

Hunya Péterné

INTERAKTIV TESZTELÉS  
PELADATRENDSZÉREK MINIMALIZÁLÁSA AID MÓDSZERREL

A pedagógiában, pszichológiában használatos nagy terjedelmű tesztek gyakorlati mérőeszközként való alkalmazásával kapcsolatban komoly problémát jelent az, hogy velük a mérés esetenként több /tiz/ órát is igénybe vehet. E probléma feloldásához választ kell keresnünk arra a kérdésre, hogy:

- a nagy méretű tesztből mi módon szerkeszthetjük meg azt az optimális mérőeszközt, mely az eredetiből maximális információt tartalmaz, a lehető legkevesebb számú elemmel /megállapodás szerint: maximális információt nyújtónak tekintjük a rövidített tesztet akkor, ha korrelációja az eredeti tesztrel megfelelően magas/,
- és lehetséges-e több egymással ekvivalens teszt megszerkesztése ugyanabból az alaptesztből kiindulva /az optimalizálttal ekvivalensnek tekintünk egy tesztet, ha az eredeti teszt főbb strukturáját meghatározó csoportok mindegyikéből tartalmaz elemet - kérdést, feladatot - és ezen teszt korrelációja az eredetivel ugyanolyan, mint az optimalizálté/?

A felvetett két probléma megoldásának egy-egy lehetséges módját mutattuk be a "Minimális elemszámú ekvivalens tesztek kiválasztása" című cikkben /Hunyáné, 1984/. Mint kiderült: lehetőségünk van arra, hogy a nagy terjedelmű - egyébként a gyakorlatban mérőeszközként nem alkalmazható - tesztből előállit-

sunk egy, a lehetőséghez mérten az eredetiből maximális információ, ugyanakkor minimális elemszámot tartalmazó tesztet, sőt ezen minimális elemszámú tesztel ekvivalens teszt sorozatot. Így nem kell a teljes tesztet használnunk. Az optimalizált és a vele ekvivalens változatok révén a teszt gyakorlati mérőeszközkénti alkalmazásba vétele lehetővé válik. Eredményünk egyúttal alátámasztja azt a jól ismert tényt, hogy nem kell mindenre rákérdeznünk akkor, hogy egy-egy mérendő területről /művelési képesség, fogalomrendszer, kreativitás, tananyag stb/ reális képet alkothassunk. Ez az elv egyébként régóta jól funkcionál a tanári gyakorlatban is.

További - szintén gyakorlat sugallta - kérdésként vetődik fel az, vajon lehet-e "egyévre szabottan" összeállítani a kérdéssort. Milyen sorrendben és milyen újabb kérdéseket kell feltennünk, ha tudjuk, hogy az előző elemekre /kérdésekre, feladatokra/ hogyan válaszolt a vizsgált tanuló. A matematikai statisztikai kezelhetőség szempontjából pontosabban fogalmazva a következő kérdésekre kell választ adnunk:

- lehetséges-e predikció a teljes tesztre vonatkozóan, ha csak néhány elemre ismerem a kísérleti személy választát;
- és ha igen, mely elemekből, azok milyen sorrendje mellett, mekkora biztonsággal következtethetünk a teljes tesztben várható teljesítményre.

#### Az elemzés módszere

Sok olyan kutatási feladat van, melyben feladatunk az, hogy meghatározzuk, melyik változó kapcsolódik legjobban egy adott jelenséghez; példánkban a logikai és a rendszerezési képesség egy-egy részterületéhez. Arra vagyunk kíváncsiak, hogy az egyes tesztelemekekre adott válaszokkal, a válaszok együttesével hogyan függ össze az e képességekben mutatott teljesítmény. A teljesítmény mérőszámaként az ezen képességeket egészében lefedő tesztek összes elemét figyelembevéve kialakított tesztértéket fogadjuk el. Ezen részteszteknek a helyét egy átfogó tesztrendszerben egy korábbi cikkben ismertettük /Hunyáné, 1984/, ezért erre itt nem térünk ki.

A bevezetésben már megfogalmazott problémánkat a következőképp írhatjuk le formálisan /a leírást a konkrét részteszt-

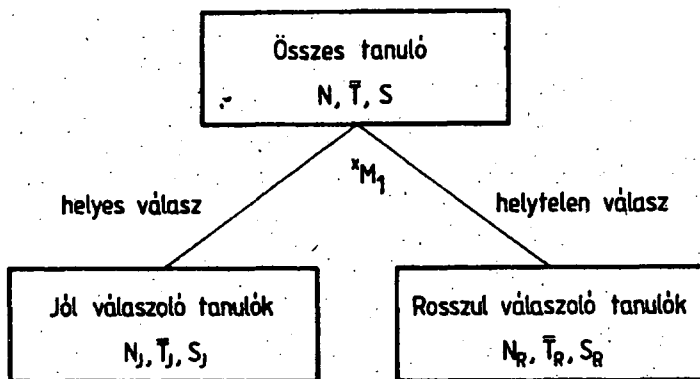
től függetlenül adjuk meg, és mivel ez nem vezethet félreértésre, a továbbiakban részteszt helyett tesztről beszélünk/.

A tesztben mutatott teljesítményt jelöljük  $T$ -vel, és ezt tekintjük feladatunk függő változójának, amelyre becslést kívánunk adni bizonyos tesztelemekek mint független változók, prediktorok segítségével. Jelöljük ezeket  $x_1, x_2, \dots, x_n$ -nel. Keressük azt az  $x_{M_1}$  elemet, amelyből a várható összteljesítményre a maximális<sup>1</sup> információt nyerhetjük. Ha az  $x_{M_1}$  elem kétértékű, a tanulókat két csoportba sorolhatjuk attól<sup>1</sup> függően, hogy az  $x_{M_1}$ -hez tartozó kérdésre helyesen vagy helytelenül válaszoltak-e /l.sz. ábra/. Azon, hogy  $x_{M_1}$  maximális információt hordoz, azt értjük, hogy a teljes teszt érték  $|T|$  varianciájából a legnagyobb hányadot az  $x_{M_1}$  segítségével előállított felbontás adja, azaz az

$$N_J \cdot \bar{T}_J^2 + N_R \cdot \bar{T}_R^2$$

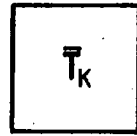
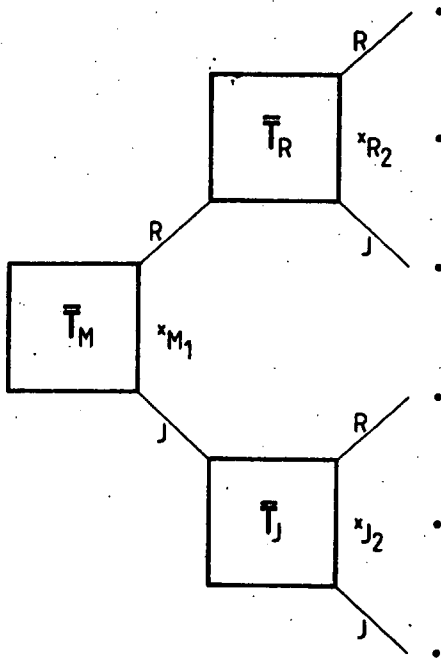
felbontás által megmagyarázott variancia a legnagyobb az összes - tesztelemekek segítségével előállítható - felbontások által elérhetőek közül.  $N_J$  és  $\bar{T}_J$  az  $x_{M_1}$ -nél helyesen válaszolók számát és tesztben elért átlagteljesítményét,  $N_R$  és  $\bar{T}_R$  a helytelenül válaszolók számát és teljesítményét jelöli.

A következő lépésben megvizsgálhatjuk azt, hogy a helyesen válaszolók csoportjában melyik lesz a fenti értelemben vett maximális információtartalmu  $x_{J_2}$  elem. Ha ugyanezt megvizsgáljuk a helytelenül válaszolókrá is, ott egy  $x_{R_2}$  elemet kapunk, amely nem feltétlenül egyezik meg  $x_{J_2}$ -vel, de lehet azzal azonos is. Ez azt jelenti, hogy - ellentétben a lépésenkénti regressziónál kapott eredménnyel - a második feltett kérdés az  $x_{M_1}$  után az  $x_{J_2}$  vagy az  $x_{R_2}$  lesz attól függően, hogy  $x_{M_1}$ -re helyes vagy helytelen választ kaptunk. Természetesen  $x_{J_2}$ -re /és ugyanugy  $x_{R_2}$ -re/ is kaphatunk helyes vagy helytelen feleletet. Ezzel az alapmintában újabb felbontást jelölhetünk ki. Egy csoportba tartoznak azok, akik  $x_{M_1}$ -re és  $x_{J_2}$ -re is helyes válaszokat adtak, egy másikba azok, akik  $x_{M_1}$ -re helyesen, de  $x_{J_2}$ -re helytelenül feleltek. További csoportokat kapunk  $x_{R_2}$  szerint. Mindegyik csoportra ismételhetjük az

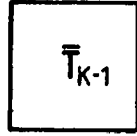


1. sz. ábra

eredeti eljárást egészen addig, míg a T varianciájából a csoportok átlagai közötti eltérések elegendően sokat nem magyaráznak meg, és ugyanakkor még a csoportok létszáma sem kicsi ahhoz, hogy statisztikai következtetéshez elegendő alapot nyújtson. Eredményül tehát az eredeti minta egy olyan hierarchikus felbontásához jutunk, melyben minden csoportot azzal jellemezhetünk, hogy a felbontás során tekintetbe vett kérdésekre helyes /jó/ vagy helytelen /rossz/ válaszokat kaptunk-e, valamint azzal, hogy a csoporton belül mekkora a T tesztteljesítmény átlagértéke és szórása. Az egyes csoportokhoz vezető kérdések általában eltérnek egymástól. Ezen eljárás eredményét szematikusan a 2.sz. ábrán mutatjuk be. A kialakult hierarchia egyik szélén azok csoportja helyezkedik el, akik minden kérdésre helyesen válaszoltak, a másik szélén pedig azok, akik minden kérdésre helytelenül. Attól függően, hogy ki melyik csoportba került /például az i-edikbe/, összteljesítményének vár-

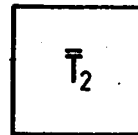


RR...RR



RR...RJ

•  
•  
•



JJ...JR



JJ...JJ

2.-sz. ábra

ható értékét a csoporthoz tartozó  $T_1$  átlaggal becsülhetjük. Azon túl, hogy az egyes csoportokhoz vezető kérdések nem azonosak, még azok száma sem feltétlenül egyező.

A fentiekben vázolt eljárás azonban két vonatkozásban tartalmaz egyszerűsítést. Az első az, hogy az elemek minősítése csak kétértékű lehet, a második pedig az, hogy csak egyetlen lépésre tekint előre. Az első úgy oldható fel, hogy egy-egy elemnél nem csupán az egyetlen felbontás magyarázó erejét vesszük figyelembe, hanem megvizsgáljuk az összes lehetséges értelmes felbontást. Ha például öt érték /1, 2, 3, 4, 5/ fordulhat elő egy elemnél, akkor ezen elem szerint elvileg  $\binom{5}{1} + \binom{5}{2} = 15$  különböző felbontás képzelhető el. Az értelmes felbontások száma azonban ennél általában kevesebb, mivel azok a felbontások, amelyeknél például az 1 és 5 tartoznak együvé a 2, 3, 4 pedig a másik csoportba, nem tartozhatók megalapozottnak. /Gondoljunk arra, hogy az elégteleneket és a jelesekét tekintenénk egymáshoz tartozónak./ Ilyenkor a megvizsgálandó felbontások azok, melyek ötfokozatú skálánk egy adott ponton való "elvágasásával" jönnek létre /például 1, 2 ; 3, 4, 5 /. Mindezeknek vizsgálata természetesen lényegesen több műveletet igényel.

A második említett probléma feloldása ugyancsak a végigvizsgálandó esetek számának növelését eredményezi. Amennyiben nem egy, hanem több kérdés alapján próbáljuk mintánkat egyszerre felbontani, a lehetséges változatok száma ijesztően megnő. A gyakorlatban kétlépéses előrelátásnál többet ritkán alkalmazunk, de az esetek többségében erre szükségünk sincs, mivel nem jutunk lényegesen jobb végeredményhez.

Az előzőkben vázolt eljárást valósítja meg az OSIRIS III programcsomag AID3 jelzésű programja /Automatic Interaction Detection/, melynek elméleti hátterét és alkalmazási lehetőségeit önálló kiadvány /Sonquist, Bakes, Morgan; 1973/ fejti ki részletesen.

#### Az elemzés eredményei

Az ismertetésre kerülő tesztek közül a tagadás teszt az egyszerűbb, mert abban minden elem kétértékű, annak megfelelően, hogy a tanuló helyesen válaszolt-e vagy nem. A helyes vá-

laszokat a továbbiakban 1-gyel, a helyteleneket 0-val jelöljük. A számítások eredményeképpen a tanulókat 16 csoportba sorolhatjuk. A csoportok homogének abban a tekintetben, hogy a hozzájuk tartozó tanulók ugyanazokra a kérdésekre ugyanolyan módon válaszoltak. Az első csoportba tartoznak például azok, akik sem a tizenegyedik, sem a nyolcadik, sem a harmadik teszteleme nem tudtak helyes feleletet adni. A negyedik csoportba például viszont azok, akik a tizennegyedik elem feladatát helyesen oldották meg, de a huszonhatodik, és a tizennegyedik kérdésre egyaránt rossz válaszokat adtak. Másrésztől a számítások eredményei azt mutatják, hogy az első csoport tanulójának átlagos teljesítménye a teljes tesztben mindössze 7,5, ezzel szemben a hatodik csoporthoz tartozóké 14. A tagadás tesztre vonatkozóan az elemzés eredményeit összefoglalóan az 1.sz. táblázatban és a 3.sz. ábrán adjuk meg.

A teszt elemeit rendre  $x_1^T, x_2^T, \dots, x_{28}^T$ -cal jelöljük, az összteljesítményt pedig  $T^T$ -vel. Az átlag jelzésére a megfelelő változó jele fölé huzott vonást használjuk. Ennek megfelelően a tesztre vonatkozó átlagérték a teljes mintában:

$$\overline{T^T} = 18,9$$

az  $i$ -dik csoportban kapott átlagot  $\overline{T}_i^T$ , a szórást pedig  $S_i^T$  jelöli.

A tagadás tesztben nyújtott teljesítmény alakulása  
az elemzés eredményeként kapott tanulócsoportokban

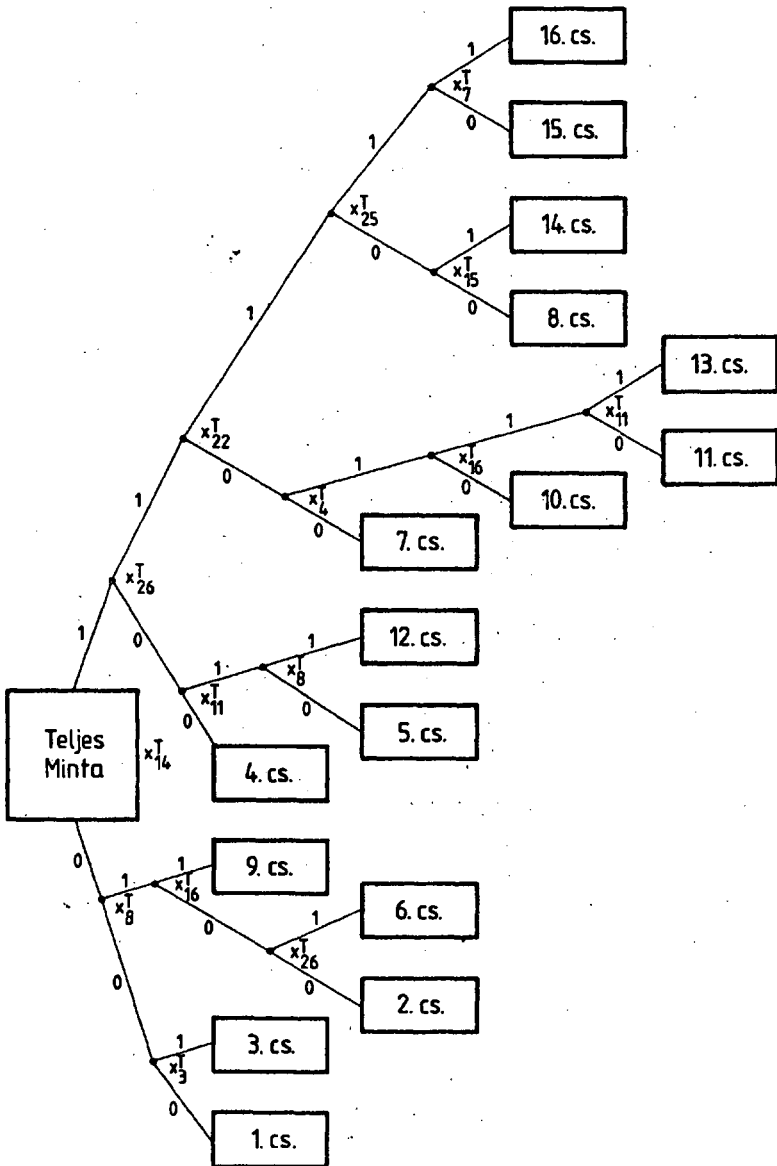
Csoport sor- száma $i$	A csoportot meghatározó tesztelemek	Átlag $\bar{T}_i$	Szórás $S_i^T$
1	$x_{14}^T/o/, x_8^T/o/, x_3^T/o/$	7,5	4,6
2	$x_{14}^T/o/, x_8^T/1/, x_{16}^T/o/, x_{26}^T/o/$	12,0	3,4
3	$x_{14}^T/o/, x_8^T/o/, x_3^T/1/$	12,0	2,6
4	$x_{14}^T/1/, x_{26}^T/o/, x_{11}^T/o/$	14,0	3,2
5	$x_{14}^T/1/, x_{26}^T/o/, x_{11}^T/1/, x_8^T/o/$	15,5	2,8
6	$x_{14}^T/o/, x_8^T/1/, x_{16}^T/o/, x_{26}^T/1/$	16,3	2,4
7	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/o/, x_4^T/o/$	16,6	2,2
8	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/1/, x_{21}^T/o/, x_{17}^T/o/$	16,8	3,1
9	$x_{14}^T/o/, x_8^T/1/, x_{16}^T/1/$	17,7	3,2
10	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/o/, x_4^T/1/, x_{16}^T/o/$	18,2	1,5
11	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/o/, x_4^T/1/, x_{16}^T/1/,x_{11}^T/o/$	18,4	1,8
12	$x_{14}^T/1/, x_{26}^T/o/, x_{11}^T/1/, x_8^T/1/$	19,2	2,6
13	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/o/, x_4^T/1/, x_{16}^T/1/,x_{11}^T/1/$	21,5	2,1
14	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/1/, x_{25}^T/o/, x_{11}^T/1/$	22,3	2,3
15	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/1/, x_{25}^T/1/, x_7^T/o/$	22,8	2,8
16	$x_{14}^T/1/, x_{26}^T/1/, x_{22}^T/1/, x_{25}^T/1/, x_7^T/1/$	26,2	1,9
Összes		18,9	5,1

Megmagyarázott variancia: 76 %

1.sz. táblázat



A minta hierarchikus felbontása a tagadás teszt kérdéseire adott válaszok alapján az összeteljesítmény függvényében.



3. sz. ábra

A 3.sz. ábrán bemutatott hierarchikus felbontás nem csupán predikciós eszköz, hanem egyuttal egy kérdezési stratégiát is szolgáltat, amelyet követve mindössze néhány elem megfelelő sorrendben való felhasználásával jól becsülhetjük a tanulók várható összteljesítményét. Eredményünk alapján először a tizenégyes elemre kell rákérdeznünk. Amennyiben az erre adott válasz helyes volt, a huszonhatos elemmel kell folytatnunk a kérdést, ha viszont nem, akkor a nyolcadikkal stb.

A megadott hierarchikus felbontás pontosan megmutatja, hogy az előzőleg kapott válasz helyességétől függően melyik elemmel kell folytatnunk a kérdést és azt is, hogy egy bizonyos pontot elérve, hogyan értékelhetjük a teljes tesztben várható összteljesítményt. Ábránkról leolvasható, hogy legalább három és legfőljebb hat kérdést kell feltennünk a választól függően.

Tekintsünk erre egy példát. Ha a tanuló a tizenégyes elemre jól válaszolt, a következő kérdésünk a huszonhatos elem lesz. Ha erre rossz feleletet kaptunk, akkor a tizenegyessel folytatjuk a tesztelést. Amennyiben erre a válasz helyes volt, a következő és egyben utolsó kérdésünk a nyolcas elem lesz. Ez az út - jól megválaszolva - a tizenkettes sorszámú csoport-hoz vezet el bennünket, melyben az 1.sz. táblázat megfelelő sorát figyelembe véve 19,2 pontos összteljesítmény várható. Ha megelégszünk egy becsléssel és nem törekszünk teljes bizonyosságra, megállhatunk a kérdések feltevésével és rögzíthetjük a kapott eredményt /19,2-t/, mint a valószínű összteljesítményt.

Megjegyezzük, hogy az összteljesítmény mérésében nem csupán az ilyen rövidített tesztelésből származhatnak hibák, hanem más forrásokból is, például valamilyen figyelemelterelő tényezőből, a tanuló pillanatnyi diszpozíciójából stb. adódóan. E módszer alkalmazási biztonságára egy mutatószám az, hogy a csoportok átlagteljesítményének különbségei mekkora hányadot magyaráznak meg a tanulói teljesítmények összvarianciájából. Ez az ismertetett példában /a tagadás tesztre/ 76 %-nak adódott. Ugyanez a lépésenkénti regressziós elemzésben /Hunyá-né, 1984/ az itt maximális hatos kérdésszám mellett 73 % volt. A most bemutatott módszer alkalmazásával a becslési pontosság akkor növelhető, ha a vizsgált minta elemszáma elegendően nagy, mert - mint erre a bevezető részben már utaltunk - a kapott

csoportokban az elemszám nem lehet kicsi. Ekkor ugyan a megmagyarázott varianciaszintje formálisan magasabb lenne, de eredményünk stabilitása, statisztikai megalapozottsága válna bizonytalanná.

Az eddig elmondottakból világos, hogy a teljesítmény-meghatározás itt leírt módszerét csak abban az esetben alkalmazhatjuk, ha folytonos kapcsolat áll fenn a tesztelő és a tesztelt között - szemben a szokásos írásos tesztelési eljárásokkal. Azokban ugyanis a feltett kérdések és azok sorrendje előre rögzített /bár e sorrendtől a válaszadó eltérhet/. Az ilyen írásos eljárás nem teszi lehetővé azt, ami itt alapvetően fontos, nevezetesen azt, hogy az újabb feladatelemet az őt megelőző kérdésre adott választól függően jelöljük ki. Ez a feltétel a hagyományos, személyes tanár-diák kapcsolatban nyilvánvalóan teljesül és a gyakorlott pedagógus az ilyen helyzethez alkalmazkodó kérdésfeltevési módot általában alkalmazza is. A kérdés az, hogy van-e lehetőség az ilyen adott stratégiát követő kérdezési eljárás valamilyen mérőeszköz felhasználásával történő automatizált megvalósítására.

Kézenfekvő eszközként kínálkozik a számítógép, amely minden további nélkül képes /többek között/ ilyen - szaknyelven fogalmazva - elágazó programok végrehajtására. Nem elegendő azonban, hogy a folyamat automatizált legyen, szükség van arra is, hogy a mérést rövid idő alatt tömegmérésekben lehessen elvégezni. Mindezen követelményeknek eleget tud tenni bármilyen párbeszédés számítógépes rendszer. A mozgathatótságot, költségeket figyelembe véve azonban az ilyen feladatkör ellátására a legalkalmasabbnak a személyi számítógépeket kell tekintenünk, amelyek ideális esetben nagyobb rendszerhez is kapcsolódhatnak. /Például feladatbankhoz való hozzáférés, statisztikai adatgyűjtés érdekében./ A számítógépes megvalósításnál természetesen a tesztelemekre adott válaszoknak, megoldásoknak géppel értékelhetőnek kell lenniük, ami enyhe korlátozást jelent a személyes kapcsolathoz viszonyítva, mivel tetszőleges verbális válaszok nem engedhetők meg. Ez azonban nem lényegbevágó megszorítás.

A másik példaként ismertetett teszttel kapcsolatban két újabb kérdést vetünk fel. Az előzőekben már feltettek a verbális rendszerezés teszténél is ugyanugy válaszolhatók meg, mint

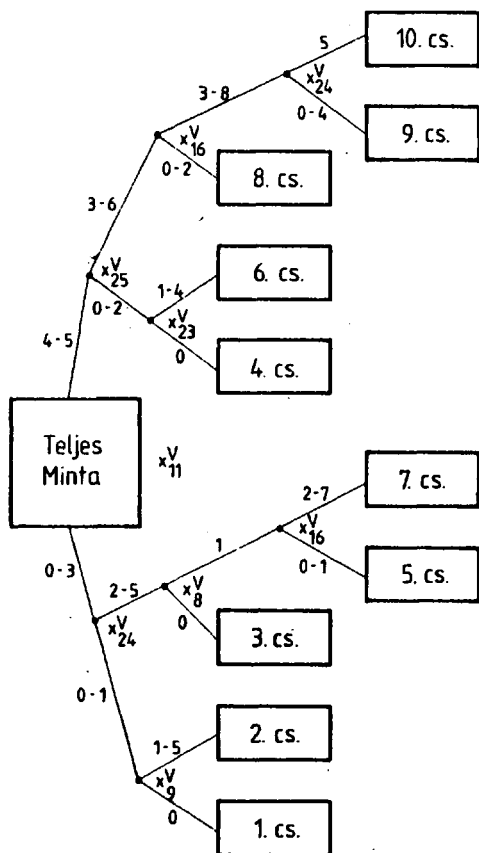
A verbális rendszerezés tesztben nyújtott teljesítmény alakulása az elemzés eredményeként kapott tanulócsoportokban

Csoport száma $i$	A csoportot meghatározó teszt- elemek	$\bar{T}_i^V$	$S_i^V$
1	$x_{11}^V/0-3/, x_{24}^V/0-1/, x_9^V /0/$	6,4	5,4
2	$x_{11}^V/0-3/, x_{24}^V/0-1/, x_9^V /1-5/$	17,0	6,9
3	$x_{11}^V/0-3/, x_{24}^V/2-5/, x_8^V /0/$	23,3	5,4
4	$x_{11}^V/4-5/, x_{25}^V/0-2/, x_{23}^V/0/$	26,3	6
5	$x_{11}^V/0-3/, x_{24}^V/2-5/, x_8^V/1/, x_{16}^V/0-1/$	31,6	6,8
6	$x_{11}^V/4-5/, x_{25}^V/0-2/, x_{23}^V/1-4/$	39,6	5,9
7	$x_{11}^V/0-3/, x_{24}^V/2-5/, x_8^V/1/, x_{16}^V/2-7/$	42,1	6,3
8	$x_{11}^V/4-5/, x_{25}^V/3-6/, x_{16}^V/0-2/$	44,3	4,6
9	$x_{11}^V/4-5/, x_{25}^V/3-6/, x_{16}^V/3-8/, x_{24}^V/0-4/$	52,4	4,0
10	$x_{11}^V/4-5/, x_{25}^V/3-6/, x_{16}^V/3-8/, x_{24}^V/5/$	59,2	2,4
	Összes	36,5	16,7

Megmagyarázott variancia: 90,4 %

2.sz. táblázat

A minta hierarchikus felbontása a verbális rendszerezés teszt kérdéseire adott válaszok alapján az összeteljesítmény függvényében.



4. sz. ábra

a tagadás teszténél. Az eredményeket a 2.sz. táblázat és a 4.sz. ábra mutatja be összefoglalóan.

Az egyik kérdés az, hogy hogyan kezelhetjük azokat a teszt-elemeket, melyeknél a válaszokat nem két kimenettel /jó, rossz/ értékeljük, hanem finomabb felbontású skálával. A másik kérdés az optimalizálttal ekvivalens változatok előállítása.

Az első kérdésre vonatkozóan /az eljárásban a csoportok részcsoporthoz való felbontásánál megvizsgálandó lehetséges változatokat illetően/ a bevezetőben már szóltunk. Mivel esetünkben az értékelés rangskálán jelentkezik, jogos az egyszerűbb keresés alkalmazása, amikor is a felbontás csak a skála valamely pontjánál való szétvágással jöhet létre. Ennek megfelelően a 4.sz. ábrán az elágazásoknál nemcsak 0 és 1 értékek jelölik az egyes utakat, hanem értékintervallumok, például a kiindulópontnál az alacsonyabb szinthez tartozók csoportjába csak azok kerülnek, akiknek válaszáat 0 és 3 közé eső pontszámmal értékeljük, a magasabb szintű csoportba pedig azok, akiket 4-gyel vagy 5-tel. Ugyanide kötődik egy másik gondolat is, amely az 5.sz. ábrán bemutatott ekvivalens teszténél élesen is felvetődik. Ha megvizsgáljuk a 10-es és a 12-es csoportokhoz vezető utat, azt találjuk, hogy azon a huszonötödik tesztelem kétszer is szerepel. Azt jelenti-e ez, hogy ugyanazt a kérdést többször is feltesszük! Ez esetleg zavarólag hathat a válaszadó tanulóra és nincs biztosítékunk arra, hogy ugyanazt a választ kapjuk, mint amit előzőleg adott. Egy ilyen esetre viszont modellünk nincs felkészülve, és ebből adódóan statisztikai következtetésünk sem lenne teljesen megalapozott. Ha figyelembe vesszük azt, hogy akár a tesztelő pedagógus, akár a számítógép megjegyezheti a korábbi válaszok értékét, akkor nyilvánvaló, hogy az ismételt kérdésfeltevésre nincs is szükség, hiszen az előzőleg kapott /és megőrzött/ érték alapján egyértelműen eldönthető, hogy melyik utat kell követnünk a kérdezőségi stratégiában. E gondolatmenettel összhangban a 10. és a 12. csoportokhoz vezető uton nem öt, hanem négy tesztelemre kell rákérdeznünk annak ellenére, hogy a hierarchikus diagramon öt lépést látunk, hiszen az utolsó lépés csupán azt jelenti, hogy az előzőleg már értékelt huszonötös elem pontszámát újra meg kell vizsgálnunk, de most már csak azt kell

vizsgálunk, de most már csak azt kell eldöntenünk, hogy arra 4-et vagy az 5-6 közül valamelyiket kapta a tanuló.

A második kérdésre, az ekvivalens teszt előállítására nem térünk ki részleteiben, mivel az AID3 programban megvalósított finomabb lehetőségek mélyebb és a pedagógus számára kevésbé érdekes ismertetését igényelné. Csupán utalunk arra, hogy az ekvivalens rövidített teszt előállításához az alternatív /de egy-egy lépésben nem optimálisnak tekintett/ felbontások magyarázóerejének összehasonlítására, az "előreteljesítés" lehetősége, és a hierarchiának bizonyos szintig való előzetes rögzítése, valamint a tesztelemekek közötti válogatás ad segítséget. Egy illy módon előállított ekvivalens változatot a 3.sz. táblázat és az 5.sz. ábra mutat be a verbális rendszerezési képesség tesztjére.

#### Záró megjegyzések

A tagadás tesztnél az eredeti tesztnek az optimalizálttal való lefedettségét illetően ugyanaz mondható el, mint a lépésenkénti regresszióval kapott első változatra /Hunyáné, 1984/, nevezetesen az, hogy a négy strukturális egységből itt is háromból választottunk ki elemeket. Ez a megállapítás a kérdéses struktúra egészét tekintve áll fenn. Egy-egy utnál lehetséges, hogy a lefedettség kevesebb strukturális egységre terjed ki. A verbális rendszerezés tesztnél mindkét változat mind a négy strukturális egységből tartalmaz elemeket.

A reliabilitás kérdését korábbi cikkünkben ugyancsak részletesen megvizsgáltuk, az ott elmondottak érvényesek jelen eredményeinkre is.

Összefoglalásul megállapíthatjuk, hogy az ismertetett eljárás alkalmazásával olyan rögzített tesztekhez és ezekre olyan kérdésfeltevési stratégiákhoz jutottunk, melyek elsősorban személyi számítógépen megvalósítva tömegméretekben teszik lehetővé az automatizált, gyors vizsgálatokat és gyors értékelést.

A verbális rendszerezés tesztben nyújtott teljesítmény alakulása az elemzés eredményeként kapott tanulócsoportokban /egyenértékű változat/

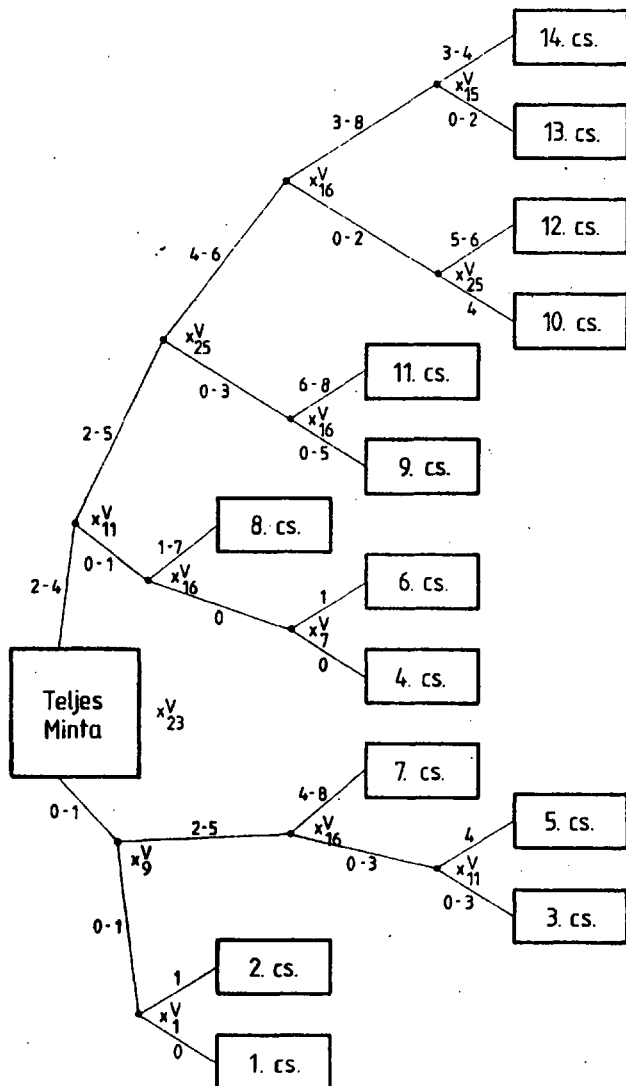
Csoport sorszáma $i$	A csoportot meghatározó teszt-elemek	$\bar{T}_i^V$	$S_i^V$
1	$x_{23}^V/0-1/, x_9^V/0-1/, x_1^V/0/$	2,8	4,4
2	$x_{23}^V/0-1/, x_9^V/0-1/, x_1^V/1/$	8,9	6,4
3	$x_{23}^V/0-1/, x_9^V/2-5/, x_{16}^V/0-3/, x_{11}^V/0-3/$	16,8	6,0
4	$x_{23}^V/2-4/, x_{11}^V/0-1/, x_{16}^V/1/, x_7^V/0/$	21,5	5,9
5	$x_{23}^V/0-1/, x_9^V/2-5/, x_{16}^V/0-3/, x_{11}^V/4/$	24,9	5,5
6	$x_{23}^V/2-4/, x_{11}^V/0-1/, x_{16}^V/0/, x_7^V/1/$	27,7	5,6
7	$x_{23}^V/0-1/, x_9^V/2-5/, x_{16}^V/4-8/$	94,4	6,9
8	$x_{23}^V/2-4/, x_{11}^V/0-1/, x_{16}^V/1-7/$	36,0	7,0
9	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/0-3/, x_{16}^V/0-5/$	38,2	4,0
10	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/4-6/, x_{16}^V/0-2/,$ $x_{25}^V/4/$	44,0	4,1
11	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/0-3/, x_{16}^V/6-8/$	45,8	4,9
12	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/4-6/, x_{16}^V/0-2/,$ $x_{25}^V/5-6/$	47,9	3,9
13	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/4-6/, x_{16}^V/3-8/,$ $x_{15}^V/0-2/$	51,3	2,9
14	$x_{23}^V/2-4/, x_{11}^V/2-5/, x_{25}^V/4-6/, x_{16}^V/3-8/,$ $x_{15}^V/3-4/$	55,8	3,5
Összesen		36,5	16,7

Megmagyarázott variancia: 90,4 %

3.sz. táblázat



A minta hierarchikus felbontása a verbális rendszerezés teszt kérdéseire adott válaszok alapján az összteljesítmény függvényében /egyenértékű megoldás/.



5. sz. ábrán

## I r o d a l o m

NAGY József: A műveleti képességek rendszere.

Acta Universitatis Szegediensis de A. József nominatae  
Sectio Paedagogica et Psychologica 25., Szeged, 1983.

HUNYA Péterné: Minimális elemszámu ekvivalens tesztek kiválasztása.

Acta Universitatis Szegediensis de A. József nominatae  
Sectio Paedagogica et Psychologica 26., Szeged, 1984.

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

SONQUIST J.A., E.L. BAKER, J.N. MORGAN: Searching for  
structure. ISR, The University of Michigan, Ann Arbor,  
Michigan, 1973.

Хуни Петернэ:

### Интерактивное тестирование

В своей работе автор исследует вопрос, каким образом можно составить из тестов большого размера (заполняющих весь лист), краткий тест с оптимально малым количеством элементов, применимый и в широких массовых измерениях, или же — вариант, эквивалентный тесту. В качестве цели намечается создание такого измерительного инструмента, предполагающего диалогическую связь между тестирующим и информатором, с помощью которого мы можем получить максимальную информацию относительно всех результатов, ожидающихся в полном тесте. Анализ в качестве математического способа использует метод получения данных структуры  $A_{12}$  3. Он устанавливает, что полученная таким образом структура, соответствующая вышеизложенным условиям, дает возможность иерархической классификации популяции. В последних группах этой структуры ожидаемые результаты значительно отличаются друг от друга. Сам краткий метод тестирования, основываясь на полученную в качестве результата анализа иерархическую систему, может быть реализован на личной вычислительной машине и применен в качестве практического метода измерения.

Hunya Péterné

### Interaktive Untersuchungen

Der Autor dieser Arbeit untersucht die Möglichkeiten, auf optimalen Wege aus grossen /ein ganzes Gebiet umfassenden/ Tests solche gekürzte, nur wenige Fragen umfassende Tests, welche so auch bei massenhaften Messungen Anwendung finden können, oder andere, hiermit gleichwertige Varianten zu gewinnen. Sein Ziel ist es, ein solches Messinstrument zu schaffen, dessen Anwendung, den Dialog zwischen Tester und Testobjekt vorausgesetzt, maximale Information bezüglich der im vollständigen Test zu erwartenden Gesamtleistung liefert. Als mathematisches Hilfsmittel wird in der Analyse die AID-3 Methode zur Darstellung der Datenstruktur verwendet. Der Autor konstatiert, dass die sich an die vorherigen Bedingungen anpassende Fragestruktur, welche so gewonnen werden kann, eine hierarchische Zerlegung der Population liefert, in deren Gruppen die zu erwartenden Leistungen erheblich voneinander abweichen. Das gekürzte Testverfahren an sich - welches auf der als Ergebnis der Analyse gewonnenen hierarchischen Struktur beruht - kann auf Kleinrechnern realisiert und in der Praxis als Messinstrument angewendet werden.