

POTE Élettani Intézete

Írásbeli tesztek jellemzése, standardizálása

Czopf János

Az írásbeli tesztelés mind gyakoribb elterjedése a módszer vitathatatlan előnyeit és vitatható hátrányait bizonyítja. Mind több közleményt olvashatunk az írásbeli tesztek kiértékelésének elvi, módszertani, sőt konkrét számítástechnikai problémáiról is. A tesztek kiértékelése utóbbiaknál inkább az egyes hallgatók válaszainak elbírálását, a hallgatók skálázását jelenti.

Miután az elmúlt esztendőkből intézetünkben is bevezetésre került ezen írásbeli módszer, fontos szerepet kapott a kérdésbank összeállításával a tesztkérdések kiértékelése is.

Az alább tárgyalt felmérésen 220 hallgató jelent meg, minden hallgató 120 azonos kérdést kapott. Az alkalmazott kérdéstípusok: kiegészítés, reláció analízis, mennyiségi összehasonlítás, többszörös kiegészítés, esetelemzés stb. voltak. Minden kérdéshez öt válaszalternatívát adtunk meg. Minden jól megválaszolt kérdés egy pontot jelentett. A véletlen ráhibázás felső határa (25 + 20) 45 pontnak adódott. Az évfolyam pontszám-eloszlása jól közelíti a normális eloszlást ( $N/\mu = 78,50$ ,  $\sigma = 11,87$ ), ezért alkalmasnak mutatkozott a tesztanalízis elvégzéséhez.

A kérdések jellemzésére - Osváth Károly psychés teljesítmények mérése során nyert tapasztalatai alapján - a következő statisztikai paramétereket választottuk: nehézségi index, diszkrimináló képesség, belső érvényesség és külső érvényesség.

A nehézségi index:  $p = f/N$  az egyes kérdésekre adott helyes válaszok relatív gyakorisága egyszerűen nyerhető, ugyanakkor igen hasznosnak mutatkozik az extrém nehéz, illetve könnyű kérdések szelektálására. A véletlen ráhibázással korrigált nehézségi index:  $p = (f_1 - (f_0/k-1))/N$  kisebb abszolút értékű, a változás irányában jól követi az előzőt. Az index mérési hibája átlagosan 2.5 %-nak mondható és maximális hiba esetén sem éri el a 3.5 %-ot. Ha azonos alternatíva-számmal dolgozunk és  $f \geq 5$ , akkor ez a hiba elhanyagolható.

A diszkrimináló képesség: arra ad információt, hogy a kérdés, mint mérő-eszköz, hogyan differenciál a legkevésbé, illetve a legjobban felkészült hallgatók közt, azaz a kérdés helyes megválaszolása milyen mértékben követeli meg az anyag mélyebb megértését. Jellemzőül az alsó és felső 25 % válaszait vizsgáltuk az összes helyes válasz függvényében.

$I_{UL} = (R_U - R_L) / f$  (1). A fenti index az első méréskor inkább csak előjelet ad, melynek segítségével egyértelműen kitűnik a rosszul, ill. ellentmondóan differenciáló kérdés. A megtartott kérdések esetén a diszkrimináló képesség szignifikáns jellegét  $\chi^2 = N (p_U - p_L)^2 / 4 p q$  értékkel igazolhatjuk.

Belső érvényesség: Bizonyos mértékben kapcsolatban van az előző mérőszámmal. Feltételezve, hogy tesztünk egyes kérdései jól mérik az egyének tárgyi ismeretét, azt várjuk el, hogy a több pontszámot elért hallgatók nagyobb valószínűséggel adjanak helyes választ a kevesebb pontszámot elértéknél. Tehát azt vizsgáljuk, hogy milyen összefüggés van a helyes válasz adás és az elért összpontszám között. Ha feltételezzük, hogy az egyes kérdésekre helyes választ adók összpontszáma közel normális eloszlású, akkor dolgozhatunk  $r_{pb} = (M_p - M_f) \sqrt{p/q} / \sigma_f$  értékekkel. (2)

Külső érvényesség: megállapítása a hallgatók teszt által képzett rangsora és a csoportvezetők évi tapasztalata, mint külső kritérium közti összehasonlítás alapján történt. Az évfolyam kb. 3.5 %-ában mutatkozott lényeges eltérés, ami első mérést figyelembevéve jónak mondható és a módszer alkalmasságát igazolja.

Két kérdést kiválasztva, a következő táblázat kívánja demonstrálni az egyes alternatívákra adott gyakoriságok, illetve a kérdéseket jellemző statisztikai paraméterek értékeit. (1 a helyes alternatívákra kapott válaszok számát jelöli a teljes évfolyamon a felső, ill. alsó 25 %-ban)

. . x / kérdés:

	1	2	3	4	5	SUM.
N	42	85	75	10	6	218
N <sub>f</sub> /25 %)	10	22	21	2	0	55
N <sub>a</sub> (25 %)	11	18	20	4	2	55

. xx / kérdés:

N	8	0	10	192	8	218
N <sub>f</sub> (25 %)	1	0	1	52	1	55
N <sub>a</sub> (25 %)	2	0	3	43	7	55

	$\frac{..x/}{\text{kérdés}}$	$\frac{.xx/}{\text{kérdés}}$
Nehézségi index	34,40	88,07
N.l.hibája	0,032	0,022
N.l. korrigált	0,18	0,85
Diszkr. kép.	0,013	0,047
D.l. szign.	0,080	13,889
Belső érvényesség	0,02	0,88

A módszer lehetőséget biztosít arra, hogy a kérdésbankban csak a jó statisztikai indexekkel rendelkező kérdések szerepeljenek, melyek mint több index tulajdonságú változók, egyszerű uton felvehetők. A következő felmérések lehetőséget adnak a megbízhatóság ellenőrzéséhez re-test módszerrel, majd ezt követően válik lehetővé az érvényesség felülvizsgálata, illetve a hitelesítés.

Miután a felmérés azok közé a ritka visszacsatolások közé tartozik, melyekből hasznos tapasztalatokat nyerhetünk az információ-közlés és a befogadás folyamatára, részletesen elemezhetjük az egyes kérdés-alternatívákra adott válaszok gyakoriságából a leadott anyag logikai strukturáját, melyet mindjobban megismerve vissz szabályozhatunk az oktatás folyamatára.

#### I r o d a l o m

1. Guilford, J.P.: Fundamental Statistics in Psychology and Education. Mc.Graw-Hill Book Company, Inc. New York, 1956.
2. Gulliksen, H.: Theory of Mental Tests. Chapman - Hall Ltd. London, 1950.
3. Lienert, G.A.: Testaufbau und Testanalyse Verlag Julius Beltz, Weinheim, 1961.
4. Ozsváth K.: Csoportos pszichológiai vizsgálatok. Tanulmányok a kat. nev. köréből. Zrinyi. 1966.
5. Czopf J.: Programozott okt. Oktatógépek. Pályamunka 1965.
6. Tigyi A.: A Multiple Choice vizsgamódszer. FPK. 1971. Budapest.