



KÁZMÉR KLÁRA

## A PISA-FELMÉRÉS MÓDSZERTANA ÉS KRITIKÁJA

**A PISA-FELMÉRÉS SAJÁTOSSÁGAI**  
2018-ban a tizenöt éves diákok ismét megmutatják, milyen képességekkel rendelkeznek az írott szövegek megértése terén, vagyis elérkezik a háromévente megszervezett PISA-felmérés következő ciklusa. Jelen írásban a PISA-felmérést, napjaink egyik legismertebb és legmeghatározóbb nemzetközi vizsgálatát fogom részletebben bemutatni Balázi Ildikó és kollégái munkája alapján (2010).

A PISA az OECD, vagyis a Gazdasági Együttműködési és Fejlesztési Szervezet egyik legnagyobb szabású felmérése, amely a Programme for International Student Assessment (Nemzetközi Tanulói Teljesítménymérés Programja) elnevezésből jött létre az 1990-es évek végén. Ez a felmérés a tizenöt éves diákok képességeit tesztelte eleinte három tudományterületen: matematikából, szövegértésből és természettudományból. A felmérések háromévente valósulnak meg, és mindegyik tesztelésben az imént említett három területből az egyik tesztelésére nagyobb hangsúlyt fektetnek, mint a többi tudományterületére (Balázi et. al. 2010: 11). A kezdetekben 32 ország 250 000 diákja vett részt a felmérésben. Napjainkban már 65 ország kb. 510 000 diákja tölti ki a PISA-teszteket

világszerte, és már nemcsak az imént említett három területből, hanem digitális szövegértésből, problémamegoldásból és pénzügyi ismeretekből is. Ez a felmérés nemcsak a diákok képességeinek feltérképezését célozza meg az említett tudományterületekből, hanem háttérkérdőívek segítségével olyan információkat tud meg az adatközlőkről, amelyek befolyásolhatják a diákok iskolai teljesítményét is. Az ilyen háttérinformációk közé tartozik pl. a családi háttér, a látogatott iskola klímája vagy a tanulók egyéni jellemzői (nembeli különbségek, motiváltság vagy annak hiánya a tanulás és olvasás terén) (OECD 2004).

### A PISA-TESZTELÉS MÓDSZERTANA

A PISA-felmérés fő módszertani jellemzői:

- *A mintavétel* – A felmérésben olyan tanulók vesznek részt, akik 15 év és 3 hónapnál idősebbek, de 16 év és 2 hónapnál fiatalabbak. Egy országból minimum 150 iskolának részt kell vennie a felmérésben, és minden iskolából legalább 35 tanulónak kell bekapcsolódnia a tesztelésbe. A diákokat véletlenszerűen választják ki a PPS módszer segítségével (probability proportional to size – elemszámmal

azonos valószínűség). Ennek a módszernek a lényege, hogy ha különböző nagyságú csoportokból választják ki a felmérésben résztvevőket, és a mintavétel ugyanazzal a valószínűséggel történik, akkor egy személy kiválasztása egy nagyobb csoportból valószínűbb, mintha egy kisebb csoportból választanák ki. A PISA készítői figyelembe veszik az iskolai mintánál az elsődleges – pl. az iskola típusa – és a másodlagos rétegezési szempontokat, amelyek országonként eltérnek (Lannert 2015).

- *A felmérés mérőeszközei* – A PISA-felmérésnek kétféle mérőeszköze van: a háttérkérdőívek és a tesztek. A tartalmi keretek foglalják össze a háttérkérdőívek és a tesztek alapjait is. Az első lépések egyikeként a tartalmi keret jön létre, amely magába foglalja az egyes tudományterületekhez kapcsolódó tartalmakat és azokat a kontextusokat, amelyekben a diákok a mért képességeket használják. A tesztek a szövegek és a hozzájuk kapcsolódó kérdések (itemek), feladatok alkotják. A kérdés típusa alapján megkülönböztetünk feleletválasztós és nyílt végű kérdést. A feleletválasztós kérdések lehetnek egyszerű feleletválasztósak és többválaszosak is. Az előbbi típusnál a diákoknak megadott számú lehetőségek közül kell kiválasztaniuk a helyes

megoldást. Ezzel szemben a másik típusnál több kijelentést kell kategóriákba csoportosítani, pl. el kell dönteni róla, hogy igaz vagy hamis, tény vagy vélemény. A nyílt végű kérdések lehetnek rövid választ vagy kifejtős választ igénylők. A rövid választos kérdések egy-két szóval megválaszolhatóak. A kifejtős kérdésekre bővebb válaszok érkeznek, melyek magába foglalják a válaszadó gondolatmenetét, magyarázatát is (Balázi 2010).

■ **Fordítás és kulturális adaptáció** – A PISA-feladatokat először angol és francia nyelven készítik el. A nemzeti változatok ezeknek a teszteknek a mintájára születnek meg az adott részt vevő országokban. A fordítás során kulturális adaptálás alkalmazható, vagyis változtatni lehet az eredeti szövegen, pl. mértékegységeket lehet cserélni az eltérő mértékegységeket használó országokban. Az elkészült nemzeti változat több ellenőrzésen megy át a tesztelés előtt, ugyanis a tesztek szövegeinek és a kérdéseknek a részt vevő országok tanítási nyelvein ugyanolyan jelentéssel kell rendelkeznie (Balázi 2010).

■ **A tesztek létrehozása és rendszerezése** – A matematikából, természettudományból és szövegértésből összeállított tesztek kérdéseket, feladatokat tartalmaznak, melyeket klaszterekbe rendeznek. Mindegyik klaszter kitöltésére a diákoknak 30 percük van. A kérdéseket 13 tesztfüzetbe rendezik, és a füzetek 4 klaszterből állnak. A 15 éves diákok véletlenszerűen kapnak 1-1 füzetet, melyet 2 órán át töltenek ki (Lannert 2015).

Az iménti információk alapján elmondható, hogy a PISA-felmérésekben a diákok nem ugyanazokat a feladatokat oldják meg, mert különböző tesztfüzetek készülnek. Ezekben a tesztfüzetekben nem szerepel az összes feladat, csak a feladatok egy része található meg bennük. A PISA-t érő kritikák egy része éppen a mérés ezen sajátosságában található.

### A PISA-FELMÉRÉSEK KRITIKÁJA

Annak ellenére, hogy a PISA-felmérés a 21. század egyik legszéleskörűbb tesztelése, egyre több kritika éri a tesztet – főleg módszertani szempontból. Az alábbiakban a PISA-tesztek módszertani kritikáit ismertetem Lannert Judit tanulmánya alapján (Lannert 2015). A PISA-val összefüggő kritikákat a következő három nagy csoportra oszthatjuk:

■ **Kultúrafüggetlenség:** A PISA filozófiája kultúrafüggetlen, vagyis nincsen tekintettel az egyes országok kultúrájára, jellegzetességeire. Ezt az is bizonyítja, hogy a diákok mindegyik részt vevő országban ugyanazokat a feladatokat oldják meg. Az indiai bírálók elmondták, hogy ha pl. egy diák soha nem látott még ATM automatát, akkor hogyan tudná megérteni és megválaszolni helyesen az azzal kapcsolatos kérdéseket. A tesztek nemcsak az adott országok kultúráját nem veszik figyelembe, hanem a tantervi sajátosságokat sem. Pl. a norvég diákok a tesztelések alkalmával nem találkoznak a következő témákkal: vízi erőművek, északi fény, olajkitermelés, haltenyésztés (Sjoberg 2012).

■ **Politikai célok:** A PISA-t bírálók kiemelik azt is, hogy a tesztek eredményei leegyszerűsített képet nyújtanak a diákok képességeiről, többek között azért is, mert nem veszik figyelembe a nemzeti sajátosságokat. A tesztelés talán éppen ezért vált olyan népszerűvé, mert egyszerű üzenetet juttat el a politika számára is. Politikai célokra könnyedén felhasználhatóak a könnyen dekódolható rangsorok. Az egyik legnagyobb kritika a PISA-t azért is éri, mert egyre nagyobb hatást gyakorol a felmérésben részt vevő országok oktatáspolitikájára, bár erre az OECD-nek hivatalosan nincs joga.

■ **Módszertani problémák:**

■ **Fordításbeli eltérések:** Az egyik terület, amelyben a kutatók hiányosságokat észleltek a PISA-tesztekkel kapcsolatban, az a fordítás volt. A tesztek különböző nyelvekre történő lefordítására nagy hangsúlyt fektetnek, és megjelenésük előtt komoly eljárásokon mennek keresztül a tesztek. Ennek ellenére a tesztelésekben előfordulnak jelentősebb fordítási hibák, pl. a 2006-os tesztelésnél a skandináv országokban történt, hogy a Dolly klónozásáról szóló szöveg alapvető különbségeket tartalmazott a norvég, a finn és a svéd verzióban is (Sjoberg 2012). A szövegek és a feladatok lefordításában levő eltérések nemcsak a különböző nyelvű szövegek között tapasztalhatóak, hanem az ugyanazon nyelvterületen levő, de eltérő országok feladatsorai között is, pl. jelentős fordításbeli különbségeket találtak az osztrák és a német tesztek összevetésekor is (Wuttke 2007). Ezeket az eltéréseket a tesztek alkotói azzal próbálják magyarázni, hogy a fordítók kapkodása és időhiánya okozhatja a szövegek és feladatok

közi különbségeket (Lannert 2015). Egy francia tanulmány felhívja a figyelmet arra is, hogy a kérdések eltérő megfogalmazása befolyásolhatja a helyes válaszok arányát (Blum-Guérin-Pace 2000).

■ **Mintavétel és hiányzó adatok:** A PISA-felmérésben tizenöt éves diákok vesznek részt, de vannak olyan országok, amelyekben a 15 évesek egy része már nem része az iskolarendszernek. Ilyen ország pl. Törökország vagy Mexikó (Wuttke 2007). A mintavétellel kapcsolatos ferdtetések közé sorolható az is, hogy az arab államokban a lányok PISA-teljesítménye kimagasló, minden bizonnyal azért, mert ebből a korosztályból többnyire a leggazdagabb családok lánygyerekei járnak iskolába. A mintavétellel kapcsolatos következő problémák közé tartozik az is, hogy a szabályok alapján a PISA-tesztelésekből kihagyhatóak a lakosság öt százalékáig azok a tanulók, akik magatartási vagy tanulási problémákkal rendelkeznek. Ezeknek a diákoknak a szándékos kihagyása a mintából szintén jelentősen torzíthatja a kapott eredményeket. Az eredményeket befolyásolhatja az a tény is, hogy a tanulók 2 órán át írják a tesztet. A hosszú időtartam következtében a diákok elfáradnak, és mivel nem jut idejük az összes szöveg elolvasására, ezért aztán tippelnek vagy nem adnak semmilyen választ sem. Ez utóbbira a PISA-bírálok is felhívják a figyelmet, hogy a tesztelésben sok a hiányzó adat. A PISA az időhiány és a fáradtság hatásával nem foglalkozik, nem méri azt (Wuttke 2007).

■ **Nem egyforma tesztfüzetek és feladatok:** A PISA-felmérésben többféle tesztfüzet készül eltérő feladatokkal, vagyis a tesztelésben részt vevők nem ugyanazokat a feladatokat oldják meg. Előfordult már az is, hogy a tesztfüzetek nem egyforma nehézségűek voltak. Pl. az eddigi egyik legkimagaslóbb eltérést két tesztfüzet között az USA-ban észlelték, ahol az egyik tesztfüzetet kitöltők átlagban 444 pontot értek el, míg a másik verzió dolgozók átlagosan 512 pontot gyűjtöttek össze (Wuttke 2007). Ahogyan azt már említettem, a tesztelésekre háromévente kerül sor, és mindegyik felmérésben más-más tudományterületre helyezik a hangsúlyt, a többi területből kisebb mértékben szerepelnek kérdések a tesztben.

■ **Rasch-módszer:** A tesztek kiértékeléséhez a Rasch-módszert használják,

mely az Item Response Theory egyik modellje (valószínűségi tesztelmélet). A Rasch-modell George Raschról kapta a nevét. A modell megalkotója abból a feltevésből indult ki, hogy a magasabb képességszintű személyek nagyobb valószínűséggel oldanak meg bármilyen típusú kérdést, mint az összes többi személy. Ez a modell csak akkor működik, ha minden kérdést a jobb képességűek oldanak meg helyesen, mint a többiek. Goldstein kritizálja a modell egydimenzióosságát, és többdimenziós modellek alkalmazását javasolja (Goldstein 2004). Az alkalmazott modell másik kritikus pontja az a feltevés, hogy az egyforma képességekkel rendelkező diákok egyforma valószínűséggel válaszolják meg a kérdéseket (Yamamoto 2002). Továbbá, a modell szerint egy kérdés akkor nehezebb, mint a többi, ha kevesebben tudják helyesen megoldani. Azok a kérdések, amelyeket többen oldanak meg helyesen, a könnyebbek közé tartoznak. Ez a modell a dichotóm kérdésekre van kidolgozva, vagyis olyan tételekre, melyeknél elsősorban két válaszlehetőség adott: igen-nem, jó-rossz. A Rasch-modell további hátránya, hogy a rész-

ben jó válaszok elemzése nem lehetséges (Lannert 2015).

Sokan támadják a PISA-t a fordítási problémák, a mintavétel és a Rasch-módszer alkalmazása miatt is. Többen újraszámolják az adatokat, és más módszereket alkalmazva nagyobb különbségeket találnak – nem a legjobb és a legrosszabbul teljesítők között, de a középszerű körében. Ezért sokan azt hiszik, hogy a PISA-rangsorolás félrevezető lehet, attól is függően, hogy milyen tesztkérdéseket tartalmaz a tesztelés. Többen állítják, hogy a Rasch-módszer nem megfelelő a PISA-tesztelésre. Szintén kritikaként merül fel, hogy a PISA-tesztelés úgy teszti a tanulókat, hogy nem veszi figyelembe a különböző anyanyelvű és kultúrájú résztvevőket.

A tesztek készítőinek nagyobb hangsúlyt kellene fektetni az imént felsorolt kritikákra, hogy a felmérésben résztvevők vizsgált képességeiről még hitelesebb képet kapjunk. 2018 áprilisa után kiderül, hogyan teljesítettek a tesztelésben a diákok, és a tesztkészítők módosítottak-e a felmérés módszertanán vagy sem.

#### FELHASZNÁLT SZAKIRODALOM:

- BALÁZSI Ildikó, OSTORICS László, SCHUMANN Róbert, SZALAY Balázs és SZEPESI Ildikó 2010. *A PISA 2009 tartalmi és technikai jellemzői*. Oktatási Hivatal, Budapest.
- BLUM A.; Guérin-Pace, F. 2000. *De Lettres et des Chi.res. Des tests d'intelligence à l'évaluation du „savoir lire“, un siècle de polémiques*. Paris: Fayard.
- GOLDSTEIN Harvey 2004. *International comparisons of student attainment: some issues arising from the PISA study*. *Assessment in Education*.
- Joachim Wuttke 2007. *Uncertainties and Bias in PISA*, [http://www.oxydiane.net/IMG/pdf/Uncertainties\\_and\\_Bias\\_in\\_PISA.pdf](http://www.oxydiane.net/IMG/pdf/Uncertainties_and_Bias_in_PISA.pdf) (letöltve: 2017. 11. 09.)
- Lannert Judit 2015. *A PISA-adatok használata és értelmezése: A módszertani kritikák tükrében*, *Educatio* 2015/2, 18-29.
- OECD 2004. *Messages from PISA 2000*. OECD, <http://www.oecd.org/education/school/programmeforinternationalstudentsessment/pisa/34107978.pdf> (letöltve: 2017. 11. 09.)
- SJOBERG, Svein 2012. *PISA: Politics, fundamental problems and intriguing results*. English manuscript. *La Revue, Recherches en Education* Numéro 14.
- Yamamoto KENTARO 2002. *Estimating PISA students on the IALS prose literacy scale*. Educational Testing Service.

