

ZERÉNYI Károly

## A Likert-skála adta lehetőségek és korlátok

### Bevezetés

A Likert-skálás kérdéseket is tartalmazó empirikus kutatásoknál nagy hangsúly helyeződik arra, hogy a kérdésekre adott válaszok feldolgozását, elemzését méréselméleti, illetve tesztelméleti szempontból alátámasztott módon végezzük el. Ugyanis egyáltalán nem mindegy, hogy a kutatás eredményeként rendelkezésre álló adatokat a mérési skálák szempontjából miként kezeljük és ennek megfelelően milyen matematikai-statisztikai műveleteket, illetve vizsgálati módszereket alkalmazunk. A Likert-skálák esetében szakirodalmi kutatásokra támaszkodva általánosan elmondható, hogy a kutatók a skálaértékeket többek között átlagolják, így élnek azzal a feltételezéssel, hogy intervallumváltozóról van szó, hiszen ebben az esetben van értelme az átlagszámításnak (*Brown, 2011; Kehl, 2012*).

Az elkövetkezőkben a fentieket alapul véve rá kívánok világítani a Likert-skálák alkalmazásának méréselméleti, valamint tesztelméleti kontextusba helyezett problematikájára, amelyre egy lehetséges megoldást a modern tesztelmélethez kötődő modellek és azokat kezelni tudó szoftverek adhatnak.

### A Likert-skála eredete és jellemzői

A Likert-skála *Rensis Likert* amerikai pszichológus és szociológus nevéhez fűződik, aki az 1932-es doktori értekezéséhez kapcsolódóan fejlesztette ki az attitűdök mérésére szolgáló módszerét (*Horváth, 2004; Bertram, 2014*), amelyet napjainkban is előszeretettel használnak kérdőíves kutatásokban. A módszer lényege, hogy különböző állításokat két szélsőséges végpont között kialakított skálán értékelnek, amely skálát általában 1-5-ig vagy 1-7-ig terjedő pontszámokkal látják el. Mind a két esetben az egyik végpont abszolút ellenkezést, míg a másik abszolút egyetértést, azonosulást testesít meg, amelyek között a válaszadó elhelyezheti véleményét az adott állítással kapcsolatban. Emellett a páratlan fokozat miatt a köztes válaszok között semleges álláspontot is ki lehet alakítani.

Abban az esetben, „ha a válaszlehetőségek száma páros (például 2, 4, 6), akkor ún. kényszerválasztásról beszélünk, hiszen a válaszlehetőségeket tekintve nincs középút” (*Rózsa, Nagybányai és Oláh, 2006, 72.*). Számos kutatásban foglalkoztak éppen a skálafokozatok számával, valamint a páros vagy páratlan válaszlehetőségek alkalmazásával, azok előnyeivel és hátrányaival. Azzal kapcsolatban is születtek vizsgálatok, hogy számít-e a válaszlehetőségek sorrendje, tehát balról jobbra vagy jobbról balra emelkednek az értékek. A vizsgálatok eredménye szerint a válaszadók előnyben részesítik a bal oldali válaszlehetőségeket, ezért ezzel a megoldással némi befolyást is lehet gyakorolni.

### A pszichológiai, pedagógiai mérések problémái és a reprezentációs méréselmélet

Az első méréselméleti megközelítés, figyelemmel a természettudományi gyökerekre, csak az additív mennyiségek (például tömeg, hosszúság, idő) mérésére korlátozódott. Tekintettel arra, hogy a pszichológiában és a pedagógiában lényegében nincsenek additív mennyiségek, ezért a klasszikus méréselmélet nem használható. Az ebben a kérdéskörben kialakult vitában egy

jelentős állomás a brit tudományos bizottság 1933-ban megjelent jelentése, amely megerősítette ezt az álláspontot, miszerint tudományosan megalapozott pszichológiai, illetve pedagógiai mérések nincsenek.

Az áttörést *Stanley Smith Stevens* hozta el, aki 1946-ban megalkotta a skálákkal kapcsolatos elméletét, amely a reprezentációs méréselmélet egyik igazodási pontja. *Stevens* javaslata szerint ki kell terjeszteni a mérés fogalmát, ugyanis a mérésnek számos formája létezik, következésképpen különböző mérési skálák hozhatók létre (*Kehl, 2012*). Az első és talán a legegyszerűbb a nominális skála, amelynek elemei között semmilyen matematikai kapcsolat nincs. A nominális adatok, mint például a gyerekek neme, a tantárgyak között nincsen hierarchia, rendezés, és legkevésbé különbségek vagy arányok (*Nahalka, 1996*). Ha az adatok között hierarchia van és rendezettek, akkor ordinális adatokról beszélünk, amelyek ordinális skálán helyezhetőek el (*Brown, 2011*). Ilyen adatnak minősül többek között az ásványok keménysége, egy futóversenyen a sorrend, amelyek esetében kisebb-nagyobb viszonyokról lehet beszélni, azonban különbségekről még nem. Amikor az adatok közötti különbség is értelmezhető, akkor már intervallumskáláról van szó. Ebben az esetben egy négyváltozós reláció áll fenn, vagyis két adat között nagyobb a „távolság”, mint másik két adat között. Tipikusan ilyenek a hőmérsékleti skálák közül a Celsius- és a Fahrenheit-skála, amelyek megfelelnek ennek a kritériumnak. Ugyanakkor az ezeken a skálákon elhelyezkedő számok arányai nem hordoznak információt. Amikor ez a feltétel is teljesül, akkor jutunk el az arányskálához. Erre jó példa a Kelvin-skála, vagy a hosszúság mérése, amelyek esetén van egy kiindulópont. Az első két skálához (nominális, ordinális) tartozó változókra a kategorikus vagy minőségi, kvalitatív, míg az intervallum- és az arányskálán mért változókra a mennyiségi, kvantitatív kifejezést szokták használni (*Nahalka, 1996*).

### **Az egyes mérési skálák esetén alkalmazható matematikai-statisztikai műveletek köre és a Likert-skála problematikája**

*Stevens* skálákkal kapcsolatos elméletének egyik legnagyobb visszhangját kiváltó eleme az a megállapítás, miszerint a mérési skálától függ, hogy az adott adatok esetén milyen statisztikai módszereket, eljárásokat lehet alkalmazni (*Kehl, 2012*). *Stevens* szerint a mérés számoknak különböző objektumokhoz vagy eseményekhez meghatározott szabályok szerint történő hozzárendelése. A megfigyelt egységek tulajdonságainál vizsgálati szempontként az egyezőség, a sorrendiség, a különbségek, valamint az arányok jelennek meg. Ennek megfelelően az egyes mérési skálák tulajdonságai meghatározzák az elvégezhető statisztikai műveleteket, amelyek az 1. táblázatban láthatóak.

1. táblázat: *Mérési skálák és tulajdonságaik*

Skála	Alapvető művelet	Matematikai csoport tulajdonság	Megengedhető műveletek
Nominális	Egyenlőség meghatározása	Permutációs csoport $x \sim f(x)$ , ahol $f$ tetszőleges, kölcsönösen egyértelmű hozzárendelés	Esetek száma Módusz
Ordinális	Sorrendiség meghatározása	Isotonikus csoport $x \sim f(x)$ , ahol $f$ tetszőleges, monoton növekvő függvény	Medián Percentilisek
Intervallum	Intervallumok/különbségek egyezőségek vizsgálata	Általános lineáris csoport $x \sim ax + b, a \neq 0$	Számtani átlag Szórás Rangkorreláció Szorzat momentum korreláció
Arány	Hányadosok egyezőségének vizsgálata	Hasonlósági csoport $x \sim ax, a \neq 0$	Mértani átlag Harmonikus átlag Relatív szórás

Forrás: Kehl Dániel (2012): Mintaelemszám tervezés Likert-skálás lekérdezések esetén klasszikus és bayesi keretek között; Doktori értekezés; PTE-KTK [6. p.]

Az 1. táblázatban szereplő megengedhető statisztikai műveletek kumulatíván értelmezhetők a skálák között oly módon, hogy az alacsonyabb rendű skálák műveleti elvégezhetőek a magasabb rendűeken is.

Az 1. táblázat utolsó oszlopában látható, hogy a kvantitatív kutatásoknál általában számított számtani átlag és szórás legalább intervallumskálás adatok megléte esetén értelmezhető, illetve megengedhető. Ennek megfelelően, ha legfeljebb csak ordinális adatok állnak rendelkezésre, akkor csupán az esetek számával, módusszal, mediánnal, valamint percentilisekkel lehetne dolgozni. Ugyanakkor a leíró statisztika mellett nem szabad megfeledkezni a különféle vizsgálati módszerekről, amelyeket nominális, illetve ordinális adatok esetén lehet alkalmazni.

**2. táblázat: Mérési skálák és vizsgálati módszerek**

	Nominális	Ordinális	Intervallum
<b>Átlag egyezése egy feltételezett értékkel</b>			Egymintás u-próba, egymintás t-próba
<b>Eloszlás egyezése egy feltételezett eloszlással</b>	Illeszkedésvizsgálatok, $\chi^2$ -próba	Illeszkedésvizsgálatok, $\chi^2$ -próba	Illeszkedésvizsgálatok, $\chi^2$ -próba
<b>Két minta összehasonlítása</b>	Homogenitás-vizsgálat, $\chi^2$ -próba	Mann-Whitney próba, összetartozó minták esetén: Wilcoxon-próba	Szórások egyezése: F-próba Várható értékek egyezése: kétmintás u-próba, kétmintás t-próba, Welch-próba
<b>Több minta összehasonlítása</b>		Kruskall-Wallis-féle eljárás	Szórások egyezése: Bartlett-próba Várható értékek egyezése: variancia-analízis
<b>Két változó összefüggése</b>	Keresztábra elemzés, $\chi^2$ -próba	Rangkorrelációs elemzés, Spearman-féle rangkorrelációs együtttható	Korrelációs számítás. regresszió számítás
<b>Kettőnél több változó kapcsolata</b>	Keresztábra elemzés, $\chi^2$ -próba		Parciális korrelációs együtttható, többváltozós regresszió, faktoranalízis, klaszterelemzés
<b>Különböző típusú változók kapcsolata</b>	Variancia-analízis, kovariancia-analízis		

Forrás: Nahalka István (1996): A statisztikai módszerek pedagógiai alkalmazásának indokai, statisztikai alapfogalmak. in: Falus Iván (szerk.) Bevezetés a pedagógiai kutatás módszereibe. Keraban Kiadó, Budapest [353-354. p.]

A vizsgálati módszerek tekintetében a 2. táblázatot alapul véve az intervallumskálás adatoknál lényegesen több lehetőség áll rendelkezésre, mint nominális vagy ordinális skála esetén. Ugyanakkor nominális vagy ordinális adatok esetén is lehet mintákat összehasonlítani, összefüggés-vizsgálatot végezni.

A pedagógiai kutatásokban is gyakran alkalmazott Likert-skálás kérdések útján kapott adatok felhasználhatóságát tekintve fontos tisztázni, hogy azok ordinális vagy intervallumskála erősségűek-e, hiszen ettől függ számos matematikai-statisztikai művelet elvégezhetősége,

illetve vizsgálati módszer alkalmazhatósága. Egy klasszikus Likert-skála esetén, ahol az egyes állításokat 1-5-ig lehet értékelni, így például az 1-es az egyáltalán nem ért egyet, az 5-ös pedig a teljes mértékben egyetért viszonyulást jelenti. Ennek megfelelően a válaszok sorrendbe tehetőek, azonban a köztük lévő „távolság” nem magyarázható, ezért az intervallumskála erősség nem teljesül. Ha nem elégedünk meg az ordinális skálán mért változók adta elemzési lehetőségekkel, akkor a klasszikus tesztelmélet keretein túl kell lépni, és a lehetséges megoldást a modern tesztelmülethez kötődően kell keresni.

### **A modern tesztelmélet és a rangskálás modell**

A pszichometriai vizsgálatok elméleti háttérének kidolgozásához kötődően létrejövő klasszikus tesztelmélet matematikailag teljes körűen alátámasztott formáját az 1968-ban Lord és Novick által megírt könyv foglalta keretbe, amelyben már a modern tesztelmélet alapjai is megjelentek. A klasszikus tesztelmélet alapvetően uralkodó elméleti keret napjainkban is a pszichológiában és a pedagógiában alkalmazott tesztek esetében, amely elméleti módszerekkel történő elemzéseknek hazánkban is jelentős múltja van, azonban ezek nem alkalmasak az objektív skálák megalkotására (Molnár, 2008).

A klasszikus tesztelmélet alapján például egy képesség fejlettségének mérése esetén a különböző tesztek eredményeinek pozitív lineáris kapcsolatban kell lenniük egymással ahhoz, hogy a mérés során intervallumskálás változókat kaphassunk. Ez a feltétel szükséges, de nem elégséges, hiszen emellett teljesülnie kell az intervallumskálánál meglévő relációnak is. A klasszikus tesztelmélettel szemben felmerült kritikákra, mint a populációfüggőség vagy a skálafüggőség problémájára nyújt egy lehetséges megoldást a modern tesztelmélet, amely már teljesíti a méréselméleti igényeket. Ennek megfelelően elérhetővé válik a legalább intervallumskálák kialakítása, a klasszikus tesztelmélettel szemben a mérés tesztől, illetve a mérés eredményeinek a mérés mintájától való függetlenítése. „A modern tesztelmélet nem a klasszikus tesztelmélet egy továbbfejlesztett, vagy „jobb” változata, hanem alapvetően más matematikai eszközökre támaszkodó, statisztikai eljárásokat használó, modelleket felállító és függvényekkel dolgozó tesztelmélet” (Molnár, 2003, 423.). Ezen elméletre gyakran használják az Item Response Theory (IRT), valamint esetenként a valószínűségi tesztelmélet megnevezést, amely szerint a tesztek feladatainak megoldása valószínűségi jellemzőkkel leírható folyamat.

A modern tesztelmélet egyik legjelentősebb modellje a dán matematikus *Georg Rasch* nevéhez fűződő Rasch-modell. A modell lényege, hogy egy adott képesség fejlettségének vizsgálata esetén valamennyi feladathoz hozzárendelhető egy feladatnehézségi mérték, továbbá valamennyi személyhez egy képességfejlettség, amelyek együttesen meghatározzák, hogy egy adott személy az egyes feladatokat milyen valószínűséggel tudja megoldani. Ugyanakkor a Rasch-modell csak dichotóm adatok esetén alkalmazható, amikor két válaszlehetőség közül választunk, vagyis például egy adott személy a feladatot meg tudta oldani vagy sem. A probléma megoldására továbbfejlesztették a Rasch-modellt, hogy más, nem dichotóm adatokból álló adatbázisok, mint a Likert-skálás kérdésekre adott válaszok elemzését is lehetővé tegyék (Molnár, 2008). Az attitűdteszteknél gyakran előforduló Likert-skálás adatok elemzésére alkalmas *David Andrich* ausztrál kutató rangskálás modellje, valamint *Masters* parciális kredit modellje (Schulz és Sibberns, 2004). A rangskálás modell kezelni tudja a páros vagy páratlan számú Likert-skálák alapján ragsorolt válaszalternatívákat azzal a feltétellel, hogy az összes állításra adott válasz azonos számú lépésből álljon, valamint

a válasz meghozatalakor megtett lépések nehézsége közel azonos legyen minden állítás esetén (*Molnár, 2008*). Abban az esetben, ha négy alternatíva közül kell választani, így például az egyáltalán nem értek egyet, nem értek egyet, egyetértek, teljesen egyetértek alternatívák esetén három azonos nehézségű lépés megtételéről beszélünk. Tehát az egyes válaszalternatívák számánál egyel kevesebb lépést kell megtenni. Amennyiben nem teljesül az azonos skálaszerkezet, illetve az egyes lépések nehézsége változó, akkor a rangskálás modell helyett a parciális kredit modell alkalmazható.

Napjainkban a számítógépes lehetőségek kiszélesedésével, és ehhez kapcsolódóan az egyre könnyebben hozzáférhető programoknak köszönhetően a modern tesztelmélet egyes modelljeinek, mint a Rasch-modell vagy a Likert-skála szempontjából releváns Andrich-féle rangskálás modell pedagógiai kutatásokban való hasznosítására adódik lehetőség. Tekintettel arra, hogy az empirikus kutatások során gyakran használt SPSS program nem tudja kezelni az IRT modelleket, ezért az elkövetkezőkben a rangskálás és a parciális kredit modellel történő adatfeldolgozást lehetővé tevő szoftverek közül kívánok néhányat bemutatni.

### **A Likert-skálás adatok elemzésére használható szoftverek**

Az elmúlt tíz évben az informatika rohamos fejlődésének köszönhetően a modern tesztelmélettel összefüggésben számos programot fejlesztettek ki, amelyek használhatók a Likert-skálás kérdések eredményeinek rangskálás modellel, illetve parciális kredit modellel történő feldolgozására. A modern tesztelmélettel foglalkozó tanulmányokban olvasható elemzések elkészítéséhez többnyire a ConQuest szoftvert, valamint a Winstep programot használják. Ezek mellett még számos megfizethető és ingyenes program is (Facets, WINMIRA, OPLM) a világhálón keresztül hozzáférhető. Az említett programok közül az első kettőről (ConQuest, Winstep) fogok bővebb leírást adni.

A ConQuest program különféle verziói egyaránt alkalmazhatók dichotóm és nem dichotóm adatok elemzésére. Ennek megfelelően a szoftver használatával elvégezhető a rangskálás, illetve a parciális kredit modellhez kapcsolódó számítások (*Wu, Adams és Haldane, 2007*). A program közvetlenül kezelni tudja az SPSS fájlokat és az eredményeket SPSS vagy Excel formátumban is elérhetővé teszi. Emellett a vizsgálatok eredményeinek megtekintésére a program széleskörű grafikus ábrázolási lehetőségeket biztosít, köztük valószínűség-térképek megjelenítését kínálja.

A fenti programhoz hasonlóan a *John M. Linacre* által kifejlesztett Winstep szoftverek is abszolút módon alkalmasak a nem dichotóm adatok elemzésére, így a szoftverrel a rangskálás adatok feldolgozása teljes körűen elvégezhető (*Linacre, 2006*). A Winstep konstrukciókat ebből kifolyólag gyakran hívják segítségül az oktatási tesztek, a pszichológiai becslések, valamint az attitűd tesztek vizsgálatához. A program hasonlóan a ConQuest szoftverhez kezelni tudja az SPSS és az Excel fájlokat, továbbá az eredményeket is mindkét formátumban képes rendelkezésre bocsátani.

### **A Likert-skálás adatok vizsgálata a modern tesztelmélet keretei között**

A Likert-skálás adatok modern tesztelméleti keretek közötti vizsgálatára alkalmas a Winstep ingyenes elérhető Ministep verziója, azonban a szoftver legfeljebb 25 állítást és 75 elemű mintát tud kezelni. A korlátok figyelembevételével mellett egy kis mintás kutatás esetén vagy gyakorlásképpen egy szűkített mintán ezzel a szoftverrel is elvégezhető a Likert-skálás adatok

vizsgálata. Első lépésként a meglévő adatokat fel kell tölteni a Ministep által elvárt struktúrába, amely az egyes itemek (attitűdállítások) megadását és a Likert-skálás pontszámok rögzítését jelenti. Emellett lehetőség van a válaszadók azonosítására szolgáló adatok (sorszám és/vagy szöveg) megadására is. Ezután már meg tudjuk vizsgálni az adatrendszer megfelelőségét a parciális kredit modellel történő elemzésre, vagyis hogy valamennyi állítás és személy esetében teljesülnek-e a modell alkalmazásának feltételei.

Az adatrendszer megfelelőségét az állítások tekintetében az „Item: measure” lekérdezés, a személyek esetében pedig a „Person: measure” vizsgálat eredményei között megtalálható infit/outfit értékek határozzák meg. Ha az infit/outfit táblában található értékek közül valamelyik 1,5-nél nagyobb, akkor azzal az állítással vagy az adott személy válaszával baj van. Ebben az esetben dönthetünk úgy, hogy nem változtatunk az adatrendszeren és nem használjuk a parciális kredit modellt, vagy a modell alkalmazása érdekében elkezdjük kivenni a problémás elemeket az adatrendszerből. Az utóbbi választásakor addig kell a problémás itemeket és személyeket szelektálni az adatbázisból, ameddig valamennyi esetben elfogadható infit/outfit értéket nem kapunk. Ebben az esetben egy olyan adatrendszert sikerült létrehozni, amely már megfelel a parciális kredit modell szabta feltételeknek és egyúttal intervallumskálán elhelyezkedő adatokat eredményez. Ha az eredeti ordinális skálán elhelyezkedő Likert-skálás adataink helyett intervallumskálán értelmezhető adatok állnak rendelkezésre, akkor már beszélhetünk átlagról, szórásról és megbízhatóbb matematikai-statisztikai vizsgálatok esetleges alkalmazásáról. Ennek érdekében egyes nagy nemzetközi méréseknél, mint a tanulók állampolgári nevelésével kapcsolatos attitűdjeit érintő IEA felmérésben (*Schulz és Sibberns, 2004*), vagy hazai vonatkozásban a középiskolások olvasás iránti attitűdjeinek vizsgálatánál is (*Kontra, 2008*), az egyes kérdésekhez kapcsolódó itemek közül az alkalmazott IRT modellel nem megfelelőket kivették.

Mindazonáltal beleütközhetünk abba a problémába, hogy egy többdimenziós jelenséget vizsgáltunk, amely a pedagógia területén gyakran előfordul. Ezzel párhuzamosan számolnunk kell azzal is, hogy az itemek és a személyek szelektálásának következményeként olyan csökkentett mintát kapunk eredményül, amely a megbízhatóság és reprezentativitás szempontjából nem lesz megfelelő. Következésképpen a Likert-skálás adatok intervallumskálán történő kezelése helyett nem marad más választás, mint az ordinális skála szabta keretek között maradni.

### **Összegzés, záró gondolatok**

A klasszikus tesztelmélet keretei között a Likert-skálás kérdések esetében a skálaértékek átlagolására nincs mód, hiszen azok nem intervallumskálán helyezkednek el, így legfeljebb az ordinális skála alapján megengedhető statisztikai műveletek elvégzésére és vizsgálati módszerek alkalmazására nyílik lehetőség. Miután az ordinális skálán elhelyezkedő adatok esetén kevesebb és kevésbé megbízható matematikai-statisztikai vizsgálatok állnak rendelkezésre, ezért a fennálló problémára egy lehetséges megoldást nyújthat a modern tesztelméleti modellek alkalmazása. A modern tesztelmélet talán legismertebb modellje a Rasch-modell, azonban ez csak dichotóm adatok elemzésére használható. A nem dichotóm adatok, mint a Likert-skálás kérdések eredményeinek vizsgálatára az Andrich-féle rangskálás modell, illetve *Masters* parciális kredit modellje alkalmazható. A Likert-skálás adatok elemzésére képes parciális kredit modell segítségével olyan intervallumskálán elhelyezkedő attitűderősségek határozhatók meg, amelyekkel már elvégezhetők a megbízhatóbb

matematikai-statisztikai vizsgálatok is. Ugyanakkor az IRT modellek alkalmazásának van egy komoly akadálya is, hogy azok komplex jelenségek (például problémamegoldás, kreativitás) esetén nem működnek.

A fent említett modellek használatára az elmúlt években egyre több szoftvert fejlesztettek ki, amelyek segítségével könnyen feldolgozhatók a különböző attitűd tesztek. A Likert-skálás kérdésekből származó adatok kezelésére többek között a Winstep, valamint a ConQuest programok alkalmasak, amelyek segítségével a rangsorolt adatok elemzése elvégezhető.

Úgy gondolom, hogy a hazai tudományos kutatásokban rendkívül hasznos és újító lehet, ha a Likert-skálás kérdések feldolgozásának nem csak a hagyományos útját választjuk, hanem a modern tesztelmélet adta lehetőségeket is kihasználjuk. Ugyanakkor a hagyományos megoldás esetén is fontos tisztázni a klasszikus tesztelméletből fakadó korlátokat. Ennek érdekében minden olyan publikációt, valamint gyakorlati példát érdemes a jövőben feltérképezni és bemutatni, amelyek elősegíthetik a fenti modellek és szoftverek kutatásokban való felhasználását.

### Felhasznált irodalom

- Bertram, D. (2014): "Likert Scales". CPSC 681 – Topic Report. URL: <http://www.alhuda.net/2012/PA/2014/topic-dane-likert.pdf>. Utolsó letöltés: 2015. július 16.
- Brown, J. D. (2011): Likert items and scales of measurement? JALT Testing & Evaluation SIG Newsletter. March 15 (1) 10-14. URL: <http://jalt.org/test/PDF/Brown34.pdf> Utolsó letöltés: 2015. július 16.
- Horváth György (2004): *A kérdőíves módszer*. Műszaki Könyvkiadó, Budapest
- Kehl Dániel (2012): Mintaelemszám tervezés Likert-skálás lekérdezések esetén klasszikus és bayesi keretek között; Doktori értekezés; PTE-KTK URL: <http://ktk.pte.hu/sites/default/files/mellekletek/2014/07/Kehl%20Daniel%20-%20disszertacio.pdf> Utolsó letöltés: 2014.09.27.
- Kontra József (2008): Középiskolások olvasás iránti attitűdjeinek vizsgálata klasszikus és modern tesztelméleti eszközökkel. In: Kereszty Orsolya (szerk.) *Új utak, szemléletmódok, módszerek a pedagógiában*. Kaposvár, Kaposvári Egyetem PFK. 153-159. URL: [http://web.t-online.hu/kontraxi/pdf/Kontra\\_Kzpolvklm.pdf](http://web.t-online.hu/kontraxi/pdf/Kontra_Kzpolvklm.pdf) Utolsó letöltés: 2016. január 13.
- Linacre, J. M. (2006): A User's Guide to Winsteps ministeps Rasch-model Computer programs URL: <http://ifile.hkedcity.net/1/001/950/public/Secondary/EI0020070012/winsteps.pdf> Utolsó letöltés: 2015.05.18.
- Molnár Gyöngyvér (2003): Az ismeretek alkalmazásának vizsgálata modern tesztelméleti (IRT) eszközökkel. *Magyar Pedagógia*. 103. évf. 4. 423-446.
- Molnár Gyöngyvér (2008): A Rasch-modell kiterjesztése nem dichotóm adatok elemzésére: a rangskálás modell és a parciális kredit modell. *Iskolakultúra*. 1-2. 66-77.
- Nahalka István (1996): A statisztikai módszerek pedagógiai alkalmazásának indokai, statisztikai alapfogalmak. in: Falus Iván (szerk.) *Bevezetés a pedagógiai kutatás módszereibe*. Keraban Kiadó, Budapest. 343-356.



- Rózsa Sándor – Nagybányai Nagy Olivér – Oláh Attila (2006): A pszichológiai mérés alapjai; Elmélet, módszer és gyakorlati alkalmazás, Bölcsész Konzorcium URL: <http://mek.niif.hu/05500/05536/05536.pdf> Utolsó letöltés: 2015.04.27.
- Schulz, W. és Sibberns, H. (2004): IEA Civic Education Study. Technical Report, IEA
- Wu, M. L., Adams, R. J., Wilson, M. R. és Haldane, S. A. (2007). ACER ConQuest Version 2: Generalised item response modelling software. Camberwell: Australian Council for Educational Research. URL: <http://www.acer.edu.au/documents/Conquest-Manualv2-hires.pdf> Utolsó letöltés: 2014.05.17.