

## A VALÓSZÍNŰSÉGI TESZTELMÉLET ADTA ESZKÖZÖK AZ IDEÁLIS MÉRÉS MEGVALÓSÍTÁSÁHOZ

Molnár Gyöngyvér

*Szegedi Tudományegyetem, BTK, Neveléstudományi Tanszék  
MTA-SZTE Képességkutató Csoport; SZTE Oktatásméleti Kutatócsoport*

*Kulcsszavak:* objektív mérés, IRT, Rasch-modell, parciális kredit modell

A tesztemléletek újabb, a nemzetközi mérésekben (OECD PISA) is egyre inkább alkalmazott generációját adják az objektív mérést is lehetővé tevő valószínűségi tesztemléletek (Bond és Fox, 2001; Horváth, 1997). Az objektív mérés megvalósításának lehetőségével, módszereivel régóta foglalkoznak a társadalomtudósok, mivel azáltal lehetőség nyílna egyetemes, mindenki által elfogadott skálák megalkotására (Write és Stone, 1979; Griffin, 1999). Hazánkban jelentős múlttal rendelkeznek a klasszikus tesztemléleti módszerekkel történő elemzések, azonban ezek nem alkalmasak az objektív mérés, az objektív skálák megalkotására, továbbá módszerei segítségével bizonyos kérdéseket nem tudunk megválaszolni.

Az előadás célja a klasszikus tesztemlélethez képest a valószínűségi tesztemlélet (IRT) adta lehetőségek, különböző valószínűségi függvényeken nyugvó elemzések bemutatása. A bemutatott elemzések alapját szimulált adatbázisok (N=2000) képezik. Az elemzésekhez a ConQuest szoftvert használjuk.

Az előadás három, különböző típusú adatok elemzésére alkalmas mérési modellt, módszert mutat be. Mindhárom modell a valószínűségi modellek családjába tartozik, illetve lehetőséget biztosít a mintafüggetlen itemkalibrációra és a személy képességparaméterének tesztfüggetlen meghatározására.

A dichotóm adatok esetében használt mérési modell a Rasch-modell (Rasch, 1960), a pedagógiai-pszichológiai teszteken nyújtott teljesítmények elemzésében leggyakrabban használt IRT-modell. Ez a modell dichotómitásánál fogva a részben jó válaszok elemzésére nem ad lehetőséget. A bemutatásra kerülő másik két modell Andrich (1979) rangskálás modellje és Masters (1982) parciális kredit modellje. Előbbi modell hátránya, hogy csak azon adatbázisok esetén alkalmazható, ahol minden egyes itemnek megegyezik a skálaszerkezete. Ezzel szemben a parciális kredit modellben minden egyes itemnek akár teljesen különböző skálaszerkezete is lehet. A parciális kredit modell alkalmazható például Likert-skálán lévő adatok elemzése során, vagy olyan itemeknél, ahol a válaszok egy része jobb, mint a többi (pl.: tévképzés-kutatásokban), vagy olyan többlépcsős itemek esetében (pl. problémamegoldásnál), ahol a diáknak több, egymástól lehetőleg független lépést kell megoldania a feladat elvégzése során. Az előadás keretében kitérünk e modellek megkülönböztető tulajdonságaira is.

A valószínűségi tesztemlélet eszközt ad a kutatók kezébe az ideális mérés megvalósításához, ahol elvárt, hogy a teljesítmény függvényében meg tudjuk mondani, hogy mit tud az adott diák, a vizsgált képesség fejlődésének milyen stádiumában van, mi várható el tőle. A Rasch-modell és továbbfejlesztett változatai egyetemes skálák létrehozásához adnak matematikai háttérrel, ahol össze tudjuk kapcsolni a tanulók képességszintjének és az itemek nehézségi szintjének két skáláját. Az OECD PISA 2003-as mérés kapcsán nemzetközi szinten már elindult egy ilyen irányú skálázásra vonatkozó kezdeményezés.