

MINIMÁLIS ELEMSZÁMU EKVIVALENS TESZTEK
KIVÁLASZTÁSA

Hunya Péterné



Pedagógiai, pszichológiai mérőeszközök készítésekor gyakran merül fel az a kérdés, hogy a kutatások során alkalmazott és eredményesnek bizonyult tesztek közül az általános használatbavétel előtt hogyan lehetne olyan optimális mérőeszközt létrehozni, amely jó közelítéssel azt méri amit az eredeti teszt, de lehetőség szerint a legkevesebb kérdéssel /elemmel/. Különösen élesen jelentkezik ez a probléma a nagyméretű, egy-egy területet /tananyag, művelési képesség, fogalomrendszer stb./ teljesen lefedő teszt esetén, amikor a mérés több /esetenként több tíz/ órát is igénybe vehet. Ezért az ilyen teszt használata mérőeszközként gyakorlatilag lehetetlen. A probléma megoldása nem csak elméleti, de gyakorlati szempontból is fontos. A nagyméretű tesztből kell megszerkesztenünk azt az optimális mérőeszközt, mely az eredetiből maximális információt tartalmaz, de a lehető legkevesebb számu elemmel. Megállapodás szerint a továbbiakban maximális információt nyujtónak tekintjük a rövidített tesztet akkor, ha korrelációja az eredeti tesztel megfelelően magas. A feladat megoldásánál abból indultunk ki, hogy a teszteknek vannak olyan elemcsoportjai, mely csoportok különböző, egymástól független dolgokat mérnek, de a csoportokon belül az elemek közel azonos dolog /fogalom, képesség, tudásszint stb./ mérését célozzák. Ezek a csoportok viszonylag jól elkülöníthetők egymástól empirikusan is a tesztelemzés során. Ezen információkat is felhasználjuk az optimalizálás során.

Ugyancsak a tesztek általános használatbavétele előtt - az optimalizálás mellett - felmerül egy másik lényeges probléma is, mégpedig az, hogy lehetséges-e több, egymással ekvivalens teszt megszerkesztése ugyanabból az alaptesztből kiindulva, és ha igen, milyen módszerrel. Az optimalizálttal ekvivalensnek tekintünk egy tesztet, ha az eredeti teszt főbb strukturáját meghatározó csoportok mind-

egyikéből tartalmaz elemet /kérdést, feladatot/ és /ezen teszt/ korrelációja az eredetivel ugyanolyan, mint az optimalizálté.

Az optimalizációs eljárást két művelleti képességet vizsgáló szubteszten mutatjuk be, nevezetesen: egy logika teszt tagadás résztesztjén és egy rendszerezési képességet minősítő teszt verbális rendszerezés szubtesztjén. Előljáróban röviden szólnunk erről a két tesztről. A JATE-n több éve folynak kutatások azzal a céllal, hogy feltárják a művelleti képességek strukturáját /Nagy, 1983 /. Mérőeszközök készültek, hogy segítségükkel mérhető legyen a művelleti képességek kialakulási folyamata az egymást követő korosztályokban. A kutatások nem egyes kiragadott művelleti képességek megismerésére irányulnak, hanem azok teljes rendszerére. Az eddigiek alapján úgy vélhető, hogy a művelleti képességeknek négy nagy relative önálló, de egymással sok szálon összefüggő rendszere van. Ezek a következők:

- logikai képességek,
- rendszerezési képességek,
- kombinatív képességek,
- bizonyítási képességek.

Az első háromhoz készített mérőeszközökkel /tesztekkel/ három korcsoportban folytak mérések. A legnagyobb létszámú a középső korcsoport volt /8. osztályosok/ és 626 tanulóra terjedt ki. A jelen cikkben bemutatott eljárások e korcsoport adataira támaszkodnak. Mivel célunk most nem az optimalizált /és velük ekvivalens/ tesztek létrehozása és ismertetése két résztesztet választottunk ki, hogy ezeken demonstráljuk a módszerek lényegét. E két részteszt az előbbieken már említett tagadás és verbális rendszerezés szubteszt volt. Az első a logikai, a második pedig a rendszerezési művelleti képességek rendszeréhez tartozik. Megjegyezzük, hogy elemzéseinkben nem a végleges, hanem egy a kutatás során alkalmazott változatot használtunk. Bár ténylegesen részteszteket vizsgáltunk, a továbbiakban a rövideg kedvéért - mivel ez félreértéshez nem vezethet - a "részteszt" kifejezés helyett a "teszt" megnevezést használjuk.

A tesztek belső szerkezete

Annak ellenére, hogy egy-egy teszt a rendszeren belül már valamilyen specifikus, fogalmilag jól körülhatárolt műveleti képességre vonatkozik és ezért e szempont szerint homogénnek tekinthető a többi teszthez való viszonyát tekintve, önmagában véve mégis meghatározott strukturát takar. E struktúra a teszt elemeinek olyan csoportokba sorolását jelenti, amelyeket az jellemez, hogy egy-egy csoporton belül egymással fogalmi, illetve a mérésekkel igazolható empirikus kapcsolatban vannak. A fogalmi csoportosítás már lényegében a teszt kialakításának szakaszában megtörténik, hiszen az elemek nyilvánvalóan nem véletlenszerűen kerülnek a tesztbe. A tesztbe tartozó elemek, illetve a szándékosan belevitt fogalmi struktúra a kísérleti mérések szakaszában természetesen jelentős változásokon, módosulásokon mehet keresztül. Végső formában a fogalmi strukturának meg kell egyeznie az empirikus strukturával. Empirikus struktúrán a tesztben szereplő elemek olyan csoportosítását értjük, amelyben az egyes csoportokhoz tartozó elemek egymással szoros sztochasztikus kapcsolatban állnak a mérési eredmények alapján, a különböző csoportokba tartozó elemek pedig statisztikusan közel függetlenek. Az empirikus struktúra feltárására több egymástól eltérő, egymást kiegészítő módszer ismeretes. A "Többváltozós vizsgálati módszerek alkalmazási lehetőségei a pedagógiai jelenségek vizsgálatában" című cikkben /Hunyáné, 1982/ három alapvető módszert mutattunk be egy pszichológiai mérés eredményeinek elemzése kapcsán. Ezek közül itt most a két tesztre a többdimenziós skálázás /multidimensional scaling/ módszerét alkalmaztuk. Ennek lényege az, hogy két /három vagy több/ dimenziós térben helyezik el a teszt elemeit a köztük lévő statisztikus kapcsolat alapján, megadva az elemeknek megfelelő pontok koordinátáit. Ezen adatok alapján a pontok síkvetületeit egy-egy kétdimenziós koordinátarendszerben ábrázolva az elemek csoportjai, illetve egymáshoz való viszonyuk igen

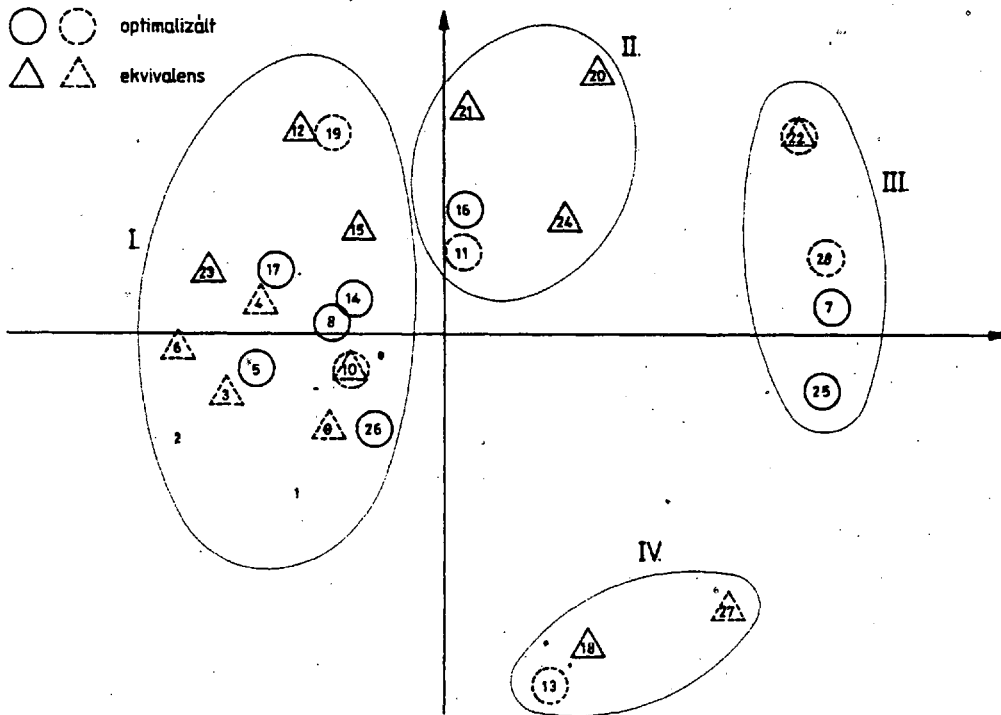
szemléletesen jelenik meg, vizuálisan is nyilvánvalóakká válik számunkra.

Ellenőrzésképpen a tagadás tesztre faktoranalízist is végeztünk, és ez jelenlegi célkitűzésünket tekintve ugyanarra az eredményre vezetett, mint a többdimenziós skálázás. A szokásostól eltérően ennek most nem számszerű eredményeit közöljük, hanem - a könnyű összevethetőség érdekében - a faktorsúlyokat ábrázoljuk grafikusán.

A tagadás teszt 28 elemet tartalmaz, melyek egy egyszerű kijelentőmondat /"Éva ír."/ tagadással képzett különböző változatait tartalmazzák. A tanulók feladata az volt, hogy megállapítsák, a változatok azonosak-e vagy egymás tagadásai.

A többdimenziós skálázásnál két és háromdimenziós megoldást is meghatároztunk, de mivel ezek nem térnek el lényegesen egymástól és a kétdimenziós könnyebben áttekinthető, ezt szemléltetjük az 1.sz. ábrán. A számításokhoz az OSIRIS programcsomag MDSCAL programját használtuk fel. A háromdimenziós megoldás ugyan finomabban mutatja az I és II-vel jelölt csoportok szétválását, de ez felismerhető és jól szemléltethető a kétdimenziós ábrán is. A 28 elemet négy csoportba sorolhatjuk, amelyek rendre 16, 5, 4 és 3 elemet tartalmaznak. A csoportokat az ábrán I, II, III, és IV-gyel jelöltük. Bár feladatkitűzésünk szempontjából a csoportok tartalma /a fogalmi struktúra/ közömbös, mégis érdemes megjegyezni azt, hogy az I és II csoportok az egyszerűbb változatokat tartalmazzák, a III és IV csoportokban pedig olyan bonyolult változatok fordulnak elő, amelyekben három tagadás van, közülük kettő az állítmányra, egy az alanyra vonatkozik. Ezen túl jellemző az, hogy a helyes válaszok gyakorisága az egyes csoportokban közel azonos és csoportonként lényegesen különböző /1.sz. táblázat/.

TAGADÁS TESZT SZERKEZETE (többdimenziós skálázás)



1. sz. ábra

A helyes válaszok arányának alakulása a tagadás
teszt strukturális egységeiben

Csoport	Elemek száma	Minimum	Maximum	Átlag
I	16	66	90	77,75
II	5	62	71	64,80
III	4	40	47	42,50
IV	3	49	50	49,33

1.sz. táblázat

A faktoranalízisben a háromtényezős megoldás /amely ugyan mindössze 32 % varianciát magyaráz meg /már világosan mutatja a többdimenziós skálázásnál is jelentkező négy csoportot. A második és harmadik faktor síkjában jól látható ennek az empirikus strukturának a megjelenése /2.sz. ábra/. A faktor-kép előállításához az MLIFO4 programot használtuk.

A helyes válaszok arányának alakulása a verbális
rendszerezés teszt strukturális egységei-
ben

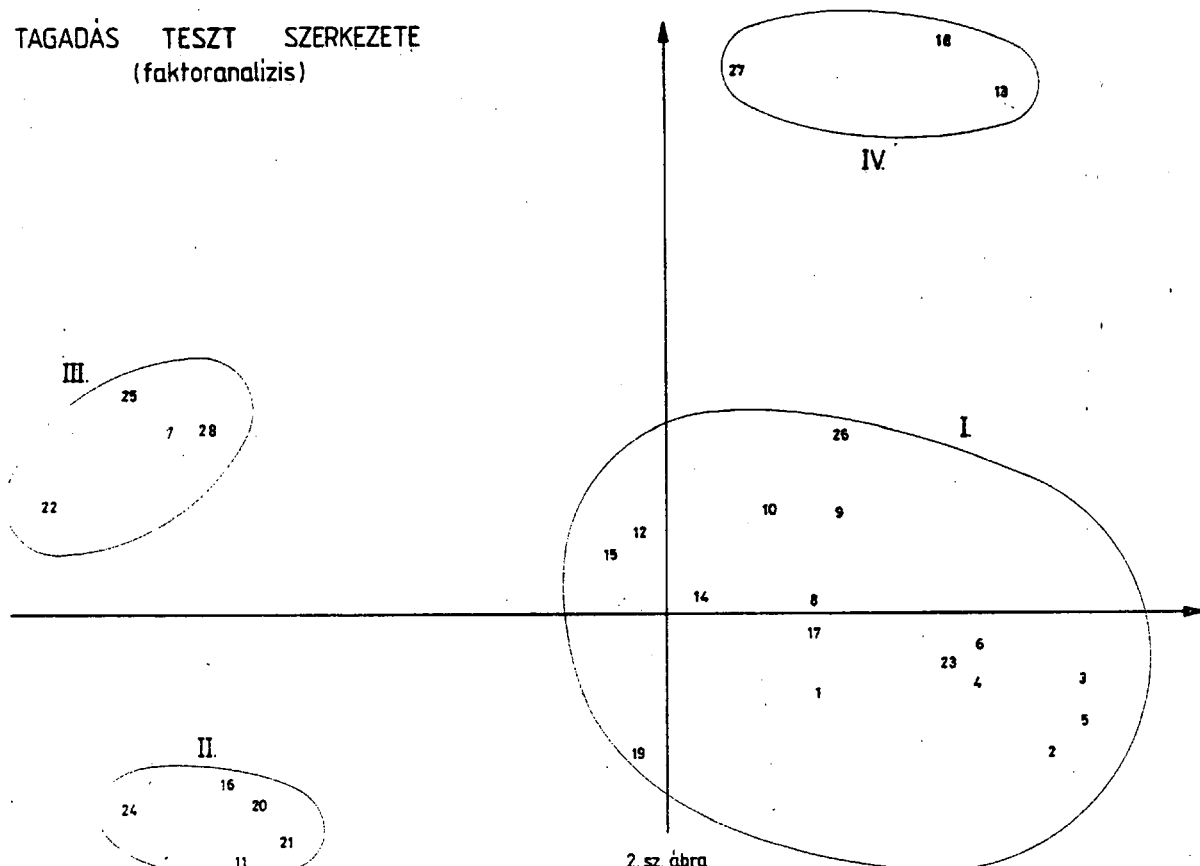
Csoport	Elemek száma	Minimum	Maximum	Átlag
I	8	77	87	80,12
II	4	55	64	60,25
III	2	31	37	34,00
IV	6	5	43	21,83

2.sz. táblázat

A verbális rendszerezés tesztben ugyancsak négy csoportot találtunk /a strukturális egységek számának egyezése a két

TAGADÁS TESZT SZERKEZETE
(faktoranalízis)

77



2. sz. ábra

tesztben csak a véletlen műve!/. Ezek a 3.sz. ábrán szemléltetett csoportok tökéletes egységben vannak a részteszt fogalmi /és ebből eredő formai/ szerkezetével. Jól elválnak egymástól a halmazműveleti és a definiálási feladatok, valamint az előbbieken belül is az egyesítési, a felosztási és a metszetképzési elemek. Az egyes csoportok elemeire kapott helyes válaszok arányának alakulását a 2.sz. táblázat mutatja.

Elemiszámra optimalizált teszt meghatározása

A tesztben nyújtott tanulói összteljesítményt az elemekre adott pontértékek összegével jellemezzük. Jelölje $x_1^T, x_2^T, \dots, x_{28}^T$ a tagadás, $x_1^V, x_2^V, \dots, x_{20}^V$ a verbális rendszerezés teszt elemeire adott pontértékeknek megfelelő véletlen változókat, T^T és T^V pedig a tesztenkénti összteljesítményt. Ekkor formálisan:

$$T^T = \sum_{i=1}^{28} x_i^T \quad \text{és} \quad T^V = \sum_{i=1}^{20} x_i^V$$

Bár önmagában sem a huszonnyolc, sem a husz elemű feladat nem okozna nehézséget a mérés során, ha a műveleti képességek teljes rendszeréről kívánunk információt kapni, a megoldandó elemi feladatok száma olyan nagy lesz, hogy a mérés elfogadható időn belül nem kivitelezhető. Éppen ezért a tesztek elemeinek számát csökkentenünk kell a gyakorlati alkalmazhatóság érdekében úgy, hogy a kapott információ azért ne csökkenjen lényegesen. Valamilyen csökkenést - általános esetben - mindenképpen meg kell engednünk.

Ez más szavakkal azt jelenti, hogy az

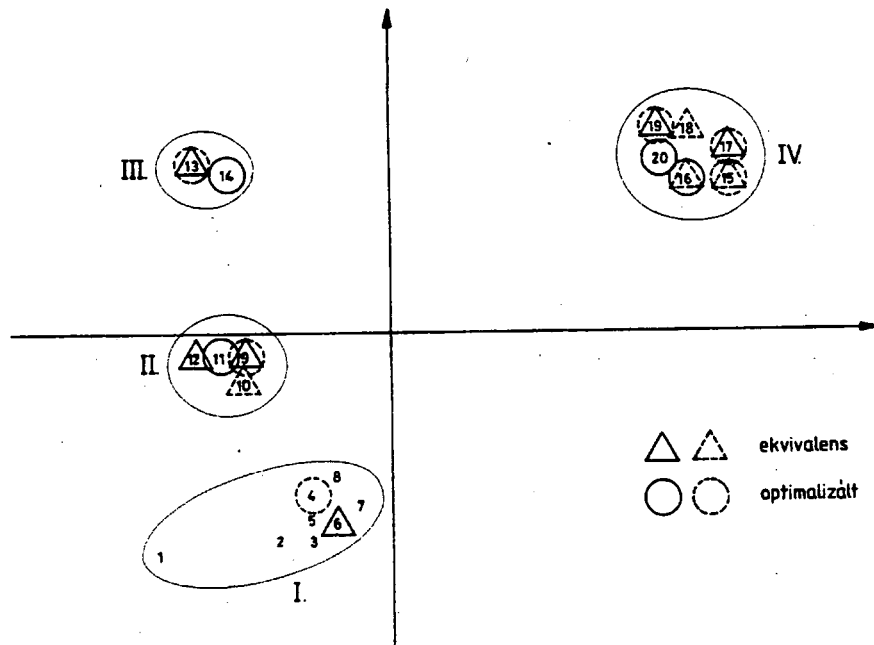
$$X^T = \{ x_1^T, x_2^T, \dots, x_{28}^T \}$$

illetve az

$$X^V = \{ x_1^V, x_2^V, \dots, x_{20}^V \}$$

halmazokból olyan részhalmazokat kell kiválasztanunk, melyekben szereplő változók felhasználásával a T^T és T^V tel-

VERBÁLIS RENDSZEREZÉS TESZT SZERKEZETE
(többdimenziós skálázás)



3. sz. ábra

jesítmények statisztikai értelemben jól becsülhetők. Formálisan fogalmazva, keressük azon

$$X_{\text{opt.}}^T = \{ x_{i_1}^T, x_{i_2}^T, \dots, x_{i_t}^T \}$$

és

$$X_{\text{opt.}}^V = \{ x_{j_1}^V, x_{j_2}^V, \dots, x_{j_r}^V \}$$

részhalmazokat, melyekre

$$T^T \approx T_{\text{opt.}}^T = F^T / x_{i_1}^T, \dots, x_{i_t}^T /,$$

valamint

$$T^V \approx T_{\text{opt.}}^V = F^V / x_{j_1}^V, \dots, x_{j_r}^V /,$$

továbbá természetesen

$$X_{\text{opt.}}^T \subset X^T \quad \text{és} \quad X_{\text{opt.}}^V \subset X^V.$$

Tekintetbe véve T^T és T^V előállítás módját jogos az F^T és F^V függvények linearitását feltételeznünk. Ez azt jelenti, hogy keressük azon $x_{i_k}^T$ és $x_{j_l}^V$ változókat, valamint w_k^T és w_l^V súlyokat, melyekre

$$T_{\text{opt.}}^T = \sum_{k=1}^t w_k^T x_{i_k}^T \approx T^T$$

és

$$T_{\text{opt.}}^V = \sum_{l=1}^r w_l^V x_{j_l}^V \approx T^V.$$

Nem fogalmaztuk meg még eddig azt, hogy mit értünk "közel egyenlő"-ségen és azt sem mondtuk meg, hogy mekkora legyen t és r , azaz az optimalizált tesztekben szereplő elemek száma. Az egyszerűség kedvéért a formulákból a továbbiakban a tagadás tesztre utaló T és a verbális rendszerezés tesztre utaló V felső indexet elhagyjuk. Ezt a félreértés veszélye nélkül megtehetjük, mivel a feladat matematikai megfo-

galmazása mindkettőre azonos.

A T_{opt} és T kapcsolatának - esetünkben egyezésének - mérésére jól használható korrelációs együtthatójuk, illetve annak négyzete /a megmagyarázott variancia részaránya/. Ennek segítségével már egzakt módon megfogalmazhatjuk az egyezésre vonatkozó követelményünket, mégpedig abban a formában, hogy a megmagyarázott variancia meghaladjon egy előre meghatározott szintet:

$$R^2 / T, T_{opt} / \geq R_{\phi}^2,$$

R_{ϕ}^2 jelöli a kiválasztott szintet, esetünkben ez 80 % ill. 95 % volt. Ezekután az optimalizációs feladat úgy fogalmazható, hogy keressük azt a minimális t -t és keressük X azon t elemű X_{opt} részhalmazát, valamint azon w_k súlyokat, amelyekre

$$R^2 / T, \sum_{k=1}^t w_k x_{i_k} / \geq R_{\phi}^2$$

teljesül.

E megfogalmazásban a jólismert lépésenkénti regresszió feladatára /Köves-Párniczky, 1980/ vezettük vissza optimalizációs problémánkat. Ennek modellje kétváltozós regressziós feladatból indul ki, amelyben a függő változó a T , a független pedig az X halmaz valamely x_i változója. A számítás azzal az x_{i_1} változóval kezdődik, amelyre a megmagyarázott variancia x_{i_1} önmagában a legnagyobb. Ezt követően lépésenként történik a modell bővítése, amely minden lépésben azzal a változóval egészül ki, amelyre az előzőleg már bevontak figyelembevételével képzett parciális korrelációs együttható négyzete maximális. Más szavakkal, azt a független változót vonjuk be, amelyik a megmagyarázott varianciát a legjobban növeli. Ezt követően minden lépésben megvizsgáljuk a már bevont változókat hatásuk szignifikanciája tekintetében és kiejtjük azt, amelyiknek hatása - a többi bevont együttes hatásához viszonyítva - nem lényeges. Ezt az eljárást /a bevont változók számának növelésével/ addig

folytatjuk, míg a modellben a megmagyarázott variancia aránya el nem éri az előre megadott R_0^2 értéket. A lépésenkénti kiválasztás eredménye egy t elemszámú $X_{\text{opt}} = \{x_{1_1}, \dots, x_{1_t}\}$ változóhalmaz és egy w_1 és w_2, \dots , súly-sorozat. Ezt az eljárást valósítja meg - több finomítással - az OSIRIS programcsomag REGRESSN elnevezésű programja, melynek alkalmazásával eredményeinket nyertük.

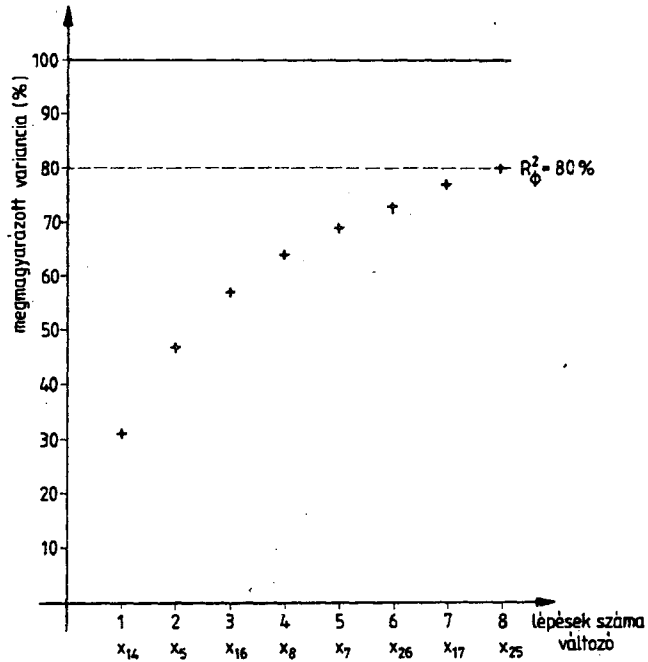
Tekintsük át ezek után a két tesztre vonatkozó optimalizálás eredményeit. Vegyük sorra először a tagadás tesztet, melynél a megmagyarázott varianciaszintet 80 %-ra választottuk, ezt a szintet a modell nyolc lépésben /azaz nyolc változó bevonásával/ éri el. Tehát a /28 elemű/ tagadás teszt egy nyolc elemű részhalmaza az az optimalizált tagadás teszt, amelyik "közel egyenlő" az eredetivel abban az értelemben, hogy az általa megmagyarázott variancia részaránya 80 % /mint ismeretes, ez kb. 0,9-es korrelációt jelent a teszt és a rövidített változat között/. Formálisan kifejezve:

$$R^2 / T^T, T_{\text{opt}}^T / \geq 80 \%$$

A 4.sz. ábrán szemléltetjük az optimalizálás folyamatát, lépések szerint megjelölve a bevont változót, és az adott lépéshez tartozó megmagyarázott varianciát. A vízszintes tengelyen a lépések számát, ill. a bevont változók sorszámát, a függőleges tengelyen az adott lépésben elért megmagyarázott variancia részarányát ábrázoljuk. Megjelöltük a rövidített tesztől megkövetelt $R_0^2 = 80 \%$ szintet is /szemben a teljes teszt által megmagyarázott 100 %-kal/.

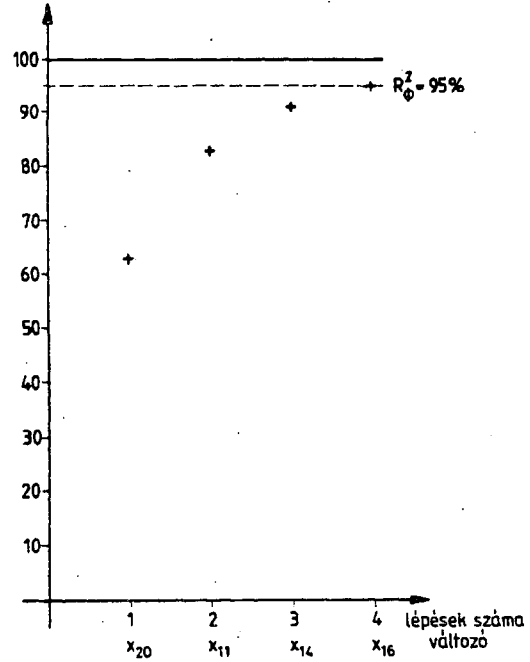
Mint a 4.sz. ábráról jól leolvasható, az optimalizálás első lépésében bevont változó a teszt 14. eleme $/x_{14}^T/$, amely önmaga a tagadás teszt varianciájának 31 %-át magyarázza, ehhez a 2. lépésben bevont 5. elem $/x_5^T/$ 16 %-kal járul hozzá. A 14. és 5. elemek által együttesen megmagyarázott variancia részaránya 47 %, ezt a 3. lépésben a 16. elem $/x_{16}^T/$ 10 %-kal, a 4. lépésben a 8. elem $/x_8^T/$ 7 %-kal növeli stb. Végül a 8. lépésben a 25. tesztelem $/x_{25}^T/$ 3 %-ot ad az előzőkhöz. Az egyes lépésekben bevont változók által oko-

TAGADÁS TESZT



4. sz. ábra

VERBÁLIS RENDSZEREZÉS TESZT



5. sz. ábra

zott variancia-növekedés mindig jelentős /szignifikáns/ volt. Más, már bevont változót sem kellett kiejteni új változó hozzávételekor amiatt, hogy a varianciához való hozzájárulás a szignifikáns szint alá süllyedt volna.

Érdeemes megvizsgálunk azt, hogy az optimalizált /8 elemű/ tagadás teszt változói az eredeti teszt mely csoportjaiba tartoznak. Az általunk megkivánt szinten jól közelítő tagadás tesztbe az I csoportból 5, a II csoportból 1, a III-ból 2 változó került be az eljárás során a lineáris regressziós modellbe, míg a IV-ből egyetlen egy sem /1. ábra, folytonos vonallal rajzolt körrel jelölt elemek/. Mivel az optimalizációs feladatot egy lineáris függvénnyel oldottuk meg, mely követelményeinknek megfelelően jó becslést ad az eredeti tesztre, ezért az optimalizált tesztben szereplő változók felsorolása mellett meg kell adnunk a hozzájuk tartozó súlyokat is.

Optimalizált tagadás teszt / $R_D^2 = 80\%$ /

Változó sorszáma /i _k /	Strukturális csoport	Változó súlya w_k^T
5	I	3,4
7	III	2,0
8	I	2,5
14	I	3,0
16	II	2,3
17	I	2,4
25	III	1,9
26	I	2,4

3.sz. táblázat

A 3.sz. táblázatban a tagadás tesztet jól közelítő T_{opt}^T előállításához szükséges tesztelemek i_k sorszámainak és ezen elemek w_k^T súlyait adtuk meg, melyeknek birtokában az kiszámítható

tó a

$$T_{\text{opt}}^T = \sum_{k=1}^8 w_k^T x_{i_k}^T$$

formulával. A megmagyarázott variancia részaránya ilyen előállítás mellett 80 % lesz. Azt mondhatjuk tehát, hogy a 28 elemű tagadás teszt helyett a mindössze 8 elemű teszttel mérve már sztochasztikusan jó becslését kapjuk az alapteszt-tel való mérésnek.

A becslés jósági kritériumát szigorubban is megválaszthatjuk, nevezetesen megkövetelhetjük azt, hogy az optimalizált teszt által megmagyarázott variancia legalább 90 % legyen, azaz $R_{\emptyset}^2 = 90\%$.

Formálisan: $R^2 / T, T_{\text{opt.}} / \geq 90\%$.

A tagadás teszt esetén ennek a kritériumnak a T_{opt}^T akkor tesz eleget - az alkalmazott eljárás esetén - ha a már meglévő 8 változó mellett még újabb tesztelemeket vonunk be, mégpedig a következő ötöt: $x_{10}^T, x_{13}^T, x_{19}^T, x_{22}^T, x_{28}^T$.

A 28 elemű tagadás tesztnek így kapott 13 elemű optimalizált résztesztje eleget tesz szigorubb feltételünknek, az elemek sorszámát és súlyát a 4.sz. táblázatban adjuk meg.

Optimalizált teszt / $R_{\emptyset}^2 = 90\%$ /

Változó sorszám / i_k /	Strukturális csoport	Változó súlya / w_k^T /
5	I	2,3
7	III	1,3
8	I	2,3
10	I	1,8
13	IV	1,3
14	I	2,2
16	II	2,1
17	I	2,1
19	I	2,2
22	III	1,4
25	III	1,5
26	I	1,9
28	III	1,3

4.sz. táblázat

Ebben a rövidített tesztben az eredeti teszt I csoportjából 7, a II csoportjából 1, a III csoportból 4 és a IV-es csoportból 1 változó szerepel. Érdekes, hogy a III jelzésű részből az eljárás valamennyi /mind a 4/ elemet bevonta a tesztbe, míg a IV-ből és II-ből csak egyet-egyét.

Megjegyezzük, hogy a jósági kritériumot emelhetnénk még tovább is /kérdés: ésszerű-e?/, például úgy, hogy $R_{\theta}^2 = 95\%$ legyen a követelmény. A tagadás teszténél ennek a kritériumnak az eredeti 28 helyett egy 18 elemű optimalizált teszt tesz eleget. Tehát még ennél a viszonylag magas jósági kritériumnál is az optimalizált tesztváltozat az eredetihez viszonyítva 36 %-kal kisebb elemszámúra redukálható.

A tagadás tesztéhez hasonlóan jártunk el az optimalizálást illetően a verbális rendszerezési teszt esetén is. Itt azonban a megmagyarázott variancia szintjét jóval magasabban állapíthattuk meg / $R_{\theta}^2 = 95\%$ /, mert ezt már négy tesztelem bevonásával elérjük. Mindössze négy elemmel elegendő tehát e képességet mérnünk ahhoz, hogy a kapott teljesítmény, a T_{opt}^V jól becsülje T^V -t, formálisan, hogy

$$R^2 / T^V, T_{opt}^V / \geq 95\%$$

legyen. Az 5.sz. ábrán adjuk meg grafikusán a megmagyarázott varianciát, lépésenként megjelölve a bevont tesztelemet. Az első lépésben bevont elem az x_{20}^V , ami önmagában 63 % varianciát magyaráz, ehhez a 2. lépésben az x_{11}^V 21 %-kal, a 3. lépésben az x_{14}^V 7 %-kal és végül az x_{16}^V 4 %-kal járul hozzá. Ezen optimális verbális rendszerezés teszt elemei közül 2 /az x_{16}^V és x_{20}^V / az eredeti teszt strukturájában a IV csoportba, az x_{11}^V a II, az x_{14}^V a III csoportba tartozik. Nem szerepel viszont I-ből egyetlenegy elem sem. /3. ábra, folytonos vonalú körrel jelölt elemek/.

Az optimalizált verbális rendszerezési teszt T_{opt}^V értékét az 5.sz. táblázatban megadott sorszámú elemek és a hozzájuk tartozó súlyok lineáris kombinációja adja meg. Ez az előzők szerint igen jó becslése a T^V -nek. Akkor is jó képet kapunk tehát a verbális rendszerezési képességről, ha

az eredeti kérdéseknek mindössze 25 %-át tesszük fel a tanulóknak.

Optimalizált verbális rendszerezés

$$/R_{\emptyset}^2 = 95 \%/$$

Változó sorszám $/i_k/$	Strukturális csoport.	Változó sulya $/w_k^v/$
11	II	3,9
14	III	1,4
16	IV	3,2
20	IV	2,8

5.sz. táblázat

A két tesztre kapott eredményeket összehasonlítva megállapíthatjuk azt, hogy ugyanazon jósági követelmény-szintet támasztva a különböző tesztek az elemszámot tekintve eltérő mértékben redukálhatók. Míg a tagadás tesztnél $R_{\emptyset}^2 = 95 \%$ mellett a csökkentés mindössze 36 % volt, ugyanilyen szintnél a verbális rendszerezés teszt esetén a kérdések száma 75 %-kal redukálható!

Ekvivalens optimalizált tesztek

A nagyméretű teszteknel egy optimális, rövidített változat előállításával mellett a bevezetőben már megjelölt másik - pedagógiai szempontból - fontos feladat a tőle különböző optimális tesztváltozatok létrehozása is. Arra, hogy ilyen változatok léteznek és meg is találhatók, az ad eredményt, hogy az alaptesztek strukturális egységeiben vannak az előzőleg optimalizáltba be nem választott elemek, mégpedig olyanok, melyek a struktúra szerint rokonságban vannak az oda bevontakkal. Az optimális tesztvariáns előállításánál feladatunk tehát először az, hogy kihagyjuk azokat a változókat, amik az előző optimalizálásnál az eredmény-teszt-

be kerültek és a megmaradt elemekre megismételjük a lépésenkénti regressziót. Ha valamely csoportban nem maradt volna elegendő változó, olyanokat is felhasználhatunk e lépésben is, melyek az előző optimalizált tesztben már szerepeltek.

Először a tagadás tesztet jól becsülő optimalizált teszttel ekvivalens változat előállítását kíséreljük meg. Az alaptesztet alkotó 28 elemből - az optimalizált tesztbe bevont 8 változót kihagyva - a visszamaradó 20 elemből próbálunk kiválasztani az előzőkhöz hasonlóan valahányat úgy, hogy az így előállított új rövidített teszttel mérve a kapott teljesítmény, $T_{1,opt}^T$ eleget tegyen annak a követelménynek, hogy

$$R^2 / T^T; T_{1,opt}^T / \geq 80 \%$$

legyen. Itt nem kell a korábban már kiválasztottak közül egyet sem ismét tekintetbe vennünk, mivel mindegyik strukturális egységben maradt várhatóan elegendő elem. A lépésenkénti regressziót a 20 változóra elvégezve az adódott, hogy a megmagyarázott variancia aránya a 9. lépés - a kilencedik bevont változó - után haladja meg a 80 %-ot. Az eredményül kapott ekvivalens változatba tartozó elemek sorszámait és a hozzájuk tartozó súlyokat a 6.sz. táblázat adja meg; elhelyezkedésüket a teljes teszt strukturájában - folytonos vonallal rajzolt háromszöggel jelölve - az 1.sz. ábrán szemléltetjük.

Az e változatban szereplő elemek száma nem tér el lényegesen az első lépésben nyert optimalizált változattól /8 helyett 9/, így a mérés ideje szempontjából ugyanolyan használati értékű eszközhöz jutottunk. Ha azonban összehasonlítjuk az elemek strukturális szempontok szerinti megoszlását, már jelentős különbséget találunk.

Tagadás teszt ekvivalens változat

$$/R_{\emptyset}^2 = 80 \%/$$

Változó sorszám $/i_k/$	Strukturális csoport	Változó sulya $/w_k/$
4	I	2,8
9	I	2,6
12	I	2,1
15	I	2,5
18	IV	2,3
20	II	2,0
21	II	1,9
23	I	2,3
24	II	2,5

6.sz. táblázat

Míg az első változat a IV csoportból, addig a második a III csoportból nem tartalmaz egy elemet sem. E problémára a validitás tárgyalásánál később még visszatérünk. Elsődleges optimalizációs kritériumunkat, a megmagyarázott variancia részarányát tekintve azonban két egyenértékű rövidített teszt-variánst kaptunk.

A fentiekhez hasonlóan járunk el a verbális rendszerezés teszt esetén is. Itt az optimalizált változatban szereplő négy elemet kihagyva, 16 változóból kiindulva hajtottunk végre lépésenkénti regressziót. Az $R_{\emptyset}^2 = 95 \%$ -os szintet a 6. lépésben értük el. Az eredményt részleteiben a 7.sz. táblázat adja meg, az e változatba tartozó elemeket a 3.sz. ábrán folytonos vonallal rajzolt háromszöggel jelöltük.

Verbális rendszerezés teszt ekvivalens
változat

$$/R_{\emptyset}^2 = 95 \%/$$

Változó sorszáma /i _k /	Strukturális csoport	Változó sulya /w _k /
6	I	6,0
9	II	2,5
12	II	1,9
13	III	2,5
17	IV	3,0
19	IV	2,6

7.sz. táblázat

E változatra is állnak a tagadás tesztre fentebb mondtak, nevezetesen az, hogy az ekvivalencia az R_{\emptyset}^2 szint tekintetében teljes egészében fennáll, de ugyanez nem mondható el sem az elemszámról, sem a strukturáról /csak az utóbbi változat tartalmaz elemet az I csoportból/. Természetesen további optimális tesztvariánsok előállítására is van lehetőség, de ezek meghatározásánál már nem hagyhatjuk ki automatikusan a korábban felhasznált összes elemet, hiszen ezzel teljes strukturális egységeket hagynánk ki mérőeszközünkéből. A verbális rendszerezés teszténél ez következne be a III elemcsoportnál /metszetképzés/, ami nem csak formális, hanem tartalmi validitási problémákhoz is vezetne. Ilyenkor e strukturális egységekből ismét fel kell használni legalább annyi elemet, amennyi az előző lépésekben nyert változatokban szerepelt. Ez ugyanis egy tapasztalati irányszám arra, hogy a rövidített változatban milyen súllyal kell szerepelnie e szerkezeti egységnek.

A rövidített tesztek validitása és reliabilitása

Térjünk most vissza az előző részben már említett validitás kérdésére. A külső validitást illetően, amennyiben a kiindulásul szolgáló résztesztekhez lenne külső kritérium-változónk, azt használhatnánk fel a validitás megítéléséhez. Ilyenkor alapjában véve ez a kritérium-változó lehetne a lépésenkénti regresszió függő változója a teljes tesztben nyújtott teljesítmény helyett. Ez az optimalizálás természetesen nem vezetne szükségképpen ugyanahhoz az eredményhez /más az optimalitás kritériuma/. Ha azonban a teszt felépítése olyan, hogy tartalmilag teljesen lefedi a mérni kívánt területet és az empirikusan is feltárt strukturális egységek arányai /maximális pontértékek/ is megfelelnek a mérendőnek, akkor magát a tesztértéket fogadhatjuk el validitási kritériumnak. E feltételezéssel élve két - műveleti képességet mérő - tesztünknel nem csupán annyit állíthatunk, hogy a rövidített tesztek statisztikusan jól mérik ugyanazt a képességet, mint amit az eredeti tesztek, hanem azt is kimondhatjuk, hogy meghatározott validitási szinthez kaptunk minimális elemszámú rövidített változatokat. Ilyenkor tehát az ismertetett eljárás egyben a validitás vonatkozásában is optimumot szolgáltat, természetesen a lépésenkénti regressziós eljárás korlátain belül /Éltető - Meszéna - Ziermann, 1982/.

A belső szerkezeti validitás vonatkozásában már az előző részben láttuk, hogy a kapott eredmények nem teljesen kielégítőek. Vannak ugyanis mindkét tesztben olyan szerkezeti egységek, amelyekből a rövidített változatok nem tartalmaznak elemet. Másrészt az elemszám csökkenése óhatatlanul a reliabilitás /megbízhatóság/ csökkenéséhez vezet /Guilford, 1965/.

A 6. ábrán azt mutatjuk be, hogy az elemszám negyedére, illetve felére csökkentése esetén a rövidített változat reliabilitása r_{TT} / milyen függvénye az eredeti reliabilitásnak r_{TT}^{opt} /.

E függést a 8. táblázatban néhány értékre számszerűen is megadtuk. A 0,90-es reliabilitás például az elemszám fele-

zésével $/r_{TT_{1/2}} / 0,82$ -re, negyedelésével pedig $/r_{TT_{1/4}} / 0,69$ -re csökken.

Rövidített teszt reliabilitása

r_{TT}	$r_{TT_{1/2}}$	$r_{TT_{1/4}}$
0,00	0,00	0,00
0,25	0,14	0,08
0,50	0,33	0,20
0,75	0,60	0,43
0,90	0,82	0,69
1,00	1,00	1,00

8.sz. táblázat

Erre való tekintettel célszerűnek tűnik a rövidített tesztek elemszámának növelése. Feladatunkat most úgy fogalmazhatjuk, hogy keressük a tesztelemek azon meghatározott elemszámú részhalmozát, amelyre lineáris regresszióval a teljes tesztben nyújtott teljesítmény a legjobban becsülhető, ugyancsak a megmagyarázott variancia részarányát véve alapul. E feladatot is a korábbiakban alkalmazott lépésenkénti regresszió módszerével oldhatjuk meg, de most nem R_{ϕ}^2 -hez /a szinthez/ keressük a t elemszámot, hanem adott t -hez választjuk a tesztelemeket és ezzel R_{ϕ}^2 -et. Ha mindkét vizsgált tesztünkönél megállapodunk abban, hogy az elemszámot a féltre /14 ill. 10/ csökkentjük, mind a külső validitást, mind a szerkezeti validitást és a reliabilitást illetően jobb eredményre számíthatunk, mint az előzőekben ismerttetett megoldásoknál, hiszen azok elemszáma lényegesen kisebb volt /8 és 9, illetve 4 és 6 /. Nem követelhetjük meg ezzel szemben, hogy a különböző változatok között ne legyen átfedés. Meg kell engednünk ezért, hogy az ekvivalens változatok keresésénél bizonyosak az első optimalizált válto-

zatból ismét szerepeljenek. Csak azokat zárjuk ki, amelyek a legkisebb elemszámu rövidített változatban szerepeltek az előző /minimális R_{\emptyset}^2 -et megkövetelő/ elemzéseknél. A többi változó bármely változatban előfordulhat. Ettől az elvtől példáink egyetlen ponton térnek el: a verbális rendszerezés teszt negyedik csoportjának változóit mindkét változatnál megengedtük.

Az eredményeket az előző részeknél alkalmazott formában a 9. és 10. táblázat tartalmazza.

Rövidített tagadás teszt /t=14/

Optimális változat
 $R_{\emptyset}^2 = 90 \%$

Ekvivalens változat
 $R_{\emptyset}^2 = 90 \%$

Változó sorszáma	Strukturális csoport	Változó súlya	Változó sorszáma	Strukturális csoport	Változó súlya
5	I	2,2	3	I	2,2
7	III	1,2	4	I	1,9
8	I	2,1	6	I	1,7
10	I	1,5	9	I	2,0
11	II	1,5	10	I	1,6
13	IV	1,5	12	I	1,8
14	I	2,1	15	I	1,6
16	II	1,8	18	IV	1,9
17	I	2,0	20	II	1,8
19	I	2,1	21	II	1,5
22	III	1,2	22	III	1,5
25	III	1,4	23	I	1,7
26	I	1,9	24	II	2,3
28	III	1,3	27	IV	1,4

9.sz. táblázat

Mint a 6-os ábráról és a 7-es táblázatról jól leolvasható, az elemszám felezése nem vezet megengedhetetlen mértékű csökkenéshez. Ezzel tehát jobb változatokhoz jutottunk, az eredeti tesztekkel való egyezés a tagadásnál lényegesen $R_{\emptyset}^2 = 90 \%$, a verbális rendszerezésnél kisebb mértékben $R_{\emptyset}^2 = 99 \%$ és $R_{\emptyset}^2 = 97 \%$ javult. A legtöbbet azonban talán a szerkezeti validitás terén nyertük. E változatoknál ugyanis már mindegyik strukturális egység megjelenik.

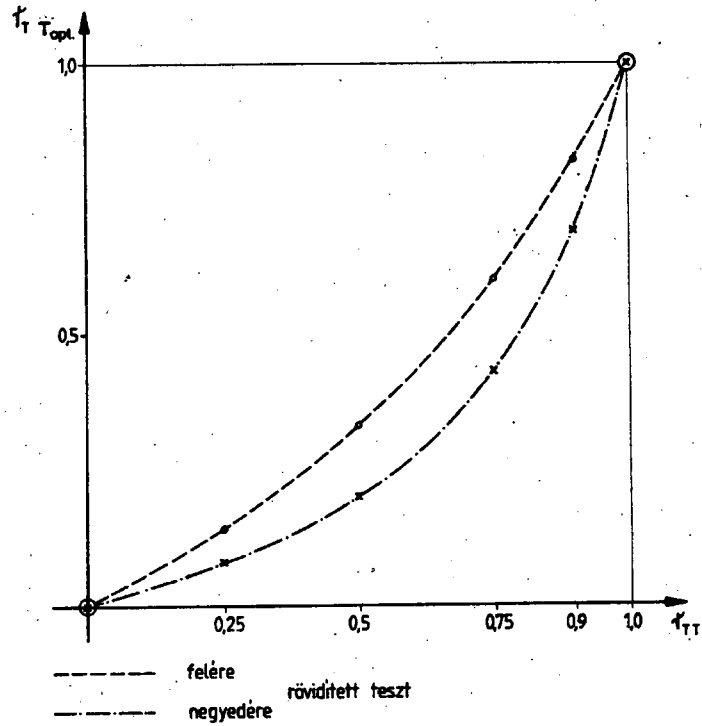
Rövidített verbális rendszerezés teszt / $t=10$ /

Optimális változat $R_{\emptyset}^2 = 99 \%$			Ekvivalens változat $R_{\emptyset}^2 = 97 \%$		
Változó sorszáma	Strukturális csoport	Változó súlya	Változó sorszáma	Strukturális csoport	Változó súlya
4	I	4,5	6	I	5,5
9	II	1,1	9	II	1,5
11	II	2,4	10	II	2,1
13	III	1,0	12	II	1,4
14	III	1,1	13	III	2,4
15	IV	1,1	15	IV	1,2
16	IV	1,1	16	IV	1,2
17	IV	1,4	17	IV	1,2
19	IV	1,2	18	IV	1,1
20	IV	1,2	19	IV	1,5

10.sz. táblázat

A két különböző - a megmagyarázott varianciát, illetve az elemszámot rögzítő - megközelítés eredményeit e vonatkozásban a 11. és 12. táblázatban mutatjuk be számszerűen. A bevont elemek elhelyezkedését a tesztek strukturájában ismét az 1. és 3. ábrán szemléltetjük s szaggatott vonallal rajzolt körrel, illetve háromszöggel jelölve az optimalizált,

RÖVIDÍTETT TESZT RELIABILITÁSA



6. sz. ábra

illetve az ekvivalens változatban megjelenő újabb elemeket, E lépésben kapott eredményeink ugyanis természetesen tartalmazták azokat az elemeket, amelyek a kisebb elemszámú rövidített változatokban szerepeltek.

Teszttelemelek száma a tagadás teszt
strukturális csoportjaiban

Strukturális csoport	Optimalizált változat		Ekvivalens változat	
	Rögzített szint $R_{\emptyset}^2 = 80\%$	Rögzített elemszám $t=14$	Rögzített szint $R_{\emptyset}^2 = 80\%$	Rögzített elemszám $t=14$
I	5	7	5	8
II	1	2	3	3
III	2	4	0	1
IV	0	1	1	2
Összesen	8	14	9	14

11.sz. táblázat

Teszttelemelek száma a verbális rendszerezés teszt
strukturális csoportjaiban

Strukturális csoport	Optimalizált változat		Ekvivalens változat	
	Rögzített szint $R_{\emptyset}^2 = 95\%$	Rögzített elemszám $t=10$	Rögzített szint $R_{\emptyset}^2 = 95\%$	Rögzített elemszám $t=10$
I	0	1	1	1
II	1	2	2	3
III	1	2	1	1
IV	2	5	2	5
Összesen	4	10	6	10

12.sz. táblázat

Konkluzió

Az optimalizációs feladatot tehát kétféleképpen is megfogalmaztuk és példánkon bemutattuk a lépésenkénti regresszió módszerével kapott megoldások eredményét. Megállapíthatjuk, hogy a méréseknél alkalmazott tesztelemelek száma lényegesen csökkenthető anélkül, hogy a kapott információ alapvetően megváltozna. Továbbmenve, mindkét változatnál több rövidített tesztvariánst is bemutattunk. Ha az optimalizálást olyan teszteken végezzük, amelyek konstrukciójukból adódóan önmagukban a külső validitást is biztosítják, a bemutatott módszer egyben maximális validitású rövidített variánsokhoz vezet. Ha ez nem áll fenn, akkor annyit állíthatunk, hogy a rövidített tesztek jó becslései az eredetieknek, de a külső validitást külön ellenőriznünk kell. A reliabilitást és a szerkezeti validitást is figyelembe véve azt a megközelítést tarthatjuk célravezetőbbnek, amelynél - a megengedett reliabilitás-csökkenés figyelembevételével megállapított - rögzített elemszám mellett keressük az optimális, rövidített változatokat.

Irodalom

ÉLTETŐ Ödön - MESZÉNA György - ZIERMANN Margit:

Sztochasztikus módszerek és modellek. Közgazdasági és Jogi Könyvkiadó Bp., 1982.

GUILFORD, J.P.: Fundamental statistics in psychology and education. Mc Graw-Hill, New York, 1965.

HUNYA Péterné: Többváltozós vizsgálati módszerek alkalmazási lehetőségei a pedagógiai jelenségek vizsgálatában. Pedagógiai Szemle, XXXII. évf. 1.szám 1982.

KÖVES Pál - PÁRNICZKY Gábor: Általános statisztika I-II. Közgazdasági és Jogi Könyvkiadó, Bp., 1981.

MLIF04: Principal component factor analysis. JATE Kalmár L. Kib.lab. 1984. /kézirat/

NAGY József: A műveleti képességek rendszere. Acta Univetsitatis Szegediensis de A. József nominatae Sectio Pedagogica et Psychologica 25., Szeged, 1983.

OSIRIS III; an integrated collection of computer programs, The University of Michigan, 1974.

Хуня Петернэ

Выбор эквивалентных тестов с минимальным элементарным числом

Автор пытается ответить на вопрос -- каким образом можно найти такое оптимальное измерительное средство из тестов, примененных в ходе исследований и показавших себя результативными, которое с хорошим приближением измеряет то, что и первоначальный тест, но, по возможности, с наименьшим количеством вопросов /элементов/. Автор исходит из того, что тесты имеют такие группы элементов, которые измеряют различные, друг от друга независимые предметы, но внутри групп эти элементы предусматривают измерение тождественных предметов /понятие, способность и т.п./. В ходе анализа тестов группы довольно хорошо отделимы друг от друга и в эмпирическом плане. Обладая информацией, они решают оптимизационную задачу. Оптимизированный тест считается максимально дающим информацию в том случае, если его корреляция по первоначальному /исходному/ тесту довольно высока. Оптимизационный метод показывается на субтесте, исследующем две операционные способности, а именно: на субтесте отрицания теста логики и на субтесте вербальной классификации теста, определяющего классификационные способности. После оптимизации возникает также вопрос, возможно ли составление большего количества тестов, эквивалентных между собой, исходя при этом из одного и того же основного теста.

Две способности в случае субтеста предоставляют оптимальным тестом эквивалентные тестовые варианты. Эквивалентным с оптимизированным тестом считается такой, если в каждой группе, определяющей основную структуру первоначального теста он содержит элемент /вопрос/ и корреляция этого теста с первоначальным идентична корреляции оптимизированного. Оптимизация осуществляется поэтому с помощью

linenar der regressien, genau es je ispol'zuetesja i pri so-
stavlenii ekvivalentnyx testov, dopolnjasja analizom struk-
tury /s mnogozismeritel'noj skaloj i analizom faktorov/.

Hunya Béterné

Die Auswahl äquivalenter Teste von minimaler Zahl
der Elemente

Die Verfasserin sucht Antwort auf die Frage, wie man aus den bei den Forschungen erfolgreich verwendeten und sich als erfolgreich erwiesenen Teste vor der generellen Anwendung einen optimalen Vermessungsapparat zusammensetzt, der in einer guten Annäherung dasselbe erfasst wie der Originaltest, allerdings mit möglichst wenigen Fragen. Sie geht davon aus, dass die Teste solche Gruppen von Elementen enthalten, die verschiedene, voneinander unabhängige Sachen ermitteln, die Elemente zielen aber in der Gruppe auf die Ermittlung nahezu identischer Inhalte /Begriff, Fähigkeit.../ ab. Im Laufe der Testanalyse lassen sich die Gruppen voneinander auch empirisch gut absondern. Im Besitz dieser Informationen löst die Verfasserin die Aufgabe der Optimierung. Sie erachtet den optimalisierten Test erst dann als einen optimalen Informationsträger, wenn seine Korrelation zum Originaltest entsprechend hoch ist. Das Optimalisierungsverfahren wird an zwei Subtesten dargestellt, die die Operationsfähigkeit untersuchen: am Negations - Subtest eines logischen Testes sowie am verbalen Systematisierungssubtest, der die Systematisierungsfähigkeit beurteilt. Nach der Optimalisierung wirft sie noch die Frage auf, ob es möglich ist, mehrere äquivalente Teste auf Grund des Basistestes zusammenzustellen.

Im Falle des Subtestes stellen die beiden Fähigkeiten Testvarianten her, die mit dem optimalen Test äquivalent sind. Sie sieht einen Test mit dem Optimalisierten als äquivalent an, wenn er Elemente /Fragen/ von allen Gruppen enthält, die die wichtigere Struktur des Originaltestes bestimmen, und die Korrelation dieses Testes zum Originaltestes ebenso ist wie die des Optimalisierten.

Die Optimalisierung wird mit der Methode der schrittweisen linearen Regression gelöst; dieselbe Methode wird auch bei der Zusammenstellung der äquivalenten Teste angewandt, die noch mit der Strukturanalyse /Verteilung auf eine Skala von mehreren Dimensionen und Analyse der Faktoren/ ergänzt wird.