

## KÖNYVEKRŐL

**Koretz, Daniel: *Measuring Up: What Educational Testing Really Tells Us***  
*Harvard University Press, Cambridge, M.A. 2008. 353 o.*

Az iskolák, iskolarendszerek eredményességének megítéléséhez, elszámoltatásához egyre több fejlett országban használják a tanulói teljesítményteszteket. A nemzetközi és az egyes államok tesztelési rendszereinek mérési eredményei összehasonlíthatóvá teszik és leírást adnak az oktatás, az iskolák, valamint a tanulók teljesítményéről. Bár egyre több állam dolgoz ki saját tesztrendszert, valamint az eredmények iránt fokozottan növekszik a társadalmi érdeklődés, az adatok helyes értelmezése és a következtetések megfogalmazása alapvető tesztelméleti ismeretek nélkül felszínes és félrevezető. A tesztelési programok empirikus elemzése napjainkban számos kutatás tárgya, azonban igen kevés azon munkák száma, amelyek a tesztelés elméleti alapvetéseivel foglalkoznak.

*Daniel Koretz Measuring Up: What Educational Testing Really Tells Us.* című könyve átfogó ismertetést nyújt a tesztelés különböző területeiről. A könyv megírását egyrészt az Amerikai Egyesült Államokban – az oktatáspolitikai sarokkövévé vált – jelentős következményekkel járó tesztelések körüli bizonytalanság, és a nyilvánosság hiányos ismeretei, másrészt a témában az egyetemen tartott kurzusa inspirálta. Tekintettel arra, hogy a tesztelés hatása egyre jelentősebb az iskoláztatásra, és az oktatásról szóló társadalmi vitát is tematizálja, szükség van arra, hogy minél szélesebb körben ismertté váljanak a tesztelés jellemzői. A könyv a tesztelmélet alapfogalmait a matematikai statisztikai apparátussal nem rendelkező olvasó számára is érhetően, a mindennapi életből – nemcsak az oktatás területéről – felvonultatott példák sokaságával mutatja be: mire használhatók a tesztek, mikor érdemes alkalmazni azokat, milyen korlátokkal bírnak, s hogyan használhatók a tanulók tudásának feltárására.

A kötet 13 fejezetből áll. Alapos betekintést nyújt a tesztelés problémakörébe; egyes fejezetek egymásba fonódnak, számos visszautalás segíti az információk összekapcsolását, ugyanakkor azok különálló fejezetként is egységet alkotnak. A tárgyalt témák sorra veszik a tesztelés egyes lépéseit a célok meghatározásától az eredmények interpretálásáig. Az egyes témák leírása mellett minden esetben felhívja a figyelmet a leggyakoribb hibákra, hibaforrásokra. Rávilágít a tesztelés lehetőségeire és elkerülhetetlen korlátaira, amelyeket megkülönböztet a tesztelési hibáktól és az eredmények torzításától.

Az első fejezet – *If Only It Were So Simple* – illusztrálja, hogy a tesztek milyen nagy mértékben jelen van az emberek mindennapjaiban, milyen meghatározó szerepet töltenek be életükben. Egy magasabb iskolafokozatba való lépés, egy jogosítvány, nyelvvizsga, egy állás megszerzésének feltétele, hogy megfelelő pontszámot érjünk el egy teszten. A tesztekkel olyan értékes információkat lehet megtudni, amelyek más eszközökkel gyakran elérhetetlenek. Ugyanakkor ezek az információk csak akkor érhetőek el, ha az egyes tesztelési programokat körültekintően tervezik meg és működtetik, valamint a kapott adatokat megfelelően interpretálják és használják fel. Az Amerikai Egyesült Államokban működő szövetségi és állami szintű mérések eredményeinek összehasonlításán keresztül mutatja be a szerző, hogy a tesztelés, annak ellenére, hogy széles körben használatos, nagyon bonyolult, alapos körültekintést igénylő, komplex folyamat.

A második fejezet – *What Is a Test* – a mintavételi alapelveket mutatja be egy újság számára a megfelelő munkatárs kiválasztására alkalmas teszt elemeinek meghatározásán keresztül. Rávilágít arra, hogy a tesztfeladatsorok jellemzően nem fedik le – hiszen nincs is rá szükség – az elvárt tudás- és képességelemek egészét, hanem ezek egy részahalmazát tartalmazzák, s ezen részahalmaz vizsgálatán keresztül vonhatók le következtetések az alaphalmazra vonatkozóan. Ebből adódóan meghatározó a tesztbe bevont tudás- és képességelemek ki-

választása és ezek pontos megfogalmazása annak érdekében, hogy a tesztekből a várt információkat valóban ki lehessen nyerni. Ebben a részben ismerteti a standardizált tesztek jellemzőit is.

A harmadik fejezet – *What We Measure: Just How Good is The Sample* – bemutatja, hogy az oktatás összetett célrendszerének nagy része lefedhető standardizált tesztekkel, bizonyos elemei viszont nehezen tesztelhetők, ilyen például a tudásvágy, vagy a tudás hasznosítása a munkaerőpiacon. Az Amerikai Egyesült Államokban a teljesítménytesztek fejlesztésében és elterjesztésében jelentős szerepet játszó kutató, *Lingvist* 1951-ben megjelent munkáján keresztül írja le a standardizált tesztek oktatási jelentőségét és korlátait. A felidézett munka előre jelezte azokat a polémákat, amelyek a teszteléssel kapcsolatban az Amerikai Egyesült Államok jelenlegi oktatáspolitikáját tematizálják.

A negyedik fejezetben – *The Evolution of American Testing* – az Amerikai Egyesült Államok tesztelési céljainak és eszközeinek időbeli változását foglalja össze az ötvenes évektől napjainkig. Ismerteti, hogyan jutottak el a diagnosztikus tesztelésről, az egyes tanulók értékeléséről a rendszer értékeléséig, az elszámoltatás mint elsődleges cél megjelenéséig; valamint a normaorientált értékeléstől a minimum-kompetencia teszteken keresztül a kritériumorientált értékelésig. Képet kapunk arról is, a törvényhozás és a létrehozott értékelési programok miként hatottak a tesztelésre és az oktatási folyamatra (*measurement-driven instruction*). A fejezet a 2001-ben elfogadott No Child Left Behind elnevezésű oktatási törvény bemutatásával zárul. Az ötödik fejezet – *What Test Score Tell Us about American Kids* – szorosan összefonódik az előzővel, segít értelmezni több, az Amerikai Egyesült Államok és a nemzetközi szervezetek által működtetett tesztelési program szolgáltatta információkat, a tanulók teljesítményében a hatvanas évektől megmutatkozó fejlődési tendenciákat.

A hatodik fejezetben – *What Influences Test Scores or How Not to Pick a School* – sor kerül a tesztpontszámok értelmezési lehetőségeinek illusztrálására. A pontszámok jól leírják a tanulók, tudás- és képességszintjét, ezek alapján, a felállított teljesítményrangsorok segítségével pedig az osztályok, iskolák iskolakörzetek teljesítményei összehasonlíthatók. Annak meghatározásához, hogy mi okozza az elért teljesítményt, milyen iskolai és iskolán kívüli tényezők határozzák meg a teljesítményt, és mekkora mértékben, háttérinformációk gyűjtésére van szükség. A teljesítmény és az iskolai, iskolán kívüli tényezők közötti lehetséges oksági kapcsolatokat részletesen megismerhetjük a fejezetből.

A hetedik – *Error and Reliability: How Much We Don't Know What We're Talking About* – és a kilencedik – *Validity* – fejezetben a tesztek jóságmutatóinak (reliabilitás és validitás) legfontosabb jellemzőit foglalja össze a szerző. A matematikai statisztikában kevésbé járatos olvasók számára is könnyen értelmezhető útmutatást nyújt a mérési, mintavételi hiba meghatározására, példákon keresztül segítve a könnyebb megértést.

A teljesítmények visszajelzésének napjainkban általánosan elterjedt módszerei jelennek meg a nyolcadik fejezetben (*Reporting Performance*). Ismerteti a teljesítménystandardok és a skálák segítségével történő eredményközlés ésszerűségének okait, ugyanakkor rámutat arra, hogy ezek alkalmazásával az iskolák kevésbé motiválhatók, és túlzottan általánosít, ezáltal torzítja a teljesítmények változásáról alkotott képet.

A tizedik fejezet – *Inflated Test Scores* – tárgya a tesztpontszámok inflálódásának jelensége. Fontos kérdésnek tekinti, hogy valós teljesítményszintet vagy változást tükröznek-e a teszteken elért pontszámok. Különböző mérési programok eredményeit bemutatva igazolja, hogy a tesztpontszámok növekedése vagy nagysága mögött nem minden esetben a tanulók jobb teljesítménye áll, különösen félrevezető lehet akkor, ha növekedése rövid időn belül túlzottan nagy mértékű. Ennek oka lehet a tesztek eredményeihez kapcsolódó negatív következmények elkerülése érdekében tett elkerülő magatartás, mint a tesztre tanítás, vagy a csalás okozta pontszámtorzulások.

A tizenegyedik fejezetben – *Adverse Impact and Bias* – a szerző értelmezi és elkülöníti egymástól a torzítás, a kedvezőtlen hatás és a mérési hiba fogalmakat, mint a teljesítményben megmutatkozó különbségek lehetséges magyarázó tényezőit. Hangsúlyt helyez annak bemutatására, hogy a tesztek eredménye nagymértékben függ a tesztet kitöltők közötti – például nyelvi, kulturális, gazdasági, nemi – különbségektől, amelyek ha elkerülnek a tesztkészítők és az eredményeket értelmezők figyelmét, a teszt validitása jelentősen sérül.

## Könyvekről

A tizenkettedik fejezetben – *Testing Students with Special Needs* – a sajátos nevelési igényű tanulók (a segítséggel élők és az angolt nem anyanyelvi szinten beszélők) tesztelésének lehetőségeit és szükségességét tárgyalja, tekintettel arra, hogy mindkét csoport tagjainak száma jelentős, az utóbbiba pedig egyre többen tartoznak. Ezek a gyerekek ma gyakran kimaradnak az értékelésekből, vagy ugyanolyan feltételek mellett töltenek ki tesztek és ugyanolyan szempontok szerint értékelik őket, mint többségi társaikat, így a teszteken elért eredményeik félrevezetően alacsonyak. Sorra veszi a tesztelés nehézségeit, illetve fejlesztésének lehetőségeit (pl. Braille-írással, nagyobb betűmérettel elkészített tesztváltozatok, a mozgásban akadályozott tanulók szóbeli válaszadásának és a válaszok elektronikus rögzítésének lehetősége).

Az utolsó fejezetben – *Sensible Uses of Tests* – összefoglalja azokat a veszélyforrásokat, amelyek a tesztelés nem megfelelő működtetésében és jellemzőiben keresendők. Emellett kiemeli a tesztek oktatásban betöltött elvitathatatlan szerepét, fontosságát a tanítási-tanulási folyamat eredményességének javítása, az oktatási rendszerben zajló folyamatok megismerése érdekében.

A könyv keletkezésének okai és a benne felsorakoztatott példák is igazolják, hogy az Amerikai Egyesült Államokban a méréseknek sokkal hosszabb története és fejlettebb kultúrája van, mint Európában. Már a hatvanas évek végétől rendszeresen végeztek az összes államra kiterjedő teszteléseket. Ugyanakkor túlzásokra, torzulásokra is több példát lehet találni mint az Európa egyes államaiban vagy a nemzetközi szervezetek által vezetett mérési programokban. Ez összefüggésben állhat azzal, hogy az Amerikai Egyesült Államokban a tesztek olyan célok kizárólagos eszközei, mint a közoktatás elszámoltatása. *Koretz* úgy véli, a tesztelés egy hatékony elszámoltathatósági program alapvető elemének kell lennie, de egyedüli alkalmazása helytelen lehet.

A könyv a nyilvánosság széles köréhez szól. Törvényhozók, kormányzati hivatalok dolgozói, tanárok és intézményvezetők számára ad átfogó képet a tesztelésről. Segítheti őket a pozitívumok megismerésében, átvételében, a lehetséges buktatók elkerülésében. Ugyanakkor a kötetet haszonnal olvashatják szülők és munkaadók, akik fontosnak tekintik a tesztek eredményein keresztül is nyomon követni, hogyan változik az oktatási rendszer és a tanulók, a potenciális munkavállalók teljesítménye.

*Tóth Edit*