

LANNERT JUDIT

A PISA-adatok használata és értelmezése

A módszertani kritikák tükrében

Az OECD egyik legnagyobb hatású nemzetközi programja a PISA néven ismert nemzetközi tanulói teljesítménymérés (Programme for International Student Assessment) szinte forradalmasította az oktatásról való tudásunkat. Ez a vizsgálat a mai napig egyedülálló módon teszi lehetővé, hogy időben és térben is összehasonlítható elemzéseket tudjanak végezni a világ oktatási rendszereiről. Napjainkra valódi nagyiparrá vált ez a terület. Már a kezdetekben is 32 ország több mint 250 ezer tanulójának teljesítményét mérték három (matematika, szövegértés és természettudomány) területen, napjainkra pedig már 510 ezer tanuló tölti ki a tesztek a világ 65 országában és már nem csak a matematika, szövegértés és természettudomány területeken, de a digitális szövegértés, problémamegoldás és pénzügyi ismeretek területeken is. Egy ekkora vállalkozásnak csodálói, irigyei és ellenlábasai egyaránt akadnak. Már a kezdeti szakaszban is sok kritika érte a felmérést és a belőle kinyert eredményeket és ez mára sem csitult el. Tanulmányomban azt foglalom össze, hogy milyen kritikák érik ma is a PISA-felmérést, és ezek hogyan árnyalják a mai PISA-ról vallott felfogásunkat és a PISA-adatok felhasználásának célját és módjait.

PISA módszertan

Érdemes kicsit részletesebben kitérni a vizsgálat módszertani sajátosságaira, mivel nagyon sok kritika ezen a téren éri a PISA-t.

Minta

15 éves tanulókat mér fel a vizsgálat, pontosabban 15 éves és három hónapostól 16 éves és két hónaposig, alkalmazkodva a két földteke közötti évszakok és ezáltal tanév eltérésehez. Minimum 150 iskolának szerepelnie kell a mintában, minden iskolában véletlenszerűen mintegy 35 tanulót választanak ki. Összességében országonként 4500 tanuló a mintanagyság. A mintavétel a PPS módszer követi (probability proportional to size), ahol az iskolai mintában egyrészt az elsődleges (Magyarország esetén pl. az iskola típusa) és

► Educatio 2015/2. Lannert Judit: A PISA-adatok használata és értelmezése, 18–29. pp.

másodlagos országoként változó rétegzési szempontokat veszik figyelembe, de az iskolaméret is számít. A kisiskolák problémájára is külön metódust találtak ki. A mintavétel a PISA felvétel egyik legkidolgozottabb része, ezt részletesen leírja a technikai jelentés. A koncepció kialakítása, a nyelvi és kulturális eltérések kezelése, a mintavétel megalapozottsága, az adatfelvétel biztonságos lebonyolítása és az adatok kezelésének és elemzésének módszertanilag megalapozott módszerei mind helyet kapnak a technikai jelentésben.

Testtfejllesztés

A matematika, szövegértés és természettudomány standard itemjeit 2012-ben 13 klaszterbe rendezték (hét matematikai, három szövegértési és három természettudományi), mindegyik klaszter kitöltése 30 percet igényelt. Az itemeket 13 tesztfüzetbe rendezték, ahol mindegyik füzet négy klaszterből állt. A 15 éves tanulókat minden országban véletlenszerűen rendelték hozzá egy-egy tesztfüzethez, akik két órán át töltötték ki a tesztek egy ötperces szünet beiktatásával.

Skálák készítésének módszertana

A PISA azon a feltevésen alapszik, hogy létezik egyfajta logikus hierarchia az adatokban (Rasch-modell). A válaszok alapján a tanulókat és az itemeket is besorolják képességszintekre, illetve nehézségi fokozatokba. Azok az itemek, amelyeket többen oldanak meg, könnyebbnek minősülnek, mint amit kevesebben tudnak megoldani. A cél egy képesség-skála teljes lefedése, ami megfelelően nagy minta esetén hozható létre. (Molnár, 2005).

A klasszikus tesztelmélet szerint a mérés pontossága növelhető az itemek számának növelésével, így általában a teljesítménytesztek legalább 70 itemet tartalmaznak. Ugyanakkor nem mindig van lehetőség ilyen sok item lekérdezésére, a PISA tesztek során is nem minden kérdést válaszol meg minden tanuló minden országban. Így a tanulókra ún. valószínűsíthető (plausible) képességértékeket számítanak a válaszok, a háttérváltozók és a modell paramétereinek feltételes eloszlása alapján. Ezek az értékek nem tesztpontszámok a hagyományos értelemben, közvetítőként szolgálnak arra, hogy megfelelően tudják becsülni a populáció jellemzőit. Amennyiben a modell korrekt, akkor ezek a plauzibilis értékek megfelelően becslik a népesség jellemzőit (Yamamoto, 2002).

A PISA négylépcsős eljárásban alakítja ki az ország-rangsorokat. Első lépésként kiszámolja az országokénti Rasch-paramétereket, majd 30 ország 15 ezer tanulójának véletlen almintáján ezeket a paramétereket az egész mintára is kiszámolja. A harmadik lépésben létrehozzák a valószínű (plauzibilis) értékeket, amelyek a mérni kívánt látens dimenzió poszterior eloszlása alapján létrehozott random számok. A negyedik lépésben varianciaanalízissel hasonlítják össze az országokat (Kreiner, 2012).

A PISA módszertani kritikája

A legtöbb kritika azért éri a PISA-t, mert – ahogy a bírálók fogalmazzák – az országokat egy dimenzió mentén sorba állítva leegyszerűsített képet ad egy ennél bonyolultabb világról, figyelmen kívül hagyva a nemzeti sajátosságokat. Sajátos paradoxon, hogy a PISA épp azért vált ennyire nagy hatásúvá, mert a média és a politika számára is egyszerű és hatásos üzenetet tudott közvetíteni. A rangsor könnyen értelmezhető és politikai célra jól haszno-

sítható termék. Ugyanakkor nem véletlen, hogy habár módszertani kritikák a kezdetektől fogva érték a PISA-t, ezek a kritikai hangok a 2012-es eredmények után erősödtek fel, amikor az ázsiai országok előretörésével Európa és az angolszász országok is hátrébb sorolódtak a PISA rangsorban. A legutolsó PISA felvétel eredményei alapján az első 10 ország közé nem került be angolszász ország, az első hét ország ázsiai (távol-keleti) és az Európai Unióból csak Hollandia fért be Svájc és Lichtenstein után. A finnek hátrébb kerültek és sokat rontottak a korábban jobb eredményt elérők, mint Új-Zéland, vagy Svédország.

A PISA-eredmények fogadtatása persze sohasem volt konfliktusmentes. Alapvetően három tényezőtől függ, hogyan fogadja egy ország az eredményt. Egyrészt számít az eredmény (jó eredmény esetén még nem hangzott el olyan, hogy esetleg mintavételi probléma okozhatta a kiugrást, míg romlás esetén elég gyakran hivatkoznak különböző okokra és vetnek fel kritikai megjegyzéseket). Másrészt függ attól, hogy politikai választások elé néz-e egy ország vagy nem. Előbbi esetben mindig kiélezettebb a fogadtatás és az adott érdekeknek megfelelő az eredmények értelmezése vagy éppen elhallgatása. Harmadrészt pedig függ a fogadtatás attól is, az, hogy vannak-e olyan szakemberek, akik rendelkeznek azokkal a kompetenciákkal, amik a PISA adatok elemzéséhez és értelmezéséhez szükségesek.

A módszertani kritikák is elsősorban azt a kérdést feszegetik, lehetséges-e egy többnyelvű és kultúrájú közegben kontextustól függetlenül mérni a tanulók és ezáltal az országok teljesítményét. A mintavétel, a tesztek fordítása, a Rasch-módszer megfelelő alkalmazása, a rangsorok mind összetűz alá kerültek. Egyre többen számolják újra az adatokat más modellek alapján és ennek eredménye általában az, hogy a rangsor eleje és vége ugyan nem változik, de a középsőben nagyon nagy eltérések is tapasztalhatók attól függően, hogy milyen módszert használnak. Ezért állítják egyre többen, hogy a PISA nem jól mér, hiszen az országok nem összehasonlíthatók egy egydimenziós skálán. A rangsorok – mondják – félrevezetőek, attól függően milyen tesztkérdéseket választunk, eltérő rangsorokat kaphatunk. ráadásul a mintavétel során is történnek hibák, a Rasch-módszer a PISA céljára nem megfelelő (Goldstein, 2004; Kreiner, 2010).

Az alábbiakban ezeket a módszertani kritikákat részletesebben is körbejárjuk.

Kultúrafüggetlen teszt, fordítási problémák

Habár a PISA tesztek fordításának komoly procedurális minőségbiztosítása van, mégis újra és újra felmerülnek fordítási, interpretálási problémák. A skandináv országok esetén a Dolly klónozásáról szóló történet finn, svéd és norvég fordítása alapvető eltéréseket tartalmazott (Sjoberg, 2014). Az ugyanolyan nyelvterületen lévő, de különböző országok esetén is meglepő eltéréseket lehet tapasztalni, így az osztrák és német tesztek összehasonlításánál is találhatók különbségek (Wutke, 2007). Ezt annak tudják be, hogy dacára a komoly procedurális eljárásnak, a fordítók gyakran kifutnak az időből és a kapkodás eredménye a fordítási eltérések.

Mintavétel, válaszadási arány, hiányzó adatok

A PISA a 15 éves tanulókat méri, de sok országban a 15 évesek egy része már nincs az iskolarendszerben. Mexikó, Törökország esetén a beiskolázás ebben a korban már 60 százalék alatt van (Wutke, 2007). Ilyen jellegű torzításnak tudható be valószínűleg az is,

hogy az arab országok esetén a lányok teljesítménye átlagon felüli, mert valószínűleg csak a legjobb helyzetben lévő családok lánygyermekai találhatóak az iskolában, ezekben az országokban ebben az életkorban.

A PISA szabályai megengedik, hogy a népesség 5 százalékáig a tanulás vagy magartászavaros tanulókat kivegyék a mintából. Ez az arány ugyanakkor nagyon változó, az OECD-országok közt 0,7%-7,3%-ig terjed. Előfordul az is, hogy bizonyos rétegek nincsenek kellőképpen reprezentálva a mintában. Ezt általában súlyozással korrigálják, de nem mindig veszik észre a hibát. Ráadásul a súlyozás nem tudja korrigálni azokat a láthatatlan torzításokat, amit például az okoz, ha szisztematikusan a legrosszabbul teljesítők hiányoznak a felmérés napján (*Baird et al*, 2011).

Van, amikor feltűnően hiányoznak a hiányzó adatok (missing). 2003-ban Lengyelország esetén nem volt egy tanuló sem, aki kevesebb, mint 25 kérdésre válaszolt volna és volt hét olyan kérdés, amire mindenki válaszolt. A külső szemlélőben felmerül a gyanú, hogy a hiányos tesztfüzeteket nem értékelték (*Wutke*, 2007).

A kétórás teszt meglehetősen megterhelő, így nem meglepő, hogy gyakran nem tudják befejezni a tanulók és így sok hiányzó (missing) adat is van. Ugyanakkor ennek aránya nagyon változó országonként. 2003-ban a negyedik blokkban az el nem ért kérdések aránya a hollandok esetén 1%, a mexikói tanulók esetén több mint egynegyed. A meg nem válaszolt kérdések aránya a holland tanulók esetén az első blokkban 2,5%, a negyedik blokk esetén 4%, ugyanez az arány a görög tanulók esetén 11%-ról 24,4%-ra duzzadt. Az idő szűkét nagyon különbözőképpen menedzselik a különböző országok tanulói. A holland tanulók szinte minden kérdést megpróbálnak megválaszolni, de a kérdőív végére megnő a találgatás aránya, az osztrák és német tanulók sok kérdést átugranak, de így a többi kérdés megválaszolására elegendő idő jut. A görög tanulók éppen ellenkezőleg, meglepődnek, amikor azzal szembesülnek, hogy már kevés idejük van. Az első blokkban még jobb válaszokat produkálnak, mint a portugál tanulók, de az utolsó blokk esetén az el nem ért és hiányzó válaszok aránya eléri a 35 százalékot, ami miatt a PISA rangsorban le is csúsznak. Ez is egy olyan dimenzió, az elfáradás és időmenedzsment, aminek a hatását nem méri a PISA modell (*Wutke*, 2007).

A PISA mintáját legfrissebben ért kritika Shanghai kiemelkedő teljesítményére reflektál. Tom Loveless, egy harvardi professzor megkérdőjelezte a kínai eredményt, mondván, több mint gyanús, hogy bár 2009-ben 12 kínai tartomány is részt vett a PISA-ban, de csak a Shanghai-i eredmények lettek publikusak, ahol bár a 15 évesek hivatalos beiskolázási aránya 84 százalék, a 23 milliós városban csupán 108 ezer 15 évest tartanak számon (összehasonlításképpen a 10 milliós Magyarországon is hasonlóan 111 ezer 15 éves van). A furcsaság magyarázata az, hogy a migráns tanulókat, gyerekeket nem számolják be a statisztikába, és ők a PISA felvételen sem vesznek részt.

Különböző tesztfüzetek és itemek

Sokaknak talán meglepő, de a PISA felvétel során a tanulók nem egyforma feladatokat oldanak meg és nem is minden kompetenciaterületen töltenek ki tesztfüzetet. Többféle tesztfüzet készül, hogy elegendő számú item legyen a modell megalkotásához, ugyanakkor életszerűen egy tanulóra nem juthat életkorához képest megterhelően nagyszámú kérdés. Mintegy 165 itemből így 60-nál több nem jut egy tanulóra. Ugyanakkor a tesztfüzetek nehézségét nem mindig sikerül egyformára kalibrálni, legalábbis erre utal, hogy

sokszor ugyanazon területet mérő különböző tesztfüzetek kitöltése eltérő eredményeket produkál. Az egyik legnagyobb különbséget az USA-ban tapasztalták, ahol az egyik tesztfüzetet kitöltők átlagosan 444 pontot értek el, míg a másikon dolgozók 512 pontot. Ez felveti annak a gyanúját, hogy az itemek csoportosítása gyakran önkényes és nem eléggé homogén (Wutke, 2007)

A tesztfüzetek mindegyike tartalmazza az adott ciklus fő kompetenciaterületét (2012-ben ez a matematika volt), de a többi kompetenciaterület nem mindegyik tesztfüzetben szerepel. Ugyanakkor minden tanulóra megadják mindegyik kompetenciaterületen a plauzibilis értékeket, amit a tanulónak a más kompetenciaterületeken adott válaszai és egyéb adatok alapján becsülnek. Ez a fajta eljárás, amennyiben korrekt, csökkenti az átlag standard hibáját. Ugyanakkor érdekes eltérések is előfordulnak, Görögország esetén 2003-ban a hivatalos természettudományi kompetencia pontszám 481 volt, ami 16 ponttal volt magasabb, mint azon görög tanulóknak az átlagos teljesítménye, akik valóban kitöltöttek természettudományos tesztkérdéseket is (Wutke, 2007).

A tesztkérdések közül – a Rasch-modell feltételeinek megfelelően – kiszűrik azokat, amelyek országonként vagy országcsoportonként másképpen működnek. Ez azt is eredményezi, hogy viszonylag kevés kérdés marad ahhoz, hogy megnyugtatóan tudjanak mérni a tesztelmélet elvárásai szerint. 2009-ben 131 szövegértési, 35 matematikai és 53 természettudományos item szerepelt a tesztfüzetekben. Ezt sokan kevésnek tartják és a PISA egyik gyengeségének is egyben (Baird et al, 2011).

IR modell, Rasch-módszer

Az IR modell egydimenziós, éppen ezért csak akkor működik, ha minden kérdés esetében igaz, hogy a jobb képességű nagyobb valószínűséggel találja el a helyes választ, mint a gyengébb képességű. Goldstein (Goldstein, 2004) kritizálta ezt a fajta modell-eljárást és olyan többdimenziós modellek használatát javasolta, amelyek azt feltételezik, hogy több dimenzió húzódik meg egy egyén teljesítménye mögött és a teljesítménye ezen dimenziókban eltérhet. Az IRT modell alapfeltevését, a feltételes függetlenséget számosan az adatelemzés legkritikusabb részének tartják. A feltételes függetlenség nem mást jelent, minthogy az azonos képességű válaszadók azonos valószínűséggel válaszolnak a kérdésekre, függetlenül egyéb, csoportjellegű tényezőtől (Yamamoto, 2002).

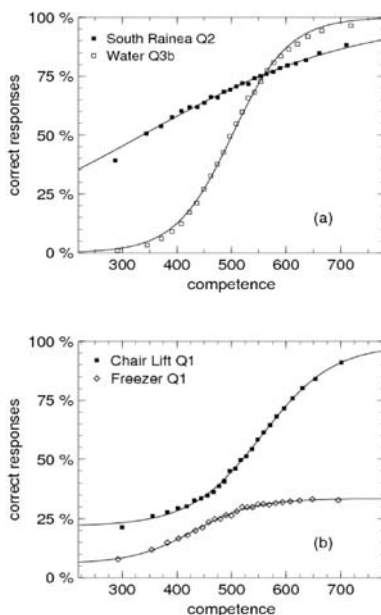
Többen is újrászámolták az adatokat az adatbázis alapján, hogy teszteljék valóban igaz-e, hogy – amint azt a PISA modellje feltételezi – az egy kompetenciaterülethez tartozó itemeknek ugyanolyan alakzatot kell felvenniük, ahol az egyedüli eltérést horizontálisan csak az item nehézségi foka jelentheti. Általában ettől eltérő alakzatokat tapasztaltak kutatók több item esetén (Wutke, 2007; Kreiner, 2010).

Wutke négy – ugyanazt a kompetenciaterületet vizsgáló – item kapcsán azt találta, hogy csak az egyik (Water q3) tudja valóban jól megkülönböztetni az adott területen a jobb és gyengébb képességű tanulókat (lásd 1. ábra). A Chair Lift Q1 item esetén azt láthatjuk, hogy alacsony kompetenciaszintnél szinte egyenes a görbe. Ez tipikus jele a találgatásnak. A Freezer Q1 item viszont már 35 százaléknál eléri a telítettséget. Ez arra utal szerinte, hogy a tanulók többsége nem találta ki, mire is kíváncsiak a tesztet készítő. A South Rainea Q2 itemnek gyenge a képesség szerinti megkülönböztető (diszkriminációs) ereje, aminek több magyarázata is lehet. A különböző alpopulációk esetén a nehézség foka különbözik, vagy a megoldási stratégiák különböznek alcsoportonként, jól sikerült ta-

lálgatás is lehet mögötte vagy a mérendő látens képesség és a terület itemjeinek többsége közötti gyenge korreláció (Wutke, 2007). Ez mind arra utal szerinte, hogy az utóbbi három item esetén további paraméterekre van szükség az értelmezésre. Az ábrán a folytonos vonalak többparaméteres modellek illeszkedését mutatja, ami még mindig tartalmaz lineáris mozgást, ami értelmezhető az item nehézségi fokaként. Ugyanakkor ezek a legjobban illeszkedő modellek +/- 30 pont eltérést mutatnak a hivatalos Rasch-modell szerinti nehézségi fokokhoz képest. Ugyanakkor, amennyiben elfogadjuk a több paraméteres modellt, a tanulók rangsorba állítása máris önkényessé válik. Érdekes azt is megfigyelni, hogy a „South Rainea” és a „Chair Lift” kérdések esetén az első megfigyelt érték a görbe alatt található, vagyis a leggyengébb 4% még gyengébben teljesít, mint ahogy azt a modell feltételezi. Ezt magyarázhatja a kooperáció hiánya is, vagyis a tanulók nem akarják kitölteni a kérdőívet, vagy nem veszik azt komolyan. Ez is egy olyan dimenzió, amit a PISA egydimenziós Rasch-modellje nem vesz figyelembe (Wutke, 2007, lásd az 1. ábrát).

A legnagyobb visszhangja egy Rasch-tanítvány dán matematikus (Svend Kreiner) kritikájának lett. Az angol és dán diákok válaszait újraelemelve Kreiner arra az eredményre jut, hogy az itemek különbözőképpen működnek és az adatok nem véletlenszerűen hiányoznak. A 2006-os PISA hatodik tesztfüzetében lévő szövegértési itemjeinek vizsgálata azt mutatta, hogy az angol tanulóknak szisztematikusan könnyebben ment a megoldás, mint a dán fiataloknak, vagyis az itemek nem tekinthetők feltételesan függetlennek. Éppen ezért a PISA rangsor módszertani megalapozottságát nem tartja megfelelőnek (Kreiner, 2012).

1. ábra Néhány item jellemzője, ami a Rasch-modelltől való eltérést mutatja. A folytonos vonal az (a) ábrán egy – a különbségeket magyarázó – két paraméteres modellre illeszkedik, míg a négy-paraméteres (b) a találgatás és félreértést is engedő modellre illeszkedik.



Forrás: Wutke, 2007

Mire jó a PISA?

Fontos hangsúlyozni, hogy a PISA módszertanát bírálók magát a PISA-felmérés jelentőségét nem vonják kétségbe, kritikájuk elsősorban a PISA továbbfejlesztésére irányul. A legtöbb kritika magát a rangsort éri, mondván, hogy ez a kapott, bonyolult információk túlzott leegyszerűsítésének veszélyét hordozza. Szofisztikáltabb módszereket és az egydimenzió helyett több dimenzió, paraméter bevonását javasolják a szakértők. Az országok összehasonlítását célzó kérdések és skála – állítják – nem hasonlíthat egy intelligenciateszthez, ez egy komplex, többdimenziós megközelítést igényel. Éppen ezért a PISA rangsort szinte mindegyik megszüntetni vagy átalakítani kívánja (Goldstein, 2004, Kreiner, 2010). A nem tipikustól – mint ami valóban országspecifikus – szerintük nem szabad „megtisztítani” a vizsgálatot. Éppen ezért más megközelítésre van szükség a kérdések tartalmi és formai megalkotásánál és a válaszok elemzésénél. A PISA-t társadalmasítani szükséges, az elemzési keret kialakításától a tanácsadók bevonásán át az eredmények és módszerek publikálásáig nagyobb nyilvánosságot és főleg transzparenciát kell biztosítani (Goldstein, 2004). A PISA – ahogy mondják – egyszerű lehetőség arra, hogy még többet tudjunk meg arról, miben különbözünk, és jóval kevésbé célszerű versenyként felfogni.

A PISA-val szembeni kritikák, amelyek annak szűklátókörűségére, egydimenziós jellegére vagy korlátozottságára utalnak, nem feltétlenül jogosak. El kell tudni különíteni a PISA-t magát attól, hogy hogyan használják és értelmezik. A PISA-t létrehozók soha nem állították, hogy általános receptet nyújtanának. Ugyanakkor tény és való, hogy a pénzek felett diszponáló döntéshozókat és a médiát leginkább az egyszerű, de hatásos üzenetek érdeklik. Éppen ezért nem valószínű, hogy a rangsorok megszűnnének. A szakmának ugyanakkor törekednie kell arra, hogy a mélyebb üzenetek is átjöjjenek, mert bár igaz, hogy a sorrendek azok, amik a legnagyobb figyelmet felkeltik, de éppen ezek a legkevésbé használható elemek. Ugyanakkor az adatbázis nyilvános, az eljárások nagy része transzparens, az eredmények jól dokumentáltak és továbbgondolhatóak. Felvállaltan nem mér „mindent”, de folyamatosan bővülnek a kompetenciaterületek. A PISA-adatokat az OECD szakértői folyamatosan dolgozzák fel és értelmezik, az eredményeket egyre több kötetben teszik közzé. Sajnálatos módon ezeket nem csak az újságírók vagy politikusok nem olvassák, de a kutatók nagy része vagy a szakma képviselői, a pedagógusok sem. Pedig a térben és időben is kiterjedt elemzések egyik legizgalmasabb produktuma, hogy a mai napig erőteljesen élő, de valójában nem működő mítoszokról rántja le a leplet.

Az oktatási mítoszok leleplezése

Ennek jelentősége éppen abban áll, hogy a hagyományokra vagy „hüvelykujj” szabályokra alapozó oktatáspolitiká, amennyiben van erre készítés, eredményesebb üzemmódra állhat. A PISA eredményekből az évek során tartósan leszűrhető tanulság, hogy a méltányosság és eredményesség kéz a kézben járnak. A hátrányos helyzetű tanulók nincsenek arra kárhóztatva, hogy gyengébben teljesítsenek, az ugyanolyan családi háttérű gyermekek között szintén nagyok lehetnek a különbségek, attól függően, hogy milyen iskolába járnak. A PISA-adatok szerint nincs összefüggés egy ország teljesítménye és a bevándorló gyermekek aránya között sem. Azok az országok teljesítenek jól, ahol nem a született tehetségben, hanem abban hisznek, hogy a szorgalmon, a befektetett energián múlik egy tanuló eredményessége. Ezen felül az is nyilvánvaló, hogy egy bizonyos szint fe-

lett már nem az oktatásra szánt erőforrások mennyisége, hanem azok minősége fontos. Így önmagában az osztályméret és az országok teljesítménye között sincs összefüggés, hiszen a pedagógus felkészültsége az, ami számít. A digitális világ pedig más tananyagot és óraszervezést igényel. Az eredményesebb országok nem új tantárgyak bevezetésével és a tananyag növelésével érnek el eredményt, hanem éppen fordítva, kevesebbet, de mélyebben tanítanak.

Az adatok másodelemzése és egyéb adatokkal való összevetése a hazai sajátosságok felderítése céljából

A PISA eredmények másik hasznos felhasználási terepe az adatok sajátos, nemzeti szempontokat is előtérbe helyező másodelemzése. Egy, az Oktatási Hivatal által kiírt pályázat keretében készült tanulmány kimutatja, hogy a magyar tanulók matematikai önhatékonyasága 2003 óta szignifikánsan romlott, aminek oka, hogy a magyar diákok a zsúfolt tananyagot a környező országokhoz viszonyítva alacsonyabb matematika óraszámiban tanulják az iskolában. Ez óhatatlanul odavezet, hogy nincs idő az iskolában a gondolkodásra, így a jobb képességű tanulók sem sajátítják el a magasabb szintű tanulási stratégiákat. Ezért hiába tanulnak a magyar tanulók sokat iskolán kívül, tanulásuk nem lesz hatékony és matematikai önhatékonyaságuk csökken (Csüllög, D. Molnár & Lannert, 2014).

Bár a PISA-adatok oksági magyarázatokra korlátozottan alkalmasak, egyéb adatbázisokkal összevetve az összecsengő eredmények erősítik a következtetések magyarázó erejét. Magyarország ilyen szempontból kifejezetten szerencsés helyzetben van, mert létezik egy, a tanulók teljesítményét évente PISA kompatibilisan mérő teljes körű felmérés, az Országos Kompetenciamérés. Ez az adatbázis egyéni szinten több mérési időpontban is összehasonlítható, így alkalmas oksági kapcsolatok feltárására is. A PISA-adatokban Magyarország egy dimenzióban kiugró, nálunk magyarázza a világon a legnagyobb mértékben a családi háttér a tanulói eredményeket. Ugyanezt az összefüggést a kompetenciamérés adatbázisa is visszaigazolja. Az igazi kérdés viszont az, hogy ez minek tudható be. A kompetenciamérés 2011-es és 2013-as tanulói szinten összekapcsolt adatain ennek kiderítésére is próbát tesz egy jelenleg zajló kutatás.¹ Az eredmények azt igazolják, hogy a pedagógusok hajlamosak a beskatulyázásra, ezáltal működtetve a szakirodalomban Pygmalion és Gólem hatásként leírt mechanizmusokat. A kompetenciamérés adatain végzett elemzés azt mutatja, hogy a pedagógusok jellemzően a hátrányos helyzetű tanulókat hajlamosak alulértékelni. Az egyik évben alulértékelt diákok között jóval magasabb a következő évben szintén alulértékelt aránya, és ugyanez jellemző a felülértékeltre is, bár kisebb mértékben. Mint ha egy stigmát vinnének magukkal az általánosból a középiskolába. Úgy tűnik ráadásul, hogy a tanári értékelés hatással van a későbbi teljesítményre, hiszen a 8. évfolyamon alulértékelt tanulók teszteredménye alacsonyabb lett 10.-re, míg a felülértékelt diákoké javult. A pedagógusok elfogult félreértékelése és rossz értelemben vett differenciálása (értsd alatta, hogy nem a módszerben, hanem a tanuló iránti elvárásokban differenciálnak) a fundamentuma a magyar oktatási rendszer legnagyobb bűnének, annak, hogy az nemhogy ki-

¹ Oktatási törzsindikátorok és mérési eredmények alapján készülő indikátorok rendszerének fejlesztése a TÁMOP-3.1.8-09/1-2010-0004 azonosító számú projekt keretében” A középiskolák összehasonlítható elemzése a KIR bázisán című kötetben Lannert Judit – Csüllög Krisztina: A diákok ambícióinak és teljesítményének alakulása és változása az általános iskolától a középiskoláig. A továbbtanulás mint a középiskolai munka egy lehetséges indikátora tanulmánya, kézirat

egyenlíti, de még fel is erősíti a családi háttér hatását. Ez a háttére annak, hogy milyen mechanizmusokon keresztül erősíti fel a magyar iskola a családi háttér hatását.

Összegzés

A PISA hatalmas vállalkozás, amelynek üzemeltetése igényli a média és a politikusok folyamatos érdeklődésének fenntartását. Ezt szolgálják a rangsorok, amelyek hatásos és egyszerű üzenetet hordoznak. A PISA jelentősége ugyanakkor abban áll, hogy az adatok mélyebb elemzésével többet tudhatunk meg arról, hogyan működnek az oktatási rendszerek. A PISA-adatok elemzéséhez szükséges módszertani tudás azonban ma még szűkös. Érthető az oktatáskutatók és pedagógusok szorongása, ha érzékelik, hogy valami befolyásolja az életüket, de nem igazán értik ennek működését, módszertanát. Éppen ezért a PISA-tudás társadalmisítására lenne szükség, mint ahogy ezt a PISA-t elsősorban módszertani szempontból bírálók is javasolják.

A PISA-t a globális hatása miatt bírálók nagy része elsősorban nem a PISA-val, hanem a saját országának a PISA adatokra adott válaszával elégedetlen. Két rossz válasz képzelhető el a PISA-rangsor kapcsán, az egyik, amikor nem vesznek róla tudomást és elhallgatják (mint a franciák a legelső időkben), vagy amikor azt versenyként felfogva direkt módon a versenyt erősítik saját rendszereikben (USA, Svédország). A PISA rangsorban előjárók általában a PISA előtt járnak (Lengyelország oktatási reformja már 1997-ben elindult, vagy a finnek pedagóguspolitikája több évtizedes fejlődés eredménye), és nem cibálják az oktatási rendszerüket az aktuális PISA-adatok tükrében, vagy másolják más országok példáját. Használják és értelmezik a PISA-adatokat, de csak egy indikátoraként az oktatásnak és folyamatosan kutatják-elemzik saját maguk is az oktatásukat. A sikeres országokra jellemző még, hogy belülről irányítottak, legyen az a tanuló motiváltsága és szorgalma (távol-keleti országok), vagy a pedagógusok felkészültsége, akik folyamatosan elemzik saját munkájuk eredményét (finnek). A középmezőnyben található a kívülről irányított országokat, amelyek elsősorban a külső kontroll vagy a piaci mechanizmusok felerősítésével válaszoltak a romló adatok láttán. Az utolsó harmadban pedig az irányítatlan országokat találjuk. Erre legjobb példa India, ahol egyáltalán nem ritka, hogy a pedagógus meg sem tartja az óráját.

Nem kérdéses, hogy habár jelenleg a PISA-t ért kritikákra ingerülten válaszol az OECD, hosszabb távon ezek minden bizonnyal termékenyítőleg hatnak a PISA további fejlődésére, hiszen az OECD érdekelt abban, hogy az adatait egyre értőbben elemezzék. Ebben még van tere a fejlődésnek, például a technikai leírás részletebben meg kellene mutatni, hogy hogyan alakították ki a teljesítményszinteket, illetve, hogy a tartalmi kerethez hogyan rendelik hozzá az itemeket.

Magyarország a nagymintás teljesítménymérések terepén akár európai nagyhatalom má is válhatna. Az Oktatási Hivatal által gondozott országos kompetenciamérés, vagy a szegedi egyetemen az elektronikus diagnosztizáláson munkálkodó műhely egyaránt jó alapot adna ehhez. Természetesen ahhoz, hogy ebbe az irányba elmozduljunk, az oktatáspolitikának fel kellene ismernie ennek jelentőségét és nagyobb költségvetéssel támogatni e műhelyek és általában az oktatási adatbázisokat létrehozó intézmények munkáját.

IRODALOM

- BLUM, A. & GUÉRIN-PACE, F. (2000): *Des lettres et des chiffres. Des tests d'intelligence à l'évaluation du „savoir lire”, un siècle de polémiques.* Fayard, Paris,
- BROWN, G; MICKLEWRIGHT, J.; SCHNEPF, S. V. & WALDMANN, R. (2005): Cross-national surveys of learning achievement: how robust are the findings?, *IZA Discussion Papers*, No. 1652.
- CARVALHO, L. M. & COSTA, E. (2014): *Seeing education with one's own eyes and through PISA lenses: considerations on the reception of PISA in European countries.* *Institut of Education*, University of Lisbon, Lisbon.
- CSÜLLÖG, K., MOLNÁR, D. É. & LANNERT, J. (2014): A tanulók matematikai teljesítményét befolyásoló motívumok és stratégiák vizsgálata a 2003-as és 2012-es PISA-mérésekben. In: *Hatások és különbségek*, Oktatási Hivatal, Budapest.
- Education at a Glance 2010. OECD Indicators. Paris: OECD.
- FROESE-GERMAIN, B. (2010): *The OECD, PISA and the Impact on Educational Policy.* Virtual Research Centre. *Canadian Teachers' Federation.* <http://files.eric.ed.gov/fulltext/ED532562.pdf>. Letöltés ideje: 2015. június 7.
- GOLDSTEIN, H. (2004): International comparisons of student attainment: some issues arising from the PISA study. In: *Assessment in Education: Principles, Policy and Practice*, vol. 11, No. 3, pp. 319-330.
- HOPMANN, T. S. & BRINEK, G. (2007): *Introduction: PISA According to PISA – Does PISA Keep What Promises?*, LIT Verlag Vienna, Berlin.
- KREINER, S. (2010): *Is the foundation under PISA solid? A critical look at the scaling model underlying international comparisons of student attainment.* Dept. of Biostatistics, University of Copenhagen, Copenhagen.
- MOLNÁR, Gy. (2005): Az objektív mérés lehetősége: a Rasch-modell. *Iskolakultúra*, március pp. 71-80.
- MONS, N., PONS, X.; VAN ZANTEN, A. & POUILLE, J. (2009): *The reception of PISA in France. Knowledge and Regulation of the Educational System.* KNOWandPOL <http://knowandpol.eu/IMG/pdf/o31.pisa.france.pdf>. Letöltés ideje: 2015. június 7.
- MORTIMORE, P. (2009). *Alternative Models for Analysing and Representing Countries' Performance in PISA.* Brussels. [Expert paper commissioned by the Education International Research Institute] <http://download.ei-ie.org/Docs/WebDepot/Alternative%20Models%20in%20PISA.pdf>. Letöltés ideje: 2015. június 7.
- PISA 2009 Results (2011): *Students on Line. Digital Technologies and Performance– Volume VI*, OECD, Paris.
- BALÁZSI, I., OSTORICS, L., SZALAI, B., SZEPESI, I. & VADÁSZ, Cs. (2013): *PISA 2012 : Összefoglaló jelentés.* Oktatási Hivatal, Budapest.
- RUBIN, D. (1987). *Multiple imputation for nonresponse in sample surveys.*: John Wiley, New York.
- RUTKOWSKI, L.; GONZALES, E.; JONCAS, M. & VON DAVIER, M. (2010): International large-Scale Assessment Data: issues in Secondary Analysis and Reporting. *Educational Researcher*, vol. 39, No. 2, pp. 142–151.
- SJOBERT, S. (2012): PISA: Politics, fundamental problems and intriguing results. English manuscript. *La Revue, Recherches en Education*. No. 14. http://www.uhr.no/documents/6b_Sjobert_PISA_English_La_Revue_no_20.03..pdf. Letöltés ideje: 2015. június 7.
- SAHLBERG, P. (2008). *Real Learning First: Accountability in a Knowledge Society.* Paper prepared for the Alberta Teachers' Association conference on Leadership in Educational. Accountability: Sustaining Professional Learning and Innovation in Alberta Schools, Edmonton, Alberta, Canada, April 18-19. www.pasisahlberg.com/downloads/Real%20Learning%20First%20FINAL.pdf Letöltés ideje: 2015. április 15.

WUTTKE, J. (2007): Uncertainties and Bias in PISA. In: PISA zufolge PISA – PISA According to PISA. Halt PISA, was es verspricht? Does PISA Keep What it Promises? Ed.: HOPMANN, S.T.; BRINEK, G. & RETZL, M. Reihe

Schulpadagogik und Pädagogische Psychologie, Bd.6. Lit-Verlag, Wien.

YAMAMOTO, K. (2002): *Estimating PISA students on the IALS prose literacy scale*. Educational Testing Service. Citeseer.