

**Aleksandar Vasić i
Dušana Šarčević¹**

**Katedra za poslovnu
psihologiju,
Visoka škola
„Pravne i poslovne
akademske studije
dr Lazar Vrkatić”,
Novi Sad**

OD ALFE DO OMEGE I NATRAG

Koeficijent α , poznat i kao koeficijent interne konzistencije, jedna je od najčešće korišćenih i istovremeno najviše zloupotrebljivanih procena pouzdanosti kompozitnih mernih instrumenata ili testova. Za tumačenje rezultata dobijenih primenom nekog psihološkog testa najvažnija zloupotreba koeficijenta α je pogrešno zaključivanje o njegovoj zadovoljavajućoj pouzdanosti i homogenosti ili jednodimenzionalnosti. Jedan izvor ovakve vrste zloupotrebe proizilazi iz same definicije koeficijenta α kao procene pouzdanosti tipa interne konzistencije. Drugi razlog ovakvoj zloupotrebi nalazi se u nedovoljnom poznavanju razlika u poimanju i definicijama pouzdanosti, homogenosti i valjanosti između klasičnog modela merenja i njegove usavršene verzije, faktorsko-analičkog modela merenja. Zbog toga su već duže u opticaju i alternativni predlozi, kao što je koeficijent ω koji predstavlja bolju procenu pouzdanosti, homogenosti i valjanosti pod Spearmanovim konfirmatornim jednofaktorskim modelom. Na osnovu podataka nastalih primenom jednog instrumenta putem elementarne algebre, demonstrirane su i upoređene procene koeficijenata α i ω . Rezultati ovih, kao i drugih analiza, govore da je povoljna procena α neophodan, ali ne i dovoljan razlog za zaključak o pouzdanosti i homogenosti testa. Da bi se o tome moglo utemeljeno zaključivati, potrebne su dodatne procene kao što je ω , zajedno sa procenama saglasnosti modela sa podacima, pogotovo u onim, za ovakvu vrstu instrumenata, čestim slučajevima koji zahtevaju proveru jednofaktorskog modela merenja.

Ključne reči: testovi, pouzdanost, homogenost, valjanost, faktorska analiza

1 Adresa autora:
dusanasarcevic@gmail.com

Primljeno: 05. 06. 2013.
Primljena korekcija:
09. 07. 2013.
Prihvaćeno za štampu:
17. 07. 2013.

– *Da li si odgonetnula zagonetku? – upita Šeširdžija, okrenuvši se opet Alisi.*
 – *Nisam, odustajem – odgovori Alisa. – Koji je odgovor?*
 – *Nemam pojma – reče Šeširdžija.*
 – *Ni ja – reče Martovski Kunić.*
Alisa pogleda oštro. – Mislim da biste mogli da pametnije provodite vreme
 – *reče ona – nego da ga traćite na postavljanje zagonetki koje nemaju*
odgovore.
 (Carroll, 1865, pp. 82–83)

U naučnom saznanju, kao i u svakodnevnom životu, od vitalnog značaja su očekivanja ili pretpostavke o stvarnosti. Ali, za razliku od zdravog razuma, naučno saznanje, i na njemu zasnovano racionalno delovanje, počiva na pretpostavkama koje uživaju relativnu empirijsku podršku. O relativnoj empirijskoj potkrepljenosti pretpostavki zaključuje se na osnovu evidencije koja je nastala primenom naučnih instrumenata. U psihološkim istraživanjima i praksi značajnu ulogu u generisanju podataka imaju kompozitni merni instrumenti ili testovi. Ovi instrumenti razvijaju se od začetka moderne psihologije zajedno sa teorijskim i matematičko-statističkim modelima, kao i iz njih izvedenim modelima merenja. Zato su testovi ne samo ambasadori teorija (Cattell, 1973), nego i matematičko-statističkih modela i na njima zasnovanih modela merenja. Dakle, osnovne gradivne komponente psiholoških modela su teorijski i matematičko-statistički modeli, kao i modeli merenja (Anderson, 2003; Maxwell & Delaney, 1990; Myers & Well, 2008; Tinsley & Brown, 2000). Svaka delatnost utemeljena na psihološkim saznanjima i veštinama, koja je usmerena na rešavanje problema u ličnom i društvenom prilagođavanju, uslovljena je ovim užim komponentama. Ali, ova složenost i uslovljenost nije dovoljno prepoznata u naučnoj i stručnoj psihološkoj javnosti, o čemu govori i prigovaranje o nedovoljnom poznavanju matematičko-statističkih modela i na njima zasnovanih modela merenja (Fajgelj, 2009; McDonald, 1985, 1999; Miles & Shevlin, 2001; Sijtsma, 2009). To je, verovatno, i jedan od osnovnih razloga što su ti modeli često pogrešno korišćeni uprkos naučno-stručnoj deklaraciji. Međutim, mora se priznati i da je malo izvora, sa časnim izuzecima (Loehlin, 2004; McDonald, 1985, 1999; Pedhazur & Pedhazur Shmelkin, 1991), u kojima su na dovoljno jasan i pristupačan način opisane i obrazložene međuzavisnosti teorijskih i matematičko-statističkih modela, kao i modela merenja. Stoga je namera ovog teksta da se doprinese razvoju svesti o složenosti naučnih modela i demonstriraju neke važne pretpostavke u okviru najčešće primenjivanih modela merenja.

U opštem, naučni modeli se mogu prikazati kao:

$$D = M + E \tag{1}$$

gde je stvarnost predstavljena podacima (D), dok stvarnost, tako kako je predstavljena, pokušavamo da opišemo, razumemo i objasnimo nekim modelom (M) (Miles & Shevlin, 2001). Naučno saznanje je tentativno i aproksimativno i zato

svaki naučni model manje ili više uspešno predstavlja stvarnost podrazumevajući uvek neku grešku (E).

Modeli merenja se takođe mogu opisati kao i naučni modeli:

$$R_m = M_m + E_m \quad (2)$$

gde je stvarnost predstavljena podacima generisanim primenom mernog instrumenta, u psihologiji obično primenom testa (R_m). Rezultati na testu sadrže deo koji se zaista odnosi na stvarnost ili pravi rezultat merenja (M_m) i, kao što je uvek slučaj sa operacijama merenja, deo koji se odnosi na grešku merenja (E_m). Analogno, koristeći pojmove iz teorije informacija, pravi rezultat predstavlja signal dok je greška merenja šum koji ugrožava prijem signala kao nosioca informacije. Razlike između modela merenja u psihologiji proističu iz specifičnosti matematičko-statističkih modela koji se koriste u definisanju pravog rezultata i greške merenja (Fajgelj, 2009; McDonald 1999).

Klasični model merenja, model pravog i rezultata greške ili klasična teorija testa (Bukvić, 1996; Crocker & Algina, 1986; Fajgelj, 2009; Gulliksen, 1950; Lord & Novick, 1968; McDonald, 1999), nije ništa drugo nego dosledna primena opšteg modela merenja:

$$Y = T + E \quad (3)$$

gde je rezultat na testu (Y) sačinjen iz dela koji se odnosi na pravi rezultat (T) i dela nastalog nasumičnim delovanjima koja dovode do toga da ukupan rezultat odstupa od pravog rezultata (E). S obzirom na to da su testovi složeni, strukturirani i kompozitni instrumenti koji su sačinjeni od većeg broja delova ili stavki, iz čijih se rezultata izvode ukupni rezultati, ovaj opšti izraz se primenjuje i na nivou stavki. Bliskost sa opštim shvatanjem merenja bila je povod za mnoge kritike koje su upućene klasičnom modelu u psihologiji. Razmene povodom tih kritika donosile su različita shvatanja ili teorije merenja, počev od Fergusonove komisije i Campbella preko klasifikacije nivoa merenja sa primerenim matematičko-statističkim modelima S. S. Stevensa do teorije stavskog odgovora (IRT: Crocker & Algina, 1986; Fajgelj, 2009; McDonald, 1999).

Istovremeno sa razvojem klasične teorije testa nastao je i faktorsko-analitički model merenja (Spearman, 1904, 1910, 1927). Za ovaj model tvrdi se da je bolja ili usavršena verzija klasičnog modela merenja (Bollen, 1989; Fajgelj, 2009; Jöreskog, 1971; McDonald, 1999; Momirović, Wolf i Popović, 1999). Naime, u faktorsko-analitičkom modelu merenja dosledno je primenjen matematičko-statistički model u definisanju osnovnih komponenata rezultata merenja kao:

$$X_j = \mu_j + \lambda_j F + \epsilon_j \quad (4)$$

gde je X_j rezultat na pojedinoj stavci, μ_j težište raspodele rezultata koje određuje metriku, dok je $\lambda_j F$ deo rezultata koji je objašnjiv predmetom merenja, odnosno faktorom (F), pri čemu je veličina ovog dela rezultata eksplicitno procenjena kao (λ_j). Greška merenja u ovom modelu (ϵ_j) definisana je kao unikvitet ili, alternativ-

no, kao specifičnost i greška merenja (Green & Yang, 2009; Lucke, 2005; McDonald, 1999; Momirović i sar., 1999; Mulaik, 1972).

Svi rezultati, nastali primenom nekog testa, mogu se predstaviti uređenim poljem brojeva ili matricom kao u Tabeli 1 u kojoj kolone predstavljaju m -ti uzorak varijabli iz nekog definisanog univerzuma M sa svim mogućim varijablama. Te varijable su u psihološkoj istraživačkoj i praktičnoj delatnosti najčešće rezultati na stavkama ili ukupni rezultati na testovima kao aproksimacije latentnih varijabli ili faktora, odnosno ponašanja, mišljenja i osećanja obuhvaćenih nekim konstruktima (Bollen, 1989, 2002; McDonald, 1999; Royce, 1963). Redovi su ispitanici iz n -tog uzorka ispitanika iz populacije N koji su opisani m -tim uzorkom varijabli. U ćelijama tabele su rezultati ispitanika, odnosno odgovori i -tog ispitanika na j -toj varijabli ili stavki (x_{ji}). Psihološko interesovanje je usmereno na redove, iako su za razumevanje informacija sadržanih u takvoj matrici značajne i kolone. Zato je važan korektiv takvog psihološkog zastranjivanja psihometrijsko usredsređivanje na kolone tabele, te otuda i redosled oznaka u označavanju izvora varijabiliteta.

Tabela 1

Početna matrica podataka nastalih primenom nekog testa

Ispitanici	Stavke							
	X_1	X_2	X_3	X_4	...	j	...	m
1	X_{11}	X_{21}	X_{31}	X_{41}	
2	X_{12}	X_{22}	X_{32}	X_{42}	
3	X_{13}	X_{23}	X_{33}	X_{43}	
4	X_{14}	X_{24}	X_{34}	X_{44}	
.				
.				
.				
i	X_{1i}	X_{2i}	X_{3i}			X_{ji}	...	
.				
.				
.				
n								

Da bi na empirijski utemeljen način moglo da se zaključuje o kvalitetu modela i njegovoj saglasnosti sa podacima, polazna matrica sa rezultatima se sažima u matricu takvog oblika da su u njoj sačuvane sve informacije o početnim rezultatima. Uobičajeno je da se takvo sažimanje podataka obavi primenom matematičko-statističkih modela sa jednom i dve varijable (Myers & Well, 2008). Matrica dobijena ovakvom redukcijom (Tabela 2) je od posebne važnosti u razmatranjima empirijske zasnovanosti modela merenja, te tako i teorijskog modela od kojeg se pošlo u izradi i proveru nekog testa. Model sa jednom varijablom je predstavljen procenama centralne tendencije (AS_j) i disperzije (σ_{jj} i SD_j) u glavnoj dijagonali i u podnožju, a model sa dve varijable procenama odnosa između varijabli (σ_{jk} i r_{jk})

kao vandijagonalnim elementima. To znači da nova matrica sadrži procene koje govore o metrici, varijabilitetu i kovarijabilitetu stavki iz nekog testa. Kovarijabilitet, procenjen na osnovu izvorne metrike, sadržan je u matrici (Σ) sa varijansama varijabli u glavnoj dijagonali i kovarijansama kao vandijagonalnim elementima. S obzirom na to da je metrika varijabli u psihološkim i srodnim merenjima po pravilu nedefinisana, uz varijanse i kovarijanse u matrici, izlažu se i procene centralne tendencije i disperzije. U standardizovanoj metrici, kada su varijable skalirane na nulu i jedinicu, reč je o matrici (R) korelacija stavki koja u glavnoj dijagonali sadrži maksimalne korelacije svake varijable sa samom sobom (Brown, 2006; Loehlin, 2004; Mulaik, 1972).

Tabela 2

Varijanse (σ_{jj}), kovarijanse (σ_{jk}), korelacije (r_{jk}) i procene parametara centralne tendencije (AS_j) i disperzije (SD_j) rezultata na stavkama nekog testa primenjenog na uzorku od n ispitanika

Stavke	X_1	X_2	X_3	X_4	...	j	...	m
X_1	σ_{11}	r_{21}	r_{31}	r_{41}	...			
X_2	σ_{12}	σ_{22}	r_{32}	r_{42}	...			
X_3	σ_{13}	σ_{23}	σ_{33}	r_{43}	...			
X_4	σ_{14}	σ_{24}	σ_{34}	σ_{44}	...			
					...			
.				
.				
.	σ_{jj}		
k	σ_{1k}	σ_{2k}	σ_{3k}		...	σ_{jk}	...	
.				
.				
.				
m								
	AS_1	AS_2	AS_3	AS_4		AS_j		AS_m
	SD_1	SD_2	SD_3	SD_4		SD_j		SD_m

Matrica varijansi i kovarijansi ili korelacija je polazište u potrazi za odgovorima na pitanje o empirijskoj zasnovanosti mernog modela nekog testa. Tokom razvoja modela merenja u psihologiji iskristalisala su se pitanja koja se odnose na osnovne kvalitete testova kojima su podaci o stvarnosti generisani (Bukvić, 1996; Crocker & Algina, 1986; McDonald, 1999) ili, po alternativnom stanovištu, samih podataka pre nego testova i njegovih stavki (Fajgelj, 2009; Thompson & Wachhaase, 2000). Kada je reč o klasičnom i faktorsko-analitičkom modelu merenja, tri važna i bliska pitanja odnose se na pouzdanost, homogenost ili jednodimenzionalnost i na valjanost psiholoških testova (Bollen, 1989; Crocker & Algina, 1986;

Fajgelj, 2009; McDonald, 1999). Tentativni odgovori na ova pitanja su zasnovani na proverama modela merenja i služe u zaključivanju o empirijskoj zasnovanosti teorijskih modela i naučno utemeljnom stručnom delovanju (Cronbach, 1951; Cronbach & Meehl, 1955; Fiske, 1971; Guttman, 1945; Momirović, 1998a).

Tokom razvoja klasičnog i faktorsko-analitičkog modela merenja u psihologiji izvedeni su različiti postupci za procenu pouzdanosti, homogenosti i valjanosti (Crocker & Algina, 1986; Fajgelj, 2009; McDonald, 1999; Momirović i sar., 1999). Jedna od tih procena dobija se preko koeficijenta pouzdanosti α tipa interne konzistencije pod klasičnim modelom merenja. Ova procena postala je gotovo neizbežna u izveštajima o rezultatima fundamentalnih i praktičnih istraživanja, ali i predmet brojnih kritičkih preispitivanja (Cortina, 1993; Gerbing & Anderson, 1988; Green, Lissitz, & Mulaik, 1977; Hattie, 1985; Miller, 1995; Schmitt, 1996; Sijtsma, 2009). Takve kritike omogućile su prepoznavanje zloupotreba te procene i nastanak alternativnih procena pouzdanosti, zajedno sa procenama drugih važnih metrijskih karakteristika (Green & Yang, 2009; McDonald, 1985, 1999; Revelle, 1979; Sijtsma, 2009). Razvoj faktorsko-analitičkog modela merenja doveo je do toga da se pouzdanost definiše zajedno sa drugim karakteristikama kao što su homogenost i valjanost (Gerbing & Anderson, 1988; Green et al., 1977; McDonald, 1999; Miller, 1995). Pod faktorsko-analitičkim modelom pretpostavke klasičnog modela su eksplicitne i, što je najvažnije, empirijski proverljive, pre nego udaljeni matematičko-statistički ideali koji su, kao i svi drugi ideali, uglavnom nedostižni.

Stoga je uži cilj ovog teksta da se demonstriraju pojmovne i matematičko-statističke osnove za procene pouzdanosti, homogenosti i valjanosti nekog testa pod klasičnim i konfirmatornim jednofaktorskim modelom merenja. Prvo je definisana i opisana najviše korišćena i najviše zloupotrebljavana procena pouzdanosti pod klasičnim modelom merenja. Zatim je definisana alternativna procena pouzdanosti koja je istovremeno i procena homogenosti i valjanosti pod konfirmatornim jednofaktorskim modelom. Prilikom definisanja α izneti su argumenti zbog kojih se povoljna procena ovog koeficijenta smatra tek jednim od važnih, ali ne i nužnih razloga za zaključivanje o pouzdanosti, homogenosti i valjanosti nekog testa. Alternativni koeficijent ω je eksplicitno definisan u okviru faktorsko-analitičkog modela merenja i predstavlja bolju procenu pouzdanosti, homogenosti i valjanosti testa. Demonstracija definicija i procena ovih koeficijenata izvedena je elementarnom algebrom pomoću matematičko-statističkog modela običnih najmanjih kvadrata. Pri tom su korišćeni podaci koji su nastali primenom jednodimenzionalnog i zadovoljavajuće pouzdanog i valjanog instrumenta samoprocene za merenje opšteg zadovoljstva životom. Reč je o skali SWLS (*Satisfaction With Life Scale*: Diener, Emmons, Larsen, & Griffin, 1985) koja je primenjena na jednom od prigodnih, ali po veličini sasvim prihvatljivih uzoraka iz domaće populacije (Vasić, Šarčević i Trogrlić, 2011).

Od početka

Niz procena pouzdanosti u okviru klasične teorije testa potekao je iz nastojanja da se reši problem korekcije za atenuaciju ili umanjeње procene povezanosti varijabli zbog greške merenja (Spearman, 1904). Razrađenu varijantu ovog pokušaja, nezavisno jedan od drugog, ponudili su Brown (1910) i Spearman (1910, 1927). Nakon toga nizali su se pokušaji procenjivanja pouzdanosti pod manje ili više sličnim matematičko-statističkim modelima i sa uglavnom istim rezultatima (Cronbach, 1947; Guttman, 1945; Hoyt, 1941; Kuder & Richardson, 1937; Rulon, 1939). Ovaj razvoj kulminirao je prekretnim doprinosom Cronbacha (Cronbach, 1951; Cronbach & Shavelson, 2004), koji je predstavljao objedinjavanje dostignuća preteča i uspostavljanje osnova za dalji razvoj modela merenja u psihologiji. Tada je definisan koeficijent koji je poznat kao Cronbachova α , mada istorijat razvoja procena pouzdanosti pod klasičnim modelom merenja čini više opravdanim naziv Guttman–Cronbachova ili, čak, Spearman–Brown–Kuder–Richardson–Rulon–Hoyt–Guttman–Cronbachova α (McDonald, 1999; Momirović i sar., 1999). U istraživanjima i praksi vremenom su predložena različita tumačenja koeficijenta α od kojih su neka manje, a neka više zasnovana, a neka su potpuno neutemeljena (Cortina, 1993; Green et al., 1977; Hattie, 1985; Miller, 1995; Pedhazur & Pedhazur Shmelkin, 1991; Schmitt, 1996; Sijtsma, 2009). Prema neutemeljenim tumačenjima, koja su važna za razumevanje rezultata na nekom testu, te otuda i od naročitoг značaja za valjanu naučnu i stručnu psihološku delatnost, koeficijent α je zadovoljavajuća procena:

- 1) pouzdanosti nekog testa na osnovu jedne njegove primene,
- 2) njegove interne konzistencije ili, alternativno, njegove homogenosti i
- 3) sa prethodno povezanim, procena zasićenosti merenja tog testa glavnim faktorom.

Koeficijent α se procenjuje na osnovu raspoloživih podataka o broju stavki, varijansama i kovarijansama tih stavki i varijansi ukupnog rezultata na testu kao:

$$\alpha = \frac{m}{m-1} \left(1 - \frac{\sum \sigma_{jj}}{\sigma_Y} \right) \quad (5)$$

gde je m broj stavki u testu, $\sum \sigma_{jj}$ je zbir dijagonalnih elemenata iz matrice varijansi i kovarijansi stavki, a σ_Y ukupna varijansa testa. Za razumevanje koeficijenta α važna je definicija varijanse ukupnog rezultata na testu:

$$\sigma_Y = \sum \sigma_{jj} + 2(\sum \sigma_{jk}) \quad (6)$$

dakle, kao zbir svih varijansi stavki i dvostrukog zbira kovarijansi stavki. Kada se uzme u obzir definicija varijanse ukupnog rezultata na testu, definicija α je:

$$\alpha = \frac{m}{m-1} \left(1 - \frac{\sum \sigma_{jj}}{\sum \sigma_{jj} + 2(\sum \sigma_{jk})} \right) \quad (5a)$$

Standardizovano rešenje podrazumeva izračunavanje α iz proseka vandijagonalnih koeficijenata matrice R :

$$\alpha_z = \frac{m(\bar{r})}{1 + (m-1)(\bar{r})} \quad (7)$$

Što je praktično isto kao i rešenje koje su nezavisno predložili Brown (1910) i Spearman (1910) pre više od jednog veka. Međutim, ovde je to rešenje prilagođeno slučaju sa više od dve stavke i podacima koji su nastali jednom primenom testa (Gerbing & Anderson, 1988).

Za izračunavanje α na osnovu varijansi i kovarijansi i na osnovu interkorelacija stavki poslužiće, kao što je rečeno, jedna od matrica koje su nastale u skorašnjem istraživanju zadovoljstva životom u Srbiji (Vasić i sar., 2011):

Tabela 3

Matrica varijansi, kovarijansi, korelacija i procene parametara centralne tendencije i disperzije rezultata na stavkama SWLS (N = 962)

Stavke	X_1	X_2	X_3	X_4	X_5
X_1 Život je blizak idealnom.	1.06	0.58	0.50	0.40	0.45
X_2 Uslovi života su odlični.	0.63	1.13	0.56	0.43	0.36
X_3 Zadovoljan je svojim životom.	0.50	0.58	0.94	0.54	0.46
X_4 Do sada je ostvario bitne stvari koje želi.	0.45	0.50	0.58	1.20	0.47
X_5 Kada bi se ponovo rodio, ne bi ništa menjao.	0.56	0.46	0.54	0.62	1.47
<i>AS</i>	2.99	3.30	3.65	3.43	2.98
<i>SD</i>	1.03	1.06	0.97	1.10	1.21

Napomena: Stavke su neznatno reformulisane u odnosu na primenjenu verziju u formi trećeg lica muškog roda.

Svi koeficijenti iznad i ispod dijagonale su pozitivni i relativno slični, a dijagonalni elementi se kreću približno oko jedinice. Ipak, izvesna odstupanja primetna su kod četvrte i pete stavke iz SWLS. Na osnovu elemenata matrice mogu se izračunati podaci koji su neophodni za procenu α . Suma varijansi je u ovom primeru $\Sigma\sigma_{jj} = 5.79$, dvostruka suma kovarijansi je $2(\Sigma\sigma_{jk}) = 10.82$, a ukupna varijansa testa iznosi $\sigma_y = 16.61$, tako da je $\alpha = .814$. Izračunavanje preko prosečne interkorelacije stavki, koja iznosi .474, daje procenu $\alpha = .818$, što je neznatno različito na trećoj decimali u odnosu na procenu α preko varijansi i kovarijansi. Drugo rešenje za procenu α je smisljeno samo ako je ukupan rezultat na testu izveden kao zbir standardizovanih rezultata na stavkama, a ne na uobičajeni način, kao prosti sumativni rezultat (Cortina, 1993; Gerbing & Anderson, 1988; McDonald, 1999).

Međutim, opšte je poznato da istraživači i praktičari zloupotrebljavaju često i daleko veću razliku između α i α_z nego što je u ovom primeru saopštavajući drugu procenu, iako preporučuju i koriste prosti sumativni rezultat na testu.

U oba slučaja, dobijene su procene zadovoljavajućih veličina zaključujući pre svega po tendenciji prema jedinici kao apsolutnoj gornjoj granici pouzdanosti. Postoje različite preporuke o prihvatljivim veličinama procene α , od kojih se najčešće navodi vrednost od najmanje .70 (Cortina, 1993; Fajgelj, 2009; Momirović, 1998b; Schmitt, 1996). Ali, sve i da je reč o prihvatljivim procenama po uobičajenim preporukama ostaje pitanje njihovog zasnovanog tumačenja. Cronbach (1951) je bio prilično neodređen ili, blaže rečeno, nedovoljno precizan u za tadašnje akademske običaje dugačkom analiziranju preteča i definisanju α . Za naučno-istraživačku i praktičnu delatnost najvažnija je pretpostavka klasične teorije testa o τ -ekvivalentnim indikatorima (Brown, 2006; Lord & Novick, 1968; Novick & Lewis, 1968). Ova pretpostavka, koja se neposredno odnosi na tumačenje α kao prihvatljive procene pouzdanosti, podrazumeva da su stavke istih metrijskih karakteristika i da se pravi rezultat na njima može menjati za neku konstantu i nikako drugačije, što je u psihološkim merenjima gotovo po pravilu izuzetak (Green & Yang, 2009; McDonald, 1999; Miller, 1995; Raykov, 1997; Sijtsma, 2009). Primena koeficijenta α je neopravdana jer, kada nije zadovoljena pretpostavka τ -ekvivalentnosti, njegove procene negativno zastranjuju u odnosu na pravu pouzdanost.

U slučajevima kada procena α dovodi u pitanje opravdanost primene nekog testa neretko se pribegava gomilanju stavki. Da je reč o još jednoj zloupotrebi α , može se jasno videti iz izraza (5) i (7). Jöreskogovo (Jöreskog, 1971) rešenje, koje podrazumeva da se u računanju α koriste procene faktorskih opterećenja stavki pod modelom kongeneričkih indikatora, nije neko naročito poboljšanje (Gerbing & Anderson, 1988). To znači da α nije uvek prikladna i prihvatljiva procena pouzdanosti, bez obzira na dobijenu vrednost, posebno kada je reč o testu sa znatnim brojem stavki, za koje se ne zna da li je zadovoljena pretpostavka τ -ekvivalentnosti (Raykov, 2001). To važi pogotovo kada nisu priložene i alternativne procene pouzdanosti koje su povezane sa proverom faktorske valjanosti i drugih mogućih provera konstruktne valjanosti (Anderson & Gerbing, 1988; Cortina, 1993; Gerbing & Anderson, 1988; Green et al., 1977; Momirović, 1998a).

Zajedno sa prethodnim, često idu i zloupotrebe koje se sastoje u tome da se na osnovu povoljne procene α za neki test, izvode i zaključci da je on interno konzistentan ili, čak, da je homogen i da je ukupan rezultat na testu znatno zasićen glavnim faktorom. Izgleda da je pojam interne konzistencije jedan od osnovnih razloga za ovakvu zloupotrebu jer niko do sada nije ponudio neko njegovo razumno objašnjenje (McDonald, 1999). Prikladno, ali verovatno ne i dovoljno objašnjenje jeste to da ovaj pojam podrazumeva generalizaciju osnovnog shvatanja pouzdanosti kao saglasnosti rezultata merenja iz dva ili više nezavisnih pokušaja. Pod uslovom, naravno, da se u slučaju α kao procene pouzdanosti na osnovu jednog merenja, ponavljanje i saglasnost između tih ponavljanja mogu izvesti na osnovu zajedništva stavki. Ovakvo pojašnjenje interne konzistencije proizilazi iz srodno-

sti opšteg shvatanja merenja i klasične teorije testa i neposredno je povezano sa pretpostavkom τ -ekvivalenosti. Ali, pojam interne konzistencije u slučaju α znači samo zajedništvo stavki u smislu njihove međusobne povezanosti, a ne i zajedništvo stavki koje proističe iz jednog zajedničkog faktora (Green et al., 1977; Hattie, 1985; McDonald, 1981, 1985; Miller, 1995; Schmitt, 1996). U klasičnom modelu merenja homogenost je jedan od elemenata u izrazu za definisanje α (Kaiser, 1968), odnosno prosečna korelacija stavki iz testa koja je u lokalnim psihometrijskim krugovima poznata i kao prvi u nizu koeficijenata h (Momirović i sar., 1999). Ali, da bi se takva procena unutrašnje konzistentnosti smatrala pokazateljem homogenosti, opet zavisi od toga da li je i u kojoj meri zadovoljena pretpostavka τ -ekvivalentnosti. Naime, tek kada su interkorelacije stavki jednake njihovoj prosečnoj interkorelaciji, kada taj skup stavki ima samo jednu glavnu komponentu, onda je α_z neposredno povezana sa karakterističnom vrednošću te komponente (Kaiser, 1968).

Postoje i druge demonstracije na simuliranim podacima koje govore da je homogenost testa nezavisna od α kao procene pouzdanosti. U klasičnoj Monte Carlo simulaciji, u kojoj su generisani podaci za fiktivni 10-ajtemski test sa petofaktorskom strukturom, pokazano je da je α u funkciji broja stavki i da nije dovoljno osetljiva na dimenzionalnost kao osnovni uslov za zaključivanje o homogenosti (Green et al., 1977). Hattie (1985) je opisao i uporedio nekoliko desetina mogućih postupaka za procenu homogenosti i , između ostalog, konstatovao da zaključivanje o homogenosti na osnovu α nije opravdano. Test koji je homogen, u smislu jednodimenzionalnosti, nije nužno i pouzdan ili interno konzistentan. Takođe, α može biti visoka čak i ako test ne meri ili, bar, ne meri na dovoljno zadovoljavajući način jedan faktor jer a) zavisi od broja stavki i njihovog ponavljanja, b) raste u funkciji rasta broja faktora stavki i c) umereno se smanjuje sa porastom komunaliteta.

Gerbing i Anderson (Gerbing & Anderson, 1988) takođe definišu homogenost kao jednodimenzionalnost u smislu jedne od osnovnih pretpostavki opšte teorije merenja. Ali, kao i drugi autori, Gerbing i Anderson definišu ovu metrijsku karakteristiku u kontekstu opšteg modelskog pristupa i povezuju je sa pojmom faktorske valjanosti. Na osnovu simulirane matrice interkorelacija za dvodimenzionalni model merenja testa, koji sadrži 10 stavki, konstatovano je da je procena α podjednako zadovoljavajuća i pre i nakon razjašnjenja modela putem konfirmatorne faktorske analize. Gerbing i Anderson su konstatovali da je α značajna za procenu pouzdanosti, ali ne i da procenjuje jednodimenzionalnost. Jednodimenzionalnost je nužan, ali ne i dovoljan uslov za zaključivanje o konstruktnoj valjanosti testova. Cilj većine istraživanja trebalo bi da bude ne samo izrada jednodimenzionalnih i pouzdanih testova, nego i provera i gradnja teorija, odnosno konstruktne valjanosti i empirijski utemeljene nomološke mreže.

Cortina (1993) je pokazao da su u slučaju jednodimenzionalnog modela merenja za testove sa različitim brojem stavki i umerenom prosečnom interkorelacijom (.50), koja se može smatrati relativno visokom za većinu psiholoških testova, procene α prihvatljive nezavisno od broja stavki. Ako se test sastoji od 14 i više

stavki, procena α je prihvatljiva čak i ako je prostor merenja definisan sa dve dimenzije a prosečna interkorelacija stavki relativno niska, ali uobičajena za većinu psiholoških testova. Rezultati za trodimenzionalni slučaj samo jasnije ističu značaj rezultata dobijenih za dvodimenzionalni slučaj. Cortina je zaključio da α a) raste u funkciji rasta prosečne interkorelacije stavki, b) neznatno opada sa odstupanjem od homogenosti definisane kao jednodimenzionalnost i, što je možda najvažnije, c) može biti prihvatljiva i relativno visoka uprkos niskoj prosečnoj interkorelaciji stavki i višedimenzionalnosti. Povoljan ishod takve procene, posebno u slučaju višedimenzionalnosti, ne znači i to da ukupan rezultat na takvom testu može da se protumači na smislen i dosledan način. Ukratko, za neki test se može konstatovati da konzistentno meri nešto, ali da to nešto može ostati nepoznato ukoliko nije primenjen i neki postupak provere konstruktne valjanosti.

Schmitt (1996) je takođe pružio par demonstracija da α nije pokazatelj homogenosti ili jednodimenzionalnosti i da takvo zaključivanje, zajedno sa zloupotrebom dužine testa u prisilnom uvećavanju α , može voditi pogrešnim tumačenjima rezultata na testu. Jedan od osnovnih razloga za takve pogreške u zaključivanju je konfuzija u upotrebi pojmova interne konzistencije i homogenosti. Schmitt zauzima umereni stav jer ističe da je u slučaju stvarnih matrica retko kada tako jasna situacija kao u slučaju simuliranih matrica. Jednodimenzionalnost nije jednoznačni ishod provere, u smislu da se na osnovu jednog pokazatelja može izvesti zaključak da li je neki test jednodimenzionalan ili ne, što je zaključak sa kojim bi se saglasili i drugi (Gerbing & Anderson, 1988; Green et al., 1977; McDonald, 1981, 1985, 1999). Međutim, uprkos ovakvim umerenim stavovima koji ističu ograničenu i u primerenim uslovima relevantnu informativnost, u poslednje vreme se ozbiljno dovodi u pitanje čak i svaka primena koeficijenta α (Green & Yang, 2009; Sijtsma, 2009).

Do kraja

U poslednjih nekoliko decenija predložene su alternativne procene pouzdanosti na osnovu jedne primene testa. Neke od tih procena, kao što je pomenuta Jöreskogova (Jöreskog, 1971) varijanta α , zasnovane su na manje ili više uspešnom povezivanju klasičnog i faktorsko-analitičkog modela merenja. Druge procene, poput Reveleovog koeficijenta β , počivaju na hijerarhijskoj taksonomskoj analizi (Revelle, 1979; Revelle & Zinbarg, 2009; Zinbarg, Revelle, Yovel, & Li, 2005). Reveleov koeficijent β nije procena pouzdanosti prve Hotellingove (Hotelling, 1933) glavne komponente koja se isto označava (Momirović i sar., 1999). Postoje i alternativni predlozi koji su i dalje u okvirima klasičnog modela merenja, neposredno se nadovezujući na Gutmanove (Guttman, 1945) definicije najvećih donjih granica pouzdanosti (Ten Berge & Sočan, 2004; Sijtsma, 2009).

Međutim, jedini alternativni predlog za procenu pouzdanosti, koji je dosledno i neposredno izveden iz faktorsko-analitičkog modela, jeste koeficijent ω (McDonald, 1970, 1981, 1985, 1999). Koeficijent ω je i pokazatelj homogenosti,

definisane kao jednodimenzionalnost testa u smislu da je ukupni rezultat na testu u zadovoljavajućoj meri zasićen glavnim faktorom. S obzirom na to da je definisan u Spearmanovoj konfirmatornoj jednofaktorskoj varijanti ovog modela merenja, koeficijent ω predstavlja i jedan od pokazatelja valjanosti tog modela. Takva varijanta faktorsko-analičkog modela merenja važi kada je rezultat u svakoj pojedinoj stavci definisan kao u izrazu (4) i može se smatrati opštim slučajem za druge varijante ovog modela. Radi procene ω potrebno je redefinisati procene iz polazne matrice kovarijabiliteta tako da je kovarijansa stavki definisana kao:

$$\sigma_{jk} = \lambda_j \lambda_k \quad (8)$$

gdje su λ_j i λ_k opterećenja stavki zajedničkim faktorom. Varijansa svake stavke je:

$$\sigma_{jj} = \lambda_j + \psi_j \quad (9)$$

Zasićenja ili opterećenja stavki faktorom, odnosno procene u kojoj se meri rezultat na svakoj stavki može objasniti faktorom, definisana su kao:

$$\lambda_j = \sqrt{\frac{\sigma_{jk} \sigma_{jl}}{\sigma_{kl}}} \quad (10)$$

McDonald (1999) je pokazao da se definicija ω , kao optimalne procene pouzdanosti i homogenosti, može izvesti na osnovu faktorsko-analičkog modela kao:

$$\omega = \frac{(\sum \lambda_j)^2}{\sigma_Y^2} \quad (11)$$

ili, alternativno, uz striktno izvođenje homogenosti u smislu jednodimenzionalnosti na osnovu konfirmatornog jednofaktorskog modela, kao:

$$\omega = 1 - \frac{(\sum \psi_j^2)}{\sigma_Y^2} \quad (12)$$

Da bi se ω mogla izračunati za konkretni primer, potrebno je izračunati opterećenja i unikvite za svaku stavku pomoću jednačina (9) i (10). U Tabeli 4 dat je primer izračunavanja opterećenja prve stavke iz SWLS na osnovu njenih kovarijansi sa svim ostalim stavkama.

Tabela 4

Izračunavanje opterećenja (λ_1) prve stavke SWLS na osnovu kovarijansi te stavke sa ostalim stavkama skale pod jednofaktorskim modelom

Koraci	λ_1
1.	$\sqrt{\frac{\sigma_{12}\sigma_{13}}{\sigma_{23}}} = \sqrt{\frac{.630 \times .497}{.576}} = \sqrt{\frac{.313}{.576}} = \sqrt{.544} = .738$
2.	$\sqrt{\frac{\sigma_{12}\sigma_{14}}{\sigma_{24}}} = \sqrt{\frac{.630 \times .454}{.498}} = \sqrt{\frac{.268}{.498}} = \sqrt{.574} = .758$
3.	$\sqrt{\frac{\sigma_{12}\sigma_{15}}{\sigma_{25}}} = \sqrt{\frac{.630 \times .561}{.460}} = \sqrt{\frac{.353}{.460}} = \sqrt{.768} = .877$
4.	$\sqrt{\frac{\sigma_{13}\sigma_{14}}{\sigma_{34}}} = \sqrt{\frac{.497 \times .454}{.578}} = \sqrt{\frac{.226}{.578}} = \sqrt{.390} = .625$
5.	$\sqrt{\frac{\sigma_{13}\sigma_{15}}{\sigma_{35}}} = \sqrt{\frac{.497 \times .561}{.536}} = \sqrt{\frac{.279}{.536}} = \sqrt{.520} = .721$
6.	$\sqrt{\frac{\sigma_{14}\sigma_{15}}{\sigma_{45}}} = \sqrt{\frac{.454 \times .561}{.620}} = \sqrt{\frac{.255}{.630}} = \sqrt{.411} = .641$

U jednačinama i u izračunavanju opterećenja za prvu stavku iz SWLS može se prepoznati postupak trijada po Spearmanu (Spearman, 1927). To je, zapravo, analiza putanja u modeliranju strukturalnim jednačinama primenom pravila praćenja putanja u izračunavanju iznosa u kojem se rezultati na stavci mogu objasniti jednom latentnom varijablom (Loehlin, 2004). Različite kombinacije kovarijansi prve sa svim ostalim stavkama daju različite procene opterećenja te stavke faktorom. Najbolje rešenje, pod matematičko-statističkim modelom sa jednom varijablom, koja je ovde niz različitih procena opterećenja, jeste da se iz tih procena izvede prosečna vrednost. Isti postupak je primenjen i za ostale stavke i njihove kovarijanse sa svim preostalim stavkama (Tabela 5). Zainteresovani/a se upućuje da izračuna prosečna opterećenja za preostale četiri stavke iz SWLS, sa molbom da toleriše razlike između dobijenih i ovde izloženih rezultata, zbog zaokruživanja vrednosti tokom izračunavanja.

Tabela 5

Opterećenja ($\bar{\lambda}$), komunaliteti ($\bar{\lambda}^2$) i unikviteti (ψ^2) stavki iz SWLS pod jednofaktorskim modelom

Stavke	$\bar{\lambda}$	$\bar{\lambda}^2$	ψ^2
1. Život je blizak idealnom.	.73	.52	.54
2. Uslovi života su odlični.	.84	.71	.41
3. Zadovoljan je svojim životom.	.75	.56	.38
4. Do sada je ostvario bitne stvari koje želi.	.73	.53	.67
5. Kada bi se ponovo rodio ne bi ništa menjao.	.74	.55	.91

Procene opterećenja stavki faktorom deluju sasvim prihvatljivo, posebno u svetlu ranije konstatovanih razlika u kovarijabilitetu nekih stavki. Ako se izuzme druga stavka SWLS, koja se po visini opterećenja jasno izdvaja, opterećenja stavki su relativno ujednačena, pozitivna i zadovoljavajućih veličina. Dakle, na osnovu ovih procena može se zaključiti da je najveći deo rezultata na stavkama iz SWLS moguće objasniti jednom latentnom varijablom. Za izračunavanje ω , na osnovu izraza (11), potrebno je još da se izračuna kvadrirani zbir opterećenja stavki generalnim faktorom koji iznosi $(\Sigma\lambda_j)^2 = 14.39$. Za izračunavanje ω , prema izrazu (12), potrebno je izračunati i zbir unikviteta koji u ovom primeru iznosi $(\Sigma\psi_j^2) = 2.91$. Za oba izračunavanja je potrebna i procena varijanse ukupnog rezultata na testu koja iznosi $\sigma_Y^2 = 16.61$ (Prilog A). U prvom slučaju koeficijent ω je .87, a u drugom .83, što je u oba slučaja više od procena pouzdanosti preko koeficijenta α . Ovo je još jedna potvrda da je α nezadovoljavajuća procena pouzdanosti i da nije indikativna za homogenost kao jednodimenzionalnost. McDonald (1999) tvrdi da je ω isto što i:

- 1) kvadrirana korelacija između ukupnog rezultata na testu i zajedničkog faktora koji aproksimira hipotetički konstrukt koji test treba da meri;
- 2) korelacija između alternativnih formi nekog testa sa istim predmetom merenja koje imaju isti zbir ili prosek faktorskih zasićenja i njihovih unikviteta i za koje važi jedan jednofaktorski model merenja – odnosno, da su obe forme homogene i,
- 3) kvadrirana korelacija između ukupnog rezultata na testu sa m -tim uzorkom stavki i ukupnog rezultata koji bi se mogao dobiti na osnovu nekog beskonačnog skupa stavki iz nekog homogenog univerzuma M , iz kojeg je izvučen taj uzorak stavki.

Pod konfirmatornim jednofaktorskim modelom mogu se izvesti još neke procene koje zajedno sa koeficijentom ω dodatno potkrepljuju zaključak o njegovoj prihvatljivosti. Pomoću izraza (8) izračunate su dve matrice koje su neophodne za dalja sagledavanja empirijske utemeljenosti modela u smislu njegove saglasnosti sa raspoloživim podacima. U Tabeli 6 su zajedno izložene procene reprodukovane

nih kovarijansi i reziduala koji predstavljaju razlike između empirijskih i reprodukovanih kovarijansi.

Tabela 6

Matrica reprodukovanih kovarijansi (σ_{jk}' - ispod dijagonale) i reziduala (iznad dijagonale) na stavkama SWLS za jednofaktorski model merenja

Stavke	1.	2.	3.	4.	5.
1. Život je blizak idealnom.		.02	-.05	-.08	.02
2. Uslovi života su odlični.	.61		-.06	-.12	-.17
3. Zadovoljan je svojim životom.	.55	.63		.03	-.02
4. Do sada je ostvario bitne stvari koje želi.	.53	.62	.55		.08
5. Kada bi se ponovo rodio, ne bi ništa menjao.	.64	.63	.55	.54	

Pomalo zaboravljeni način razmatranja saglasnosti modela sa podacima, baš kao i pregled matrice varijansi i kovarijansi i korelacija stavki, jeste upravo pregled reziduala (Prilog B). Većina reziduala u primeru je oko nule, s tim da su jedni pozitivnog, a drugi negativnog predznaka, što znači da neke procene precenjuju, a druge potcenjuju – naučni model nikada nije jednak stvarnosti. Bolja apsolutna procena, koja je prikladna za veće matrice ali koja ipak ne može zameniti pregled reziduala, jeste kvadratni koren prosečnih kvadriranih reziduala ili RMSR (*Root Mean Square of Residuals*: Loehlin, 2004) koji ovde iznosi .08. Ova vrednost se može smatrati zadovoljavajućom i može se pretpostaviti da bi pod složenijim matematičko-statističkim modelima bila bliža prihvatljivoj vrednosti do .05 (Loehlin, 2004).

Pomoću običnih najmanjih kvadrata mogu se izvesti još neke procene saglasnosti modela sa podacima. Jedan od tih apsolutnih pokazatelja jeste indeks saglasnosti ili GFI (*Goodness of Fit Index*: Jöreskog & Sörbom, 1996). Za neke druge, po svojstvima povoljnije relativne procene saglasnosti modela, kao što je kvadratni koren prosečne greške u aproksimaciji ili RMSEA (*Root Mean Square Error of Aproximation*: Loehlin, 2004; Kaplan, 2000; Steiger & Lind, 1980), potrebni su matematičko-statistički modeli koji zahtevaju iterativno mašinsko izračunavanje. Ovde će se koristiti samo procena GFI koja se može izvesti i pomoću modela običnih najmanjih kvadrata (McDonald, 1999). Za izračunavanje GFI potrebna je funkcija diskrepancije:

$$q_u = (1/m^2) \sum_j \sum_k (\sigma_{jk} - \sigma'_{jk})^2 \quad (13)$$

što je prosek kvadriranih razlika između empirijskih i reprodukovanih kovarijansi. Takođe, potrebna je i procena saglasnosti modela koja uvažava raspon empirijskih kovarijansi (McDonald, 1999):

$$c = (1/m^2) \sum_j \sum_k (\sigma_{jk}) \quad (14)$$

da bi se zatim mogao izračunati i apsolutni pokazatelj saglasnosti:

$$\text{GFI} = 1 - q_u / c \quad (15)$$

Za odabrani empirijski primer izračunate su vrednosti za $c = .12$ i $q_u = .002$, tako da procena saglasnosti GFI iznosi .98. To je vrednost koja je veoma bliska potpunoj jediničnoj saglasnosti i deluje dosta usaglašeno sa procenama koje su dobijene pomoću složenijih matematičko-statističkih procena pokazatelja saglasnosti modela sa podacima iz izvornog istraživanja (Vasić i sar., 2011). Takođe, to je još jedna u nizu međusobno usaglašanih procena koja gotovo van svake sumnje upućuje na zaključak da je SWLS pouzdana, homogena i valjana mera u strukturalnom smislu jednostavnog konstrukta kojim se opisuju i objašnjavaju individualne razlike u zadovoljstvu životom.

Natrag ka naravoučeniju

Postoji više alternativnih tumačenja koeficijenta α , od kojih su neka manje a neka više utemeljena, dok su neka tumačenja potpuno neosnovana. Ovde je izdvojeno nekoliko međusobno povezanih neutemeljenih tumačenja α , po kojima je to zadovoljavajuća procena pouzdanosti i dovoljan razlog za zaključivanje da je neki test homogen u nekom smislu interne konzistencije, odnosno da je ukupan rezultat u znatnoj meri zasićen glavnim faktorom. U većem broju simulacija demonstrirana je neutemeljenost tumačenja koeficijenta α kao prihvatljive procene pouzdanosti, posebno u onim slučajevima kada ne važi pretpostavka o suštinskoj τ -ekvivalentnosti (Kaiser, 1968; McDonald, 1999; Novick & Lewis, 1968; Raykov, 1997, 2001). Takođe, pokazano je i da je taj koeficijent neadekvatna osnova za zaključivanje o homogenosti, posebno kada se za neki test ne može tvrditi da je proste, jednofaktorske strukture (Cortina, 1993; Green et al., 1977; Hattie, 1985; Miller, 1995; Schmitt, 1996; Sijtsma, 2009). U većini zaključaka zagovara se modelski pristup koji podrazumeva lančanje analitičkih postupaka i konsultovanje više procena prilikom provere pouzdanosti, homogenosti i valjanosti nekog testa (Anderson & Gerbing, 1988; Cortina, 1993; Gerbing & Anderson, 1988; Hattie, 1985; Miller, 1995). Iako postoje stanovišta koja u potpunosti dovode u pitanje korisnost α (Green & Yang, 2009; Revelle & Zinbarg, 2009; Sijtsma, 2009), po svedecu sudeći za sada je dovoljno da se zauzme umeren stav da ovaj koeficijent zajedno sa nekim drugim procenama može biti makar i delimično informativan u okvirima modelskog pristupa (Gerbing & Anderson, 1988; Schmitt, 1996).

Ovde je, saglasno sa rezultatima ranijih demonstracija (Green & Yang, 2009; McDonald, 1981, 1985, 1999; Sijtsma, 2009), ali sada na realnoj matrici, ponovo pokazano da je koeficijent α nužan, ali ne i dovoljan razlog da se za neki test može zaključiti kako je pouzdan i homogen. Takođe, prikazano je da α ima smisla tek u

kontekstu šireg modelskog pristupa, koji podrazumeva lančanje postupaka i procena u traganju za dovoljnom količinom nalaza potkrepljujućih za takav zaključak (Anderson & Gerbing, 1988; Gerbing & Anderson, 1988; McDonald, 1985, 1999; Miller, 1995). U skladu sa drugim stanovištima (McDonald, 1985, 1999; Zinbarg et al., 2005), pored α potrebno je izračunati i procenu pouzdanosti koja je neposredno izvedena iz Spearmanovog konfirmatornog jednofaktorskog modela. Ova procena, poznata i kao koeficijent ω , istovremeno je procena pouzdanosti, homogenosti i, zajedno sa drugim procenama saglasnosti mernog modela sa podacima, govori o faktorskoj valjanosti ili, čak, generalizabilnosti.

Dakle, α je u odnosu na ω manje povoljna procena pouzdanosti, naročito kada pretpostavka o suštinskoj τ -ekvivalentnosti nije zadovoljena. Ipak, ovo je jedna od najčešće korišćenih procena pouzdanosti, možda i zato što je povezana sa nedovoljno jasnim pojmom interne konzistencije. Takođe, koeficijent α , čak i u slučajevima povoljnih procena, ne govori o homogenosti kao jednodimenzionalnosti zato što je definisan u okviru klasičnog modela merenja na osnovu međusobne povezanosti stavki u nekom testu. Pojam homogenosti je dosledno definisan kao jednodimenzionalnost, ne u klasičnom nego u faktorsko-analitičkom modelu merenja. Iz svega toga može se zaključiti da je α bez dodatnih procena nedovoljna i najčešće neadekvatna procena pouzdanosti i sasvim sigurno tek nužan, ali ne i dovoljan, osnov za zaključivanje o homogenosti nekog psihološkog testa.

Za razliku od koeficijenta α , koeficijent ω je neposredno definisan u okviru faktorsko-analitičkog modela merenja. Tačnije, ovaj koeficijent je definisan pod takozvanim Spearmanovim konfirmatornim jednofaktorskim modelom. Iako postoje i druge procene pouzdanosti, te procene nisu kao koeficijent ω eksplicitno definisane u okviru faktorsko-analitičkog modela nego u okviru klasičnog modela merenja i hijerarhijske taksonomske analize. Po svemu sudeći, ω je bolja procena pouzdanosti u smislu najveće donje granice pouzdanosti, a u slučajevima pretpostavljene jednodimenzionalnosti, to je i veoma dobra procena homogenosti i valjanosti. U nizu procena saglasnosti konfirmatornog jednofaktorskog modela sa podacima generisanim nekim psihološkim testom, koeficijent ω je jedna od procena koje govore o empirijskoj zasnovanosti tog modela. Ukratko, koeficijent ω predstavlja u okviru faktorsko-analitičkog modela merenja dobru procenu najveće donje granice pouzdanosti, homogenosti i valjanosti (McDonald, 1985, 1999).

Naravno, činjenica da je koeficijent ω definisan pod konfirmatornim jednofaktorskim modelom, može navesti na sumnju u to da je takva procena primarena kada je reč o psihološkim testovima kojima se mere složene biopsihosocijalne karakteristike (Lucke, 2005; Zinbarg et al., 2005). Cattell (1973) je bio jedan od retkih doslednih zagovornika konstruisanja i primene tzv. faktorski složenih testova. Ali, njegove egzotične ekskurzije u psihometrijske vode dovele su do toga da je, između ostalog, zaboravljeno i njegovo upozorenje da se homogenost lako može zameniti za ono što je on nazvao *naduvanim specifičnim faktorima*. Međutim, jednako je lako da se za neki pouzdani, homogeni i valjani test, u nedostatku nalaza nastalih primenom više različitih analitičkih postupaka, pogrešno zaključi da je mera naduvanog specifika, a ne neke relevantne biopsihosocijalne karakte-

ristike. Jedina relativno sigurna brana od takvih pogrešaka jeste stručna procena sadržinske valjanosti i istrajan istraživački rad na pletenju nomološke mreže koji, između ostalog, podrazumeva lančanje postupaka primereno složenosti pojava. U psihološkim merenjima česta pretpostavka o složenoj faktorskoj strukturi može izgledati kontradiktorno u odnosu na pretpostavku o homogenosti kao jednodimenzionalnosti. Ali, treba imati na umu da se sa pretpostavkom o složenoj faktorskoj strukturi pitanje homogenosti ili jednodimenzionalnosti samo pomera negde niže u faktorskoj hijerarhiji. Takođe, nije suvišno istaći da postoje i predlozi za procenu koeficijenta ω , čak i za faktorski složene hijerarhijske modele merenja. Ali, to je već druga priča o Alisi i njenim doživljajima u svetu čuda.

Reference

- Anderson, J. C., & Gerbing, D. W. (1988). Structural equation modeling in practice: A review and a recommended two-step approach. *Psychological Bulletin*, *103*, 411–423.
- Anderson, N. H. (2003). *Empirical direction in design and analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Bollen, K. A. (1989). *Structural equations with latent variables*. New York: Wiley.
- Bollen, K. A. (2002). Latent variables in psychology and social sciences. *Annual Review of Psychology*, *53*, 605–634.
- Brown, T. A. (2006). *Confirmatory factor analysis for applied research*. New York/London: The Guilford Press.
- Brown, W. (1910). Some experimental results in the correlation of mental abilities. *British Journal of Psychology*, *2*, 296–322.
- Bukvić, A. (1996). *Načela izrade psiholoških testova*. Beograd: Zavod za udžbenike i nastavna sredstva.
- Carroll, L. (1865). *Alice in Wonderland*. New York: Harper Collins.
- Cattell, R. B. (1973). *Personality and mood by questionnaire*. San Francisco: Jossey-Bass.
- Cortina, J. M. (1993). What is coefficient alpha? An examination of theory and applications. *Journal of Applied Psychology*, *78*, 98–104.
- Crocker, L., & Algina, J. (1986). *Introduction to classical and modern test theory*. New York: Holt, Rinehart, and Winston.
- Cronbach, L. J., & Meehl, P. E. (1955). Construct validity in psychological tests. *Psychological Bulletin*, *52*, 281–302.
- Cronbach, L. J. (1947). Test „reliability”: Its meaning and determination. *Psychometrika*, *12*, 1–16.
- Cronbach, L. J. (1951). Coefficient alpha and the internal structure of tests. *Psychometrika*, *16*, 297–334.

- Cronbach, L. J. (with editorial assistance by R. J. Shavelson) (2004). My current thought of coefficient alpha and successor procedures. *Educational and Psychological Measurement*, 64, 391–418.
- Diener, E., Emmons, R. A., Larsen, R. J., & Griffin, S. (1985). The Satisfaction With Life Scale. *Journal of Personality Assessment*, 49, 71–75.
- Fajgelj, S. (2009). *Psihometrija: metod i teorija psihološkog merenja*. Beograd: Centar za primenjenu psihologiju.
- Fiske, D. W. (1971). *Measuring the concepts of personality*. Chicago: Aldine.
- Gerbing, D. W., & Anderson, J. C. (1988). An updated paradigm for scale development incorporating unidimensionality and its assessment. *Journal of Marketing Research*, 25, 186–192.
- Green, S. B., & Yang, Y. (2009). Commentary on coefficient alpha: A cautionary tale. *Psychometrika*, 74, 121–135.
- Green, S. B., Lissitz, R. W., & Mulaik, S. A. (1977). Limitations of coefficient alpha as an index of test unidimensionality. *Educational and Psychological Measurement*, 37, 827–838.
- Gulliksen, H. (1950). *Theory of mental tests*. New York: John Wiley & Sons.
- Guttman, L. A. (1945). A basis for analysing test-retest reliability. *Psychometrika*, 10, 255–282.
- Hattie, J. (1985). Methodology review: Assessing unidimensionality of tests and items. *Applied Psychological Measurement*, 9, 139–164.
- Hotelling, H. (1933). Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology*, 24, 417–441, 495–520.
- Hoyt, C. (1941). Test reliability estimated by analysis of variance. *Psychometrika*, 6, 158–160.
- Jöreskog, K. G. (1971). Statistical analysis of sets of congeneric test. *Psychometrika*, 36, 109–133.
- Jöreskog, K., & Sörbom, D. (1996). *LISREL 8: A user's reference guide*. Lincolnwood, IL: Scientific Software International.
- Kaiser, H. F. (1968). A measure of the average intercorrelation. *Educational and Psychological Measurement*, 28, 245–247.
- Kaplan, D. (2000). *Structural equation modeling: Foundations and extensions*. Thousand Oaks, CA: Sage.
- Kuder, G. F., & Richardson, M. W. (1937). The theory of the estimation of test reliability. *Psychometrika*, 2, 151–160.
- Loehlin, J. C. (2004). *Latent variable models: An introduction to factor, path, and structural equation analysis*. Mahwah, NJ: Lawrence Erlbaum.
- Lord, F. M., & Novick, M. R. (with contributions by A. Birnbaum) (1968). *Statistical theories of mental test scores*. Reading, MA: Addison-Wesley.

- Lucke, J. F. (2005). The α and ω of congeneric test theory: An extension of reliability and internal consistency to heterogeneous tests. *Applied Psychological Measurement, 29*, 65–81.
- Maxwell, S. E., & Delaney, H. D. (1990). *Designing experiments and analysing data: A model comparison perspective*. Mahwah, NJ: Lawrence Erlbaum.
- McDonald, R. P. (1970). The theoretical foundations of common factor analysis, principal factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology, 23*, 1–21.
- McDonald, R. P. (1981). The dimensionality of tests and items. *British Journal of Mathematical and Statistical Psychology, 34*, 100–117.
- McDonald, R. P. (1985). *Factor analysis and related methods*. Hillsdale, NJ: Lawrence Erlbaum.
- McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah, NJ: Lawrence Erlbaum.
- Miles, J., & Shevlin, M. (2001). *Applying regression and correlation: A guide for students and researchers*. London: Sage.
- Miller, M. B. (1995). Coefficient alpha: A basic introduction from the perspectives of classical test theory and structural equation modeling. *Structural Equation Modeling, 2*, 255–273.
- Momirović, K. (1998a). O realnoj egzistenciji psiholoških konstrukata. U K. Momirović (Ur.), *Realnost psiholoških konstrukata* (str. 1–8). Beograd: Institut za psihologiju i Institut za sociološka i kriminološka istraživanja.
- Momirović, K. (1998b). Još sedam malih tabela: poredbena analiza metrijskih karakteristika testova EPQ103 i PNEI dobijenih u dve sasvim različite eksperimentalne situacije. U K. Momirović (Ur.), *Realnost psiholoških konstrukata* (str. 242–253). Beograd: Institut za psihologiju i Institut za sociološka i kriminološka istraživanja.
- Momirović, K., Wolf, B. i Popović, D. A. (1999). *Uvod u teoriju merenja I: Interne metrijske karakteristike kompozitnih mernih instrumenata*. Priština: Fakultet za fizičku kulturu.
- Mulaik, S. A. (1972). *The foundations of factor analysis*. New York: McGraw-Hill.
- Myers, J. L., & Well, A. D. (2008). *Research design and statistical modeling*. Mahwah, NJ: Lawrence Erlbaum.
- Novick, M. R., & Lewis, C. L. (1968). Coefficient alpha and the reliability of composite measurements. *Psychometrika, 32*, 1–13.
- Pedhazur, E. J., & Pedhazur Shmelkin, L. (1991). *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum.
- Raykov, T. (1997). Cronbach's coefficient alpha, and violations of essential tau-equivalence with fixed congeneric components. *Multivariate Behavioral Research, 32*, 329–352.

- Raykov, T. (2001). Estimation of congeneric scale reliability using covariance structure analysis with nonlinear constraints. *British Journal of Mathematical and Statistical Psychology*, *54*, 315–328.
- Revelle, W. (1979). Hierarchical cluster analysis and the internal structure of tests. *Multivariate Behavioral Research*, *14*, 57–74.
- Revelle, W., & Zinbarg, R. E. (2009). Coefficients alpha, beta, omega, and the glb: Comments on Sijtsma. *Psychometrika*, *74*, 145–154.
- Royce, J. R. (1963). Factors as theoretical constructs. *American Psychologist*, *18*, 522–528.
- Rulon, P. J. (1939). A simplified procedure for determining the reliability of a test by split halves. *Harvard Educational Review*, *9*, 99–103.
- Schmitt, N. (1996). Uses and abuses of coefficient alpha. *Psychological Assessment*, *8*, 350–353.
- Sijtsma, K. (2009). On the use, the misuse, and the very limited usefulness of Cronbach's alpha. *Psychometrika*, *74*, 107–120.
- Spearman, C. (1910). Correlation calculated from faulty data. *British Journal of Psychology*, *2*, 271–295.
- Spearman, C. (1927). *The abilities of man: Their nature and measurement*. London: Macmillan.
- Spearman, C. E. (1904). The proof and measurement of association between two things. *American Journal of Psychology*, *15*, 72–101.
- Steiger, J. H., & Lind, J. C. (1980). *Statistically based tests for the number of common factors*. Paper presented at the annual meeting of the Psychometric Society, Iowa, IA.
- Ten Berge, J. M. F., & Sočan, G. (2004). The greatest lower bound to the reliability of a test and the hypothesis of unidimensionality. *Psychometrika*, *69*, 613–625.
- Thompson, B., & Wacha-Haase, T. (2000). Psychometrics is datametrics: The test is not reliable. *Educational and Psychological Measurement*, *60*, 174.
- Tinsley, H. E. A., & Brown, S. D. (2000). Multivariate statistics and mathematical modeling. In H. E. A. Tinsley & S. D. Brown (Eds.), *Handbook of applied multivariate statistics and mathematical modeling* (pp. 3–36). San Diego, CA: Academic Press.
- Vasić, A., Šarčević, D. i Trogrlić, A. (2011). Zadovoljstvo životom u Srbiji. *Primenjena psihologija*, *4*, 151–177.
- Zinbarg, R. E., Revelle, W., Yovel, I., & Li, W. (2005). Cronbach's α , Revelle's ω , and McDonald's ω_H : Their relations with each other and two alternative conceptualizations of reliability. *Psychometrika*, *70*, 123–133.

**Aleksandar Vasić &
Dušana Šarčević**

Department of
Business Psychology,
Faculty of Legal and
Business Studies
"Dr Lazar Vrkatić"
in Novi Sad,
University Union,
Belgrade

FROM ALPHA TO OMEGA AND BACK

Coefficient α , also known as the coefficient of internal consistency, is one of the most used and at the same time the most misused reliability assessment method of the composite psychological instruments or tests. For the interpretation of results generated through the application of psychological tests, the most important misuse of the coefficient α is the wrong conclusion about homogeneity or unidimensionality. One source of this kind of misuse has originated from the very definition of the coefficient α . The other source is insufficient differentiation of the definitions of reliability, homogeneity and validity between classical measurement model, also known as theory of the true and error score, and its developed version, the factor-analytic model of measurement. Because of that, there has been for some time an alternative index for the assessment of reliability, homogeneity and validity under the so-called spearmanian confirmatory one-factor model, which is known as coefficient ω . Using the data generated with the one scale resembling standard composite psychological test, the usefulness and limitations of the coefficients α and ω are described through the simple algebra. Based on these results and in conjunction with previous analyses, it is showed that the acceptable assessment of the coefficient α is necessary but not sufficient reason for the conclusion about the reliability and homogeneity of the test. For the reasonable and sound conclusion of this kind it is necessary to use other assessments such as coefficient ω , together with the indices of factor validity and goodness of fit under the hypothesis of spearmanian one-factor model.

Key words: psychological tests, reliability, homogeneity, validity, factor analysis

Prilog A

Rutine za izračunavanje ω u raspoloživim programima

Mašinsko izračunavanje koeficijenta ω može se obaviti pomoću bilo kog komercijalnog ili javnog programa za statističke analize koji sadrži i modul za konfirmatornu faktorsku analizu (CFA). Ako program poseduje i modul koji omogućava programiranje makroa u datom okruženju, takvo izračunavanje može se obaviti još lakše. Alternativno, prvo se može koristiti modul za CFA, a zatim se odgovarajući rezultati mogu iskoristiti u nekom drugom programu koji omogućava elementarne algebarske ili matrične operacije sa skalarima, vektorima i matricama. U tabeli ispod su prikazani koraci ili rutine za takva izračunavanja.

Rutine za izračunavanje ω

Koraci u izračunavanju na osnovu λ (izraz 11)	Koraci u izračunavanju na osnovu ψ (izraz 12)
1. Izračunati zbir opterećenja pod konfirmatornim jednofaktorskim modelom.	1. Izračunati unikvitete pod konfirmatornim jednofaktorskim modelom.
2. Izračunati kvadrat zbira opterećenja.	2. Izračunati zbir unikviteta.
3. Izračunati varijansu kompozita.	3. Izračunati varijansu kompozita.
4. Podeliti kvadrat zbira opterećenja sa varijansom kompozita.	4. Podeliti zbir unikviteta sa varijansom kompozita.
	5. Oduzeti rezultat deljenja iz prethodnog koraka od maksimalne, jedinične pouzdanosti.

Onima koji preferiraju korišćenje gotovih programskih rešenja može dobro poslužiti niz rutina u javnom programskom okruženju R koje je sačinio William Revelle. Korisni linkovi su: <http://www.inside-r.org/packages/cran/psych/docs/omega> i <http://personality-project.org/r/psych.manual.pdf>

Prilog B

Provera pretpostavke o τ -ekvivalentnosti iz klasičnog modela merenja

U proveri empirijske zasnovanosti jednodimenzionalnog konfirmatornog modela može se proveriti i pretpostavka o τ -ekvivalentnosti. Recimo, za sve stavke iz SWLS može se uzeti samo najveće opterećenje jedne od stavki ($\lambda = .84$) ili prosečno opterećenje na osnovu procena opterećenja iz Tabele 5 ($\lambda = .76$). Ista opterećenja za sve stavke su posledica pretpostavke o τ -ekvivalentnosti. Kakve su tada procene reziduala, RMSR i GFI? Šta se na osnovu takvih procena može zaključiti?

U tabeli ispod izložene su procene reziduala pod uslovom maksimalnog opterećenja (ispod glavne dijagonale) i prosečnog opterećenja (iznad glavne dijagonale). Pod uslovom maksimalnog opterećenja sve reprodukovane kovarijanse su iste ($\sigma'_{jk} = .71$, RMSR = .12 i GFI = .89. Pod uslovom prosečnog opterećenja reprodukovane kovarijanse su $\sigma'_{jk} = .71$, RMSR = .07 i GFI = .98. Dakle, u slučaju maksimalnog opterećenja prosti reziduali, i druge ovde izračunate procene saglasnosti, jasno ukazuju na nezasnovanost pretpostavke. U drugom slučaju, sa prosečnim opterećenjem, jedino prosti reziduali ukazuju na problematičnost pretpostavke, dok su procene RMSR i GFI praktično iste kao i u empirijskom primeru.

Matrica reprodukovanih kovarijansi pod uslovom maksimalnog (ispod dijagonale) i prosečnog opterećenja (iznad dijagonale) na stavkama SWLS za jednofaktorski model merenja

Stavke	1.	2.	3.	4.	5.
1. Život je blizak idealnom.		.05	-.08	-.12	-.02
2. Uslovi života su odlični.	-.01		-.00	-.08	-.12
3. Zadovoljan je svojim životom.	-.21	-.14		.00	-.04
4. Do sada je ostvario bitne stvari koje želi.	-.26	-.21	-.13		.04
5. Kada bi se ponovo rodio, ne bi ništa menjao.	-.15	-.25	-.18	-.09	

Na osnovu ovakvih procena, koje su dobijene pod dva uslova kao posledica pretpostavke o τ -ekvivalentnosti, može se zaključiti sledeće:

1. Saglasno ranijim stanovištima, takva pretpostavka ne važi kod većine psiholoških testova, a u retkim slučajevima teorijski i psihometrijski dobro zasnovanih testova može se govoriti tek o njenoj uslovnoj aproksimaciji.
2. Saglasno modelskom pristupu, u razmatranju empirijske zasnovanosti nekog mernog modela potrebno je koristiti prikladne matematičko-statističke modele i više različitih procena. Pri tome ne treba podleći pomodarstvu nego treba konsultovati i neke klasične pokazatelje, poput prostih ili standardizovanih reziduala.
3. Procene su činjenice. Saglasnost više takvih procena možda daje težinu jednom u odnosu na alternativne zaključke. Ali, čak i tada još uvek ima dovoljno prostora za dodatnu proveru (ne)saglasnih procena i empirijske zasnovanosti konstrukta i njegove nomološke mreže.