду и Танака (La Du & Tanaka, 1989) su показали да у случају препараметризованог модела не постоји утицај погрешне спецификације на вредности индекса. Међутим, у случају недовољно параметризованог модела (код кога популационе вредности једног или више параметара нису нула, али су фиксиране на нулу) пронађен је мали али статистички значајан утицај на неке од индекса фитовања. Марш, Бала и Хау (Marsh *et al*., 1996, према: Hu & Bentler, 1998) су такође пронашли утицај погрешне спецификације модела на вредности индекса TLI, IFI, NFI, CFI i RNI, али њихово истраживање није показало и ступањ осетљивости индекса на погрешну спецификацију модела.

• Пристрасност малог узорка. Када говоримо о методама оцењивања у структурном моделовању, потребно је да напоменемо да су методе оцењивања засноване на две претпоставке (Hu & Bentler, 1998). Једна је да је модел $\Sigma=\Sigma(\theta)$, тј. да матрица коваријанси популације одговара матрици коваријанси претпостављеног модела. Друга претпоставка је да су процене и тестови засновани на великим узорцима, што се у стварности не дешава често. Уколико величина статистика варира у узорцима различите величине, такав статистик ће бити пристрасан оценитељ одговарајућег популационог параметра. Постојање утицаја величине узорка на вредност статистика нас онемогућава да поуздано и ваљано процењујемо фитовање модела. Поједини аутори (Bearden, *et al*., 1982; према: Hu & Bentler, 1998) су још 1982. године пронашли да је аритметичка средина индекса NFI у позитивној вези са величином узорка и да његова вредност има тенденцију да буде мања од 1 када је узорак мали. Њихови рани налази су истакли веома важан проблем, а то је могућност систематске пристрасности индекса подесности.

Болен (Bollen, 1990) сматра да постоје два типа ефекта величине узорка на меру фита:

а) Величина узорка (N) директно улази у калкулацију индекса фита.

б) Аритметичка средина дистрибуције узорковања индекса фита је у вези са величином узорка. Замислимо да имамо велики број насумичних узорака свих могућих величина и претпоставимо да исправни модел важи за сваки узорак (за сваки узорак се процењује модел и рачунају се индекси фита). Прикупимо све мере фитовања оцењене у узорку одређене величине, добијемо дистрибуцију узорковања мера фитовања и израчунамо њихову аритметичку средину. Када добијемо аритметичке средине за узорак сваке величине, добићемо аритметичку средину дистрибуције узорковања за узорке различите величине (N-ове). Ако је N повезана са аритметичким срединама дистрибуције узорковања, добијамо други ефекат величине узорка (Bollen, 1990).

На ову тему је урађен велики број истраживања. Ху и Бентлер (Hu & Bentler, 1998) на основу прегледа већег броја радова из ове области, извештавају да величина узорка мање утиче на инкременталне индексе друге и треће групе, него на апсолутне индексе и инкременталне индексе прве групе. Индекс TLI, а нарочито индекси CFI, IFI, RNI не показују пристрасност.

• Ефекти метода оцене. Методи оцењивања који се користе у анализи структуре коваријанси су развијени под претпоставком мултиваријатне нормалности (Hu *et al*., 1992). Међутим, у пракси претпоставка о оваквој дистрибуцији често није потврђена. У прилог оваквој тврдњи може се навести извештај који је објавио Micceri (1989) да су од 440 психометријских и мера постигнућа добијених на великим узорцима који су објављени у чланцима, пројектима и тестовима све имале дистрибуцију која је значајно одступала од нормалне. Међутим, неки аутори наводе (Micceri, 1989; Bentler & Dudgeon, 1996) да упркос томе што често нису задовољени услови мултинормалне расподеле, истраживачи користе методе оцене *максималне веродостојности* (Maximum likelihood) и *генерализованих најмањих квадрата* (generalised least squares) који су развијени под претпоставком нормалне дистрибуције, чиме се може озбиљно нарушити тестирање статистичких хипотеза Бентлер и Dudgeon (Bentler & Dudgeon, 1996) наводе да овакав тренд веома забрињава с обзиром на то да поверење истраживача у анализу структуре коваријанси за разумевање односа међу неексперименталним подацима расте.

Због тога су развијени ADF (asymptotic distribution free) методи да би се добили статистици неосетљиви на расподелу уз услов да је узорак велики (Hu *et al*., 1992). Статистици фитовања засновани на овој теорији су неосетљиви на одступања од мултинормалне дистрибуције у великим узорцима (Hu *et al*., 1992).

Рађен је већи број истраживања која су процењивала конзистентност индекса фитовања у различитим методима оцене (Hu & Bentler, 1998). Иако су аутори тих истраживања сугерисали различите препоруке, једно истраживање сумира резултате осталих. Динг, Велицер и Харлоу (Ding *et al.*, 1995, према: Hu & Bentler, 1998) извештавају да на све фит индексе, изузев на TLI, утиче метод оцењивања.

Ху, Бентлер и Кано (Hu, Bentler & Kano, 1992) наводе да су емпиријске студије са Монте Карло симулацијама показале да постоје оправдани разлози за сумњу у релевантност ADF теорије и израчунавање индекса подесности на основу такве претпоставке. Наиме, у арбитрарним дистрибуцијама адекватност модела може бити веома лоша уколико су узорци мали или број степени слободе велики (Hu & Bentler, 1998; Hu *et al*., 1992). Ху и Бентлер (Hu & Bentler, 1998) препоручују истраживачима употребу ML метода оцењивања, јер већина индекса заснованих на том методу даје боље процене фитовања модела него у случају индекса заснованих на GLS и ADF.

У оквиру ове области, развијена је и асимптотска теорија робустности којом су истраживачи покушали да установе услове под којима ML и GLS методи процене могу исправно евалуирати модел са подацима који одступају од нормалне расподеле (Hu & Bentler, 1998). Ови аутори наводе резултате низа истраживања која су показала да је то могуће под условом да су латентне варијабле (заједнички фактори или грешке), које иначе сматрамо некорелисаним, међусобно независне, а да када су заједнички фактори корелисани, они морају имати слободно процењене параметре варијанси и коваријанси (Hu & Bentler, 1998). Када су нормално дистрибуиране варијабле некорелисане, оне су и независне. Међутим, уколико варијабле немају нормалну дистрибуцију, то није нужно случај. У таквој ситуацији, параметри оцењени на основу ML и GLS метода оцене могу имати поуздане вредности, уколико постоји робустност у великим узорцима. Ипак, пошто не знамо какав је поступак процеса генерисања стварних података, не можемо да предвидимо да ли постоји независност фактора и грешака или самих грешака, па су практичне импликације асимптотске теорије робустности непознате (Hu & Bentler, 1998).

- Ефекат нарушавања нормалности и независности. Добри индекси фитовања би требало да буду осетљиви на погрешну спецификацију модела и стабилни кроз различите моделе процене, величине узорка и дистрибуције. У свом истраживању Ху и Бентлер (Hu & Bentler, 1998) су показали да у односу на друге изворе проблема који могу утицати на индексе фитовања, нарушавање мултинормалне расподеле има мањи утицај. Њихово мишљење је да индекси могу задржати висок генерални фит упркос постојању неких мањих погрешних спецификација.

Индекси фитовања који се најчешће наводе у истраживањима: карактеристике и услови примене

На основу прегледа литературе дошло се до закључка да се од већег броја индекса фитовања један број чешће примењује. Избор приказаних индекса је начињен на основу процене њиховог квалитета. Параметри у односу на које се дискутује њихова примена су најчешће утицај величине узорка и нарушавања мултинормане дистрибуције на њихове вредности.

»Индекс подесности модела« (The goodness-of-fit index GFI) Јореског и Сорбома је први фит индекс који је био стандардизован. Рачуна се према формули (Kline, 2005):

$$GFI = 1 - F_m/F_b, \qquad (1)$$

где је F_m фитинг функција опсервираног модела, а F_b фитинг функција нултог модела. Овај индекс говори о томе колико се опсервиране варијансе или коваријансе може објаснити моделом, што га чини аналогним индексу R^2 у мултиплој регресији (Hoyle, 2000). Вредност GFI = 1 говори о савршеном фиту модела, а вредности близу 0 о веома лошем фиту. Вредности веће од 0.90 сматрају се индикатором доброг фитовања модела (Kline, 2005). Мјулаик и сарадници (Mulaik *et al.*, 1989) наводе да пошто је GFI заснован на ML методу процене, он даје добре процене фита када су задовољени услови за његову примену (мултинормална дистрибуција) и када је узорак већи од 200 испитаника.

»Нормирани фит индекс« (Normed fit index – NFI) Бентлера и Бонета пореди χ^2 за тестирани модел и χ^2 за основни модел који претпоставља да су мерене варијабле у потпуности независне (Thompson, 2005). Вредност χ^2 за задати модел се пореди са χ^2 сатурисаног модела ($\chi^2 = 0$) и независног модела, а вредност индекса NFI говори колико су дискрепанца, односно χ^2 задатог модела, ближи сатурисаном моделу који савршено фитује у односу на независни модел који има најгори могући фит (Bentler & Bonett, 1980). Вредност NFI рачуна се према следећој формули (Bollen, 1990):

$$NFI = \Delta_1 = (F_b-F_m)/F_b, \qquad (2)$$

где је F_m фитинг функција евалуираног модела, а F_b је фитинг функција независног модела. Бентлер и Бонет (Bollen, 1990) овај индекс зову »нормирани« јер му је минимална вредност 0, а максимална 1. Аутори (Hu & Bentler, 1999; Miles & Shavlin, 2007; Mulaik, 2007; Thompson, 2005) сматрају да вредности статистика NFI близу 1, тј. веће од 0.95 показују добро фитовање подацима. У случају овог индекса, иако статистик који се процењује не мора да има неку одређену дистрибуцију, претпоставља се да је функција фитовања иста за оба модела (Hu & Bentler, 1998). Овај индекс не показује конзистентност у узорцима различите величине и од свих компаративних индекса фитовања најосетљивији је на нарушавање нормалне дистрибуције нарочито у малим узорцима. Такође, има тенденцију да буде потцењен (тј. да вредности буду неприхватљиво мале за моделе који добро фитују) када не постоји нормална дистрибуција података (DiLalla, 2000).

Боленов »Индекс инкременталног фитовања« (Incremental fit index – IFI или BL 89) рачуна се према формули (Bollen, 1990):

$$IFI= \Delta_2= (F_b-F_m)/ [F_b – df_m/(N-1)], \quad (3)$$

где су F_m и df_m фитинг функција и степени слободе за евалуирани модел, а F_b фитинг функција за независни модел. IFI се још зове и „ненормирани" јер његове вредности могу бити испод 0, тј. изнад 1. Вредности близу 1 су индикација доброг фитовања модела (Hu & Bentler, 1999). Индекс IFI показује конзистентност у узорцима различите величине (Dilalla, 2000). Неки аутори (DiLalla, 2000) сматрају да је у GLS методу оцене много боље користити индекс IFI него TLI.

На основу једначина (2) и (3) видимо да у рачунање NFI не улази величина узорка, док у рачунање IFI улази. Болен (Bollen, 1990) наводи резултате неколико истраживања у којима је Монте Карло симулацијама показано да су аритметичке средине дистрибуције узорковања индекса NFI веће у већим узорцима него у малим, а да је веза IFI и величине узорка близу 0. На основу тога, може се закључити да величина узорка не утиче директно на вредност NFI, али да аритметичка средина дистрибуције узорковања јесте повезана са величином узорка. У случају индекса IFI, величина узорка директно утиче на његову вредност, док аритметичка средина дистрибуције узорковања није повезана са величином узорка. Другим речима, може се десити да иако је модел добар, у малим узорцима, величина NFI буде мала и да то сугерише одбацивање модела (Bollen, 1990). Разлике између ова два индекса ће бити веће у мањим узорцима, а конвергираће ка 0 како N тежи ∞ (Bollen, 1990).

Боленов »Индекс релативног фитовања« (Relative fit index – RFI ili BL 86) рачуна се према следећем обрасцу (Bollen, 1990):

$$RFI = \rho_1 = 1\text{-}[(F_m/df_m)/(F_b\text{-}df_b)], \qquad (4)$$

где су F_m и df_m фитинг функција и степени слободе за евалуирани модел, а F_b и df_b фитинг функција и степени слободе за независни модел. Овај индекс има нормирани максимум који износи 1, али му минимум није нормиран на 0 (Bollen, 1990). Вредности близу 1 говоре да модел добро фитује подацима (Bollen, 1990; Hu & Bentler, 1998; Hu & Bentler, 1999).

»Такер-Луисов индекс« (Tucker-Lewis index TLI или eng. non-normed fit index NNFI) рачуна се на следећи начин (Bollen, 1990):

$$TLI = \rho_2 = [(F_b/\ df_b) - (F_m/df_m)]/\ [(F_b\text{-}df_b) - (1/(N\text{-}1))], \qquad (5)$$

где су F_m и df_m функција дискрепанце и степени слободе за евалуирани модел, а F_b и df_b функција дискрепанце и степени слободе за независни модел. Вредности близу 1 говоре о добром фитовању евалуираног модела, али овај индекс није нормиран на опсег 0-1 (Bentler, 1990). Овај индекс даје најбоље резултате у ML методу оцене, док у GLS даје најлошије. Такође, погодан је за поређење хијерархијских модела (DiLalla, 2000).

Као и у случају претходног пара фит индекса ((2) и (3)), Болен (Bollen,1990) наводи резултате неколико истраживања у којима је на основу аналогног закључивања и изведених симулација показано да величина узорка не утиче на рачунање индекса RFI, али да утиче на рачунање TLI. Величина узорка је у вези са дистрибуцијом узорковања RFI, али није у вези (или је у јако слабој вези) са аритметичком средином дистрибуције узорковања TLI. Другим речима, индекс TLI има велику предност што омогућава доста добру процену фита у узорцима различите величине. Као и код индекса (2) и (3), у случају индекса (4) и (5), разлика њихових вредности ће тежити 0, како N тежи ∞ (Bollen, 1990).

Претходно наведено омогућава да сагледамо предности и мане индекса NFI, RFI, IFI, TLI. Предности индекса NFI и RFI су у томе што имају нормирани максимум, што олакшава интерпретацију. Међутим, њихове вредности имају тенденцију пораста у великим узорцима, што интерпретацију резултата добијену у узорцима различите величине чини опасном. Такође, процена фита модела у мањим узорцима може бити превише песимистична. С друге стране, индекси IFI и TLI су стабилни у узорцима различите величине, али њихове вредности нису нормиране. Болен (Bollen, 1990) због тога препоручује да се у извештајима о фитовању модела наводе и једне и друге мере, како би се стекао комплетнији увид и извели поузданији закључци.

Бентлеров »Индекс компаративног фитовања« (Comparative fit index – CFI), као и NFI, процењује фитовање модела релативно у односу на независни модел. Код индекса који нису централизовани (као што су CFI, RMSEA, RNI), χ^2 = df представља савршени фит (за разлику од централизованих где је χ^2=0 савршени фит модела). Приликом рачунања нецентрализованих параметара, као чинилац се користи разлика хи-квадрата и степени слободе (χ^2 – df). Вредност CFI добија се по следећој формули (Bentler, 1990):

$$CFI= 1\text{-}max\ [(T_m\text{-}df_m),0]/max\ [(T_m\text{-}df_m),\ (T_b\text{-}df_b),\ 0], \qquad (6)$$

Бентлер (Bentler, 1990: 11) каже да „изгледа да је CFI најбољи индекс. Има опсег од 0 до 1, малу узорачку варијабилност и није централизован“. Овај индекс показује прецизност у различитим методима оцене, конзистентност у узорцима различите величине и користан је за поређење хијерархијских модела (DiLalla, 2000). Као и код претходно наведених параметара, вредности блиске 1 (тј. веће од 0.95) су пожељне и сматрају се индикатором доброг фита (Hu & Bentler, 1999; Miles & Shavlin, 2007; Mulaik, 2007; Thompson, 2005).

»Релативни нецентрализовани индекс« (Relative noncentrality index – RNI) МекДоналда и Марша рачуна се на следећи начин (Hu & Bentler, 1999; McDonald & Marsh, 1990):

$$RNI= [(T_b\text{-}df_b) – (T_m\text{-}df_m)]\ /\ (T_b\text{-}\ df_b), \qquad (7)$$

где су F_m и df_m дискрепанца и степени слободе за евалуирани модел, а F_b и df_b дискрепанца и степени слободе за независни модел. Овај индекс показује конзистентност у узорцима различите величине (Hu & Bentler, 1998). Вредности близу 1 говоре о добром фитовању евалуираног модела, али овај индекс није нормиран на опсег 0-1 (Bentler, 1999).

»Квадратни корен просечне квадриране грешке апроксимације« (Root-mean-square error of approximation – RMSEA) процењује колико добро ће се понашати параметри у репродукованим коваријансама популације, односно колико добро се модел може одржати у популацији (Schutz, 1998). Модел за који је процењено да ће репродуковати коваријансе популације ће имати вредност 0. RMSEA процењује количину грешке апроксимације за одређени број степени слободе модела и узима у обзир величину узорка. Другим речима, RMSEA показује колики је ступањ дискрепанце између претпостављене и емпиријске матрице коваријанси по једном ступњу слободе (Hoyle, 2000). RMSEA се рачуна према обрасцу (Hu & Bentler, 1999):

$$\text{RMSEA}= (F_o/df_m)^{-1/2}\text{, gde je } F_o= \max\,[(T_m\text{-}df_m)/N\text{-}1),\ 0], \qquad (9)$$

За разлику од NFI и CFI, вредности RMSEA би требало да буду што ближе 0 како би могли да сматрамо да модел фитује подацима. Вредности испод 0.06 се узимају као добре (Hu & Bentler, 1999; Miles & Shevlin, 2007; Thompson, 2005). Вредности у опсегу од 0.06 до 0.08 се узимају као прихватљиве, док се вредности ≥ 0.1 узимају као индикатор лошег фита (Cudeck, 2000; Kline, 2005). У исписима и извештајима се користи 90% интервал поузданости за RMSEA. Ако је доња граница интервала поузданости RMSEA мања од 0.05, нећемо одбацити нулту хипотезу. Такође, ако горња граница интервала поузданости прелази горњу граничну вредност (≥ 0.08), сматрамо да модел не фитује и одбацујемо нулту хипотезу (Kline, 2005). Овај индекс мери апсолутни фит модела, али је осетљив на недостатак парсимоније (Dilalla, 2000).

Бенлтеров SRMR – Стандардизовани квадратни корен просечног квадрата резидуала (standardized root mean square residual) је стандардизовани индекс RMR (root-mean-square residual) Јореског и Сорбома и представља меру укупне разлике између опсервираних и предвиђених корелација. Рачуна се према формули (Hu & Bentler, 1999):

$$\sqrt{\left\{2\sum_{i=1}^{p}\ \sum_{j=1}^{i}\left[(s_{ij} - \sigma_{ij})/(s_{ii}\,s_{jj})\right]^2\right\} / \, p(p+1)}\,, \qquad (10)$$

где су s_{ij} опсервиране коваријансе, s_{ii} и s_{jj} опсервиране стандардне девијације, p број опсервираних варијабли, а σ_{ij} репродуковане коваријансе. Као и у случају индекса RMSEA, вредности индекса SRMR би требало да буду што ближе 0. Вредности испод .10 се сматрају прихватљивим (Hu & Bentler, 1999; Kline, 2005).

Како изабрати »прави« фит индекс: које индексе је најбоље приказати у извештају о истраживању?

Ху и Бентлер (Hu & Bentler, 1998; Hu & Bentler, 1999) у својим веома опсежним истраживањима о граничним критеријумима индекса фитовања наводе да већина индекса добијених ML методом оцене показује боље резултате него индекси добијени GLS или ADF методом оцене. Такође наводе да индекси NFI, RFI, GFI, CN, AGFI, CAK и CK нису добри и да би требало избегавати процењивање подесности модела помоћу њих (Hu & Bentler, 1998). Ови аутори (Hu & Bentler, 1998; Hu & Bentler, 1999) наводе да је индекс SRMR заснован на ML методу процене најосетљивији на погрешну спецификацију факторских коваријанси или латентних структура, а да су TLI, IFI, RNI, CFI, RMSEA, GFI, Mc засновани на истом методу процене најосетљивији на моделе са погрешно специфицираним факторским засићењима. На основу матрице корелација добијене између индекса фитовања оцењених ML методом како би се проценило који индекси ће се понашати слично у зависности од величине узорка, дистрибуције и погрешне спецификације модела, добијена су два кластера корелисаних фит индекса. У једном кластеру су се налазили NFI, RFI, GFI, AGFI, CAK и CK са високим интеркорелацијама. У другом кластеру су се налазили TLI, IFI, RNI, CFI, Mc и RMSEA са такође високим интеркорелацијама. Индекс SRMR се није понашао као остали ML засновани индекси фитовања. На основу овога Ху и Бентлер (Hu & Bentler, 1999) закључују да је приликом писања извештаја о подесности модела најбоље користити "двоиндексну стратегију", односно у извештају навести вредност ML заснованог индекса SRMR и неким од ML заснованих индекса TLI, IFI, RNI, CFI, GFI, Mc или RMSEA, како би се разликовали добри модели од лоших, укључујући моделе са погрешно специфицираним факторским коваријансама и/или факторским засићењима. Ху и Бентлер (Hu & Bentler, 1998) наводе да у малим узорцима (≤250) вредности индекса TLI, RMSEA и Mc могу навести истраживача на погрешне закључке, те стога предлажу обазривост у закључивању у таквим ситуацијама и наводе да су Марш, Бала и Хау (Marsh *et al*., 1996; према: Hu & Bentler, 1998) направили нормирану верзију индекса TLI који је погодан за мале узорке.

Када се у структуралном моделовању за процену модела користи ML или GLS метод, Ху и Бентлер (Hu & Bentler, 1998) сматрају да је потребно навести вредности индекса TLI, IFI, RNI, CFI, RMSEA, SRMR и друге, с тим што су TLI и RMSEA мање пожељни када је реч о малим узорцима. Њихово истраживање је показало да су ови индекси добри, јер су умерено осетљиви на погрешну спецификацију једноставних модела (тј. на погрешно спецификоване факторске коваријансе) и више осетљиви на погрешну спецификацију комплексних модела (тј. на погрешно спецификована факторска засићења). Такође, њихово истраживање је мање осетљиво на дистрибуцију и величину узорка на коме се брши истраживање.

Међу индексима чију употребу Ху и Бентлер препоручују су и Mc и RMSEA засновани на ML и GLS методима (Hu & Bentler, 1998). У случају мањих узорака, ови аутори препоручују употребу ових индекса заснованих на ML методу, јер GLS и ADF засновани индекси потцењују њихове популацијске вредности и имају много већу варијансу у малим узорцима.

У случају процене модела путем ADF, предлажу SRMR, TLI, IFI, RNI и CFI (Hu & Bentler, 1998). Ипак, ови аутори сматрају да већина индекса заснованих на ML методу дају боље показатеље у односу на индексе добијене путем GLS или ADF, јер тако процењени индекси у мањим узорцима потцењују популацијске вредности и имају много већу варијансу у таквој ситуацији, на основу чега истраживачи могу погрешно да одбаце тестиране моделе.

Ху и Бентлер наводе да је за поуздано процењивање фитовања модела неопходно користити више индекса подесности, са чиме се слажу и други аутори (Bentler, 2007; Hu & Bentler, 1998; Markland, 2007; Schutz, 1998; Thompson, 2005).

Индекс чија се употреба препоручује без обзира на методу процене је SRMR, с тим што се препоручује ML метода процене када је реч о мањим узорцима (≤250). Ху и Бентлер (1998) наводе да би извештај требало да садржи индекс SRMR и још један од TLI, IFI, RNI, CFI, Mc или RMSEA индекса.

Поставља се и питање која је вредност индекса фитовања прихватљива, јер употреба индекса фита за евалуацију модела, намеће и проблем критеријума, односно доње/горње границе индекса која се сматра прихватљивом. Адекватни гранични, »cut-off« критеријум би требало да има за последицу минималне грешке типа I и типа II (Hu & Bentler, 1999). Иако смо претходно навели прихватљиве граничне вредности које су предвидели аутори индекса, потребно је рећи да је опсежна анализа коју су спровели Ху и Бентлер (Hu & Bentler, 1999) потврдила да за индексе TLI, IFI, RNI и CFI доња гранична вредност износи 0.95, за SRMR горња гранична вредност износи 0.08, а за RMSEA горња гранична вредност износи 0.06.

Закључак

Моделовање структуралним једначинама представља важан део савремене методологије истраживања у разним пољима психологије и другим друштвеним наукама. Истраживања указују на то да за адекватну и ефикасну примену SEM-а у процени подесности модела, веома важан ослонац представљају индекси подесности (индекси фитовања) који омогућавају веома прецизну квантификацију (Hu & Bentler, 1999). Бројност развијених индекса поставила је захтеве избора оних који су најквалитетнији и најадекватнији за примену. Од 21 индекса који су поменути у овом раду, детаљније је разматрано 10 који се по мишљењу истраживача сматрају квалитетнијим (хи-квадрат, GFI, NFI, IFI, RFI, TLI, CFI, RNI, RMSEA, SRMR). Међутим, за њихову најефикаснију примену важно је размотрити услове у којима се ови фит индекси примењују. Поузданост процене помоћу индекса фитовања је у великој мери зависна од формалних карактеристика узорка (као што су његова величина, дистрибуција података) и/или од карактеристика тестираног модела. Поред тога, за ефикасност примене одабраних индекса, потребно је узети у обзир и граничне вредности које се сматрају прихватљивим.

Истраживач се суочава и са проблемом избора индекса које ће приказати у резултататима истраживања у којима су теоријски модели тестирани помоћу SEM-а. Једна од предложених стратегија је да се у интерпретацији резултата ослонимо на већи број индекса имајући у виду напред наведене критеријуме, а да у извештају о истраживању, следећи предложену »двоиндексну стратегију« (Hu i Bentler, 1998), поред вредности хи-квадрата, прикажемо SRMR и један од фит индекса из групације инкременталних или апсолутних индекса.

Коришћена литература

Anderson, J.C. & D.W. Gerbing (1988): Structural equation modeling in practice: a review and recommended two-step approach, *Psychological Bulletin*, 103*,* 411-423.

Barrett, P. (2007): Structural equation modeling: adjudging model fit, *Personality and Individual differences*, 42*,* 815-824.

Bentler, P.M. (2007): On tests and indices for evaluating structural models, *Personality and Individual Differences*, 42, 825-829.

Bentler, P.M. (1990): Comparative fit indexes in structural models, *Psychological Bulletin*, 107, 238-246.

Bentler, P.M. & D.G. Bonett (1980): Significance tests and goodness of fit in the analysis of covariance structures, *Psychological Bulletin*, 88, 588-606.

Bentler, P.M. & P. Dudgeon (1996): Covariance structure analysis: statistical practice, theory and directions, *Annual Review of Psychology*, 47, 541-570.

Bentler, P.M. & A. Mooijaart (1989): Choice of structural model via parsimony: a rationale based on precision, *Psychological Bulletin*, 106, 315-317.

Bollen, K.A. (1990): Overall fit in covariance structure models – two types of sample size effects, *Psychological Bulletin*, 107, 256-259.
Byrne, B.M. (1989): A primer of LISREL: *Basic applications and programming for confirmatory factor analysis.* New York: Springer-Verlag
Cudeck, R. (2000): Exploratory factor analysis; in H.E.A. Tinsley & S.D. Brown (eds.): *Applied multivariate statistics and mathematical modleing* (265-296). San Diego: Academic Press.
DiLalla, L.F. (2000): Structural equation modeling: uses and issues; u H.E.A. Tinsley & S. D. Brown (eds.): *Applied multivariate statistics and mathematical modleing* (pp. 439-464). San Diego: Academic Press.
Fajgelj S. (2004): *Metode istraživanja ponašanja*. Beograd: Centar za primenjenu psihologiju.
Fassinger R.E. (1987): Use of structural equation modeling in counseling psychology research, *Journal of Cognitive Psychology*, 34, 425–436.
Hoyle, R.H. (2000): Confirmatory factor analysis; in H.E.A. Tinsley & S.D. Brown (eds.): *Applied multivariate statistics and mathematical modeling* (pp. 465-497). San Diego: Academic Press.
Hu, L. & P.M. Bentler (1998): Fit indices in covariance structure modeling: sensitivity to underparameterized model misspecification*, Psychological Methods*, 3, 424-453.
Hu, L. & P.M. Bentler (1999): Cutoff Criteria for Fit Indices in Covariance Structure Analysis: Conventional Criteria versus New Alternatives, *Structural equation modelling*, 6, 1-55.
Hu, L., P.M. Bentler & Y. Kano (1992): Can test statistics in covariance structure analysis be trusted? *Psychological Bulletin*, 112, 351-362
Jöreskog, K.G. (1978): Structural analysis of covariance and correlation matrices, *Psychometrika*, 43, 443-477.
Kline, R.B. (2005): *Principles and practice of structural equation modeling.* New York: The Guilford Press.
La Du, T.J. & S.J. Tanaka (1989): The influence of sample size, estimation method and model specification on goodness-of-fit assessments in structural equation models, *Journal of Applied Psychology*, 74, 625-636.
Lazarević, Lj.B. (2007a): Kognitivna struktura budućih trenera i nastavnika iz ugla kibernetskog modela. U zborniku radova *Međunarodna naučna konferencija Analitika i dijagnostika fizičke aktivnosti* (325-331). Beograd: Fakultet sporta i fizičkog vaspitanja i Olimpijski komitet Srbije.
Lazarević, Lj.B. (2007b): *Kibernetički model K. Momirovića i Katel-Hornova teorija inteligencije kao teorijski modeli faktorske strukture baterije za ispitivanje intelektualnih sposobnosti KOG 9.* Master rad. Beograd: Filozofski fakultet.
Markland, D. (2007): The golden rule is that there are no golden rules: a commentary on Paul Barrett's recommendations for reporting model fit in structural equation modelling, *Personality and Individual Difference*, 42*,* 851-858.
Marsh, H.W., J.R. Balla & R.P. McDonald (1988): Goodness-of-fit indexes in confirmatory factor analysis – the effect of sample size, *Psychological Bulletin*, 103, 391-410.
McDonald, R.P. & H.W. Marsh (1990): Choosing a multivariate model: noncentrality and goodness-of-fit, *Psychological Bulletin*, 107, 247-255
McIntosh, C.N. (2007): Rethinking fit assessment in structural equation modelling: a commentary and elaboration on barrett, *Personality and individual difference*, 42, 859-867.
Micceri, T. (1989): The unicorn, the normal curve and other improbable creatures, *Psychological Bulletin*, 105, 156-166.
Miles, J. & M. Shevlin (2007): A time and place for incremental fit indices, *Personality and Individual Difference*, 42, 869-874.
Mulaik, S.A., L.R. James, J.V. Alstine, N. Bennett, S. Lind & C.D. Stilwell (1989): Evaluation of goodness-of-fit indices for structural equation models, *Psychological Bulletin*, 105, 430-445.

Mulaik, S. (2007): There is a place for approximate fit in structural equation modelling, *Personality and Individual Difference*, 42, 883-891.
Novović, Z., P. Čolović, M. Babić i G. Mišić-Pavkov (2006): Struktura kliničke i staračke depresivnosti: sličnosti i razlike, *Psihologija*, 39, 425-437.
Schutz, R.W. (1998): Assessing the stability of psychological traits and measures; in Duda, J. L. (eds.): *Advances in sport and exercise psyhology measurement* (393-408). West Virginia University.
Smederevac, S., P. Čolović, D. Mitrović, Ž. Nikolašević i B. Đekić (2006): Nasledni i sredinski činioci dimenzija Ajzenkovog PEN i alternativnog petofaktorskog modela ličnosti, *Psihologija*, 39, 407-423.
Thompson, B. (2005): *Exploratory and confirmatory factor analysis – understanding concepts and applications.* Washington DC: American Psychological Association.
Vernon, T. & S. Eysenck (2007): Introduction, *Personality and individual difference*, 42.
Wright, D.B. (2006): The art of statistics: a survey of modern techniques; in P.A. Alexander & P.H. Winne (eds.): *Handbook of educational psychology* (879-901). New Jersey: Lawrence Erlbaum Associates

Примљено 27.02.2008; прихваћено за штампу 24.04.2008.

Ljiljana Lazarevic

APPLICATION OF FIT INDICES IN TESTING THE THEORETICAL MODELS IN PSYCHOLOGY: POSSIBILITIES AND LIMITATIONS

Abstract

This paper deals with the fit indices used in Structural Equation Modelling (SEM) for testing theoretical models and the difficulties that can occur during the testing of theoretical models in different fields of psychology. The paper discusses the basic assumptions of SEM and presents the indices used for assessing the fit of theoretical models. This paper also presents the procedures for calculating the basic statistic for assessing the fit of models (χ^2), as well as for calculating the most commonly used fit indices, in order to gain a better insight into the advantages and potential difficulties that can occur during their usage. We mention the difficulties regarding the assessment of fit of the model based on χ^2 and the discussed fit indices stemming from the sample size, data distribution and assessment methods, wrong specification of model and disturbance of normality and independence of latent variables, as well as the ways in which these difficulties can be overcome. This paper provides a proposal for the approach to presenting the fit indices in reports on studies where theoretical models were tested via SEM.
Key words: structural equation modelling, fit of theoretical models, fit indices, psychology.

Лиляна Лазаревич

ПРИМЕНЕНИЕ ИНДЕКСА ПРИГОДНОСТИ В ТЕСТИРОВАНИИ ТЕОРЕТИЧЕСКИХ МОДЕЛЕЙ В ПСИХОЛОГИИ: ВОЗМОЖНОСТИ И ОГРАНИЧЕНИЯ

Резюме

В работе рассматриваются индексы пригодности, используемые в моделировании при помощи структуральных уравнений (SEM) в целях тестирования теоретических моделей, а также трудности, которые могут возникнуть при тестировании теоретических моделей в различных областях психологии. Рассмотрены основные положения SEM и представлены индексы оценки пригодности теоретических моделей. В работе излагаются и приемы расчета основного статистика для оценки пригодности моделей (χ^2), а также наиболее часто применяющихся индексов пригодности, в целях более четкого осознания преимуществ и потенциальных трудностей, которые могут возникнуть при их использовании. Приводятся трудности в оценке пригодности моделей на основании χ^2 и рассматриваемых индексов пригодности, вытекающие из масштабов корпуса, распределения данных и методов оценки, ошибочной спецификации моделей и нарушения нормальности и независимости латентных вариабл, а также способы преодоления указанных трудностей. В работе предлагается подход в презентации индекса пригодности в сообщениях об исследованиях, в которых теоретические модели тестировались при помощи SEM.

Ключевые слова: моделирование структуральными уравнениями, пригодность теоретических моделей, индексы пригодности, психология.