

Trevor G. Bond

A tesztelés sokféle szerepe

Példák a pedagógiai értékelés nemzetközi gyakorlatából

A tanulói teljesítmények szummatív értékelési formáit a különböző oktatási rendszerekben számos célra alkalmazzák. Az eredményeket több országban az iskolák összehasonlítására is felhasználják.

Az azonban már ritkább, hogy a mérések eredményeit az oktatás tervezésében, a tantervek, tankönyvek készítésének segítésére is alkalmazzák.

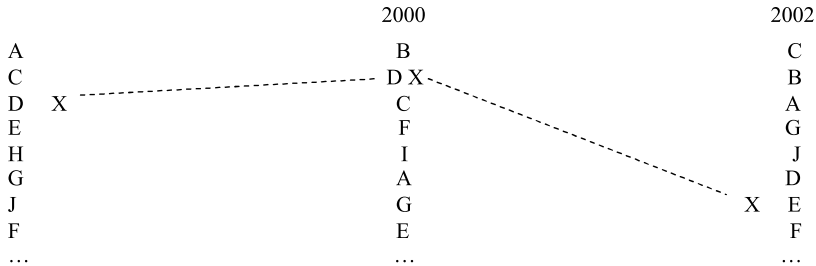
Bár a tesztek többsége közvetlenül a tantervre épül, a diákok alacsony teljesítményszintjét általában a színvonalatlan tanításnak és tanulásnak tulajdonítják, ahelyett, hogy ezek fényében a tantervi követelményeket, az elvárt teljesítmények szintjét és a tananyag elrendezésének sorrendjét vagy esetleg a tanítás módszereit vizsgálnák felül. Számos esetben iskolai vagy kerületi szintre vetítik ki a pontszámok eloszlását. Különösen Nagy-Britanniára jellemző, hogy erre egy valóságos külső tesztelési ipar épült ki, amely főleg az iskolák hatásának kimutatására szakosodott. Oregonban ezzel szemben a teljes iskolarendszert átfogó szisztematikus értékelés működik, amely hosszútávon követi az egyes tanulók fejlődését. Ez lehetőséget nyújt a különböző adatok közötti kapcsolatok sokoldalú elemzésére, melynek révén a diákok fejlődése külső standardokhoz is viszonyítható, de egyben a megelőző mérés óta hozzáadott értékkel is összevethető. A következőkben azt mutatom be, hogy a jó tesztelési eljárás központi szerepet játszik a pedagógiai hozzáadott érték mérésében, valamint az oktatás és tantervfejlesztés segítésében.

Az oktatás során rendszerint szummatív értékelési formákat alkalmazunk, többnyire azzal a céllal, hogy egyes tanulók tudását mérjük fel. A mérési eljárás végén a tanuló jegyet kap, amiből az következik, hogy az értékelés végcélja a teljesítményről szóló bizonyítvány kiállítása, azt igazolandó, hogy az oktatási folyamatban milyen konkrét szintet értek el egyes diákok. Tanulmányom alapvetése az, hogy a jó tesztelés központi jelentőséggel bír a pedagógiai értékelés folyamatában – és nemcsak azért, mert a teljesítmény mérésének kritériuma a megfelelő színvonalú felmérés. Az értékelésnek ezt a részét általában jól végezzük. A mérés azonban azért is fontos lenne, hogy megállapíthassuk, milyen mértékben járult hozzá az iskolai oktatás a tanuló tudásához. Az osztálytermi oktatás segítésében a tesztelésnek lényegi szerepet kellene játszania, de meglátásom szerint ez nem mindig van így. Sajnálatos módon a tantervfejlesztés az oktatásnak az az aspektusa, amely szinte soha nem veszi figyelembe a mérések eredményeit.

Némely oktatási rendszerre sajnos az jellemző, hogy azokat a felméréseket, melyek eredeti célja a diákok tudásának értékelése, szándékosan, vagy fel sem ismerve az alapvető problémákat, az iskolák közötti összehasonlításokra használnak fel. Az eredmények közvetlenül az egyes tanulókat jellemzik, majd ezek összesítése alapján végzik el az iskolák vagy iskolai körzetek összehasonlítását. Természetesen csak a tanulók tudásának felmérése önmagában nem nyújt hiteles képet az iskolában folyó munka színvonaláról. Ez sok helyen már annyira közismert, hogy az ilyen adatokkal nem lehet igazán érvelni, és más elemzési lehetőségek után néznek. Nagy-Britanniában ez a gyakorlat egy teljes, a hivatalos mérési rendszeren kívül álló oktatási mérésipar kiépülését eredményezte. Számos iskola és helyi oktatási hivatal súlyos összegeket fizet a rendelkezésükre álló költségvetésből a független vállalkozások által végzett oktatási felmérésekért. A függetlenek-

kel való együttműködés célja, hogy megállapítsák a helyi oktatás által hozzáadott komponenst. A külső mérések alapján, különböző összefüggéseket felhasználva a tanulók várható teljesítményszintjeire vonatkozó előrejelzéseket fogalmazzanak meg. Az előre jelzett kimenetek és a diákok által ténylegesen elért eredmények közti különbségben látják annak indikációját, hogy az egyes iskolák és az ott dolgozó tanárok milyen többlet értéket nyújtottak az elmúlt év során.

Igen sajnálatos az, hogy bár ma már kifinomult eszközök állnak rendelkezésünkre, sok országban az iskolák, iskolai körzetek és megyék összehasonlítására pusztán az iskolai mérésekből származó jegyeket vagy egyéb egyszerű adatokat használják fel. Dolgozatomban hangsúlyozom, hogy az eredmények összeszámolása önmagában még nem jelent oktatási mérést. Például a hozzáférhető adatokat felhasználva, három évre vetítve összehasonlítjuk a legjobb iskolákban tanulók teljesítményszintjeit (ld. az 1. ábrán), hogy aztán megvizsgáljuk X iskola eredményességét. 1998-ban az eredmények azt mutatták, hogy X iskola elég eredményes volt a többi jó iskolához képest. 2000-ben még annál is valamelyest jobban produkált, ám 2002-ben a többi iskolával összevetve már határozottan visszaesett. Egyértelmű, hogy ennek mind az iskolai vezetés, mind a tanárok számára konzekvenciái voltak.



1. ábra. X iskola rangsora a többi jó iskolához képest (1998-2002)

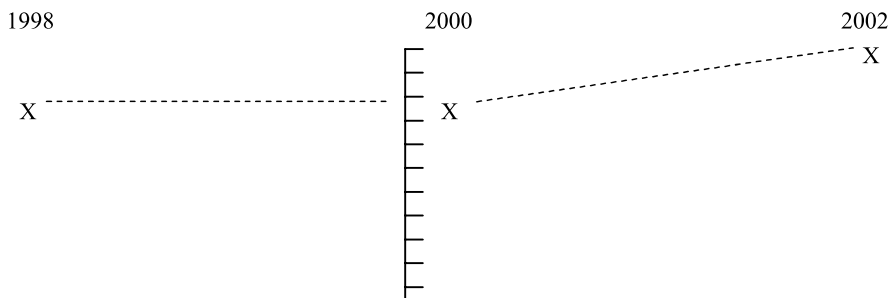
A mérés célja

Ezzel ellentétben az a folyamat, amely az USA Oregon államában zajlott le. Az oktatási hatóságok az elmúlt több mint húsz év alatt valamennyi kulcsfontosságú tanulási területen kialakították az objektív mérések teljes rendszerét. Ez ma már a matematika, a nyelvek és a természettudományok minden területére kiterjed. Nemcsak hogy számba veszik a tanulók felmérésre adott válaszait, és bizonyítványt adtak – mondván: „Bob 47 százalékot ért el matematikából”, „Mary 67 százalékot ért el nyelvekből”, és „Az iskolai átlag természettudományokból 54,3 százalék volt” -, hanem valamennyi diák, iskola és oktatási körzet teljesítményeit egységes és összehasonlítható mércéken mutatják be. Erre az úgynevezett RITS alapú, lineáris intervallum szintű mérési skálát alkalmazzák. A RITS a Rasch uniTS szavak összevonásából keletkezett megnevezés, ezt a skálát alkalmazzák valamennyi hivatalos oktatási teszteredményre. Így Bob ideai matematikai eredménye lehet 215 RITS, két évvel később pedig 245 – azaz két év alatt Bob 30 RITS egységnyit fejlődött a matematika terén. Fontos megjegyeznünk, hogy itt nem csupán a 30 matematikai RITS különbségnek van jelentősége, hanem annak is, hogy konkrétan ez a 215 és 245 közötti változást tükrözi. Az állami tantervek és a mérés dokumentumai ugyanis kvalitatív leírásokat is tartalmaznak arról, milyen matematikai készségek kialakulása kell ahhoz, hogy valaki ebben a konkrét sávban 30 értékkel növelje a RITS skálán mért eredményeit.

Pedagógiai, fejlődési perspektívából szemlélve a kérdést hallatlanul fontos, hogy ne csak az aktuális oktatási teljesítményre figyeljünk, hanem egy adott idő alatt bekövetkezett változásokra, a teljesítmények növekedésére is. A fejlesztésben érdekeltnek számára

ugyanis a növekedés az oktatás hatásának, eredményességének a leglényegesebb mutatója. Ebben a vonatkozásban az oregoni mérési rendszert számos módon lehet alkalmazni. Először is általa az egyes diákok teljesítményét összevethetjük az állami standarddal az adott tanévben – megmutatva, hogy hol áll az elvart szinthez viszonyítva. Másodsorban pedig felhasználhatjuk annak jellemzésére is, hogy az adott tanuló mennyit fejlődött az eltelt időszakban. Amennyiben az oktatási mérés ezen elvét vesszük alapul, és visszatérünk az előző példában szereplő iskolák teljesítményéhez, sokkal helytállóbb összehasonlításokat tehetünk X iskola adott időszakra jellemző oktatási teljesítményeiről.

A 2. ábra azt mutatja, hogy a RITS használatával az X iskola teljesítménypontjait egy lineáris skálán lehet megjeleníteni. Ez szinte úgy működik, mint egy vonalzó. Ha kidolgozunk egy ilyen skálát, azt mindenütt egységesen lehet alkalmazni. Ez az, ami lényegesen megkülönbözteti a Rasch-modell alapján készült valós mérési rendszereket bármely más értékelési formától. Sajnálatos, hogy bár ez a lehetőség a pszichológiában - s így az oktatásban is - rendelkezésünkre áll, az értékelés gyakorlata többnyire még mindig az 1930-as években kidolgozott alapelveken nyugszik. Úgy tűnik, túl korán elégedtünk meg kevésbé rigorózus adatelemzési formákkal – mivel alábecsültük a társadalomtudósok képességeit a szigorúbb mérési standardoknak megfelelő adatelemzési módszerek kidolgozására. Georg Rasch az 1960-as évek elején vetette meg az alapját az e tanulmány kereteit jelentő mérési elveknek.



2. ábra. X iskola eredményei (1998-2002) mérési skálára vetítve

A 2. ábrán látható, hogy 2000-ben X iskola teljesítménye 1998 óta stabil volt, s hogy 2002-ben ez a teljesítmény a skálán valójában növekedett – a társiskolákhoz képest az 1. ábrán jelölt relatív visszaesés ellenére. Vegyünk csak egy hasonló példát a fizikai mérések rendszeréből. Az, hogy milyen magas egy gyermek, nem azoktól a gyermekektől függ, akikkel együtt őt megmérték, hanem annak a mércének az osztásaitól, amellyel a gyermeket megmérték. A testmagasságot centivel kell mérni, és ez igaz az oktatási megméréstetésre is: azt a célt kell kitűznünk, hogy az iskolai teljesítményt ne csupán standard, de egyben egységes mércével is vizsgáljuk. Nem pusztán más iskolák sorrendjével vessük össze, hanem az önmaga korábbi állapotát is objektív skálán fejezzük ki. Természetes, hogy a többi iskola is megvizsgálható a mércével, de amit az oktatási értékelés révén kell az iskolának és bármely tanulójának nyújtania, az az, hogy ezt a teljesítmény standard módon mérje fel, és elsősorban azt legyen képes kimutatni, ez a teljesítmény miképpen változott a mérce szerint, nem csupán azt, hogy az miképpen viszonyul a többi tanulóhoz vagy a többi iskolához.

A vonalzó elvét követve elmondhatjuk: így nemcsak azt tudjuk megállapítani, hogy a 2002-es mérésben szereplők teljesítménye mennyiben volt jobb a 2000-esben szereplőkéitől, hanem azt is, hogy átlagban milyen jellegű volt ez a javulás. Ezután megnézhetjük a tantervi dokumentációban, mely elemek és készségek vannak e két szint között, mivel ezek jelentik azt a teljesítményt, melyet a tanulók – tanáraik segítségével – elérték 2002-ben. A valós mérések így a változás mennyiségének és típusának megragadására egyaránt alkalmasak.

A szükséges óvatosság

Az eddig előadottak azonban nem jelentik azt, hogy rendelkezésre állna egyetlen oktatási teljesítményskála, mely átfogja az iskolázás összes szintjét az általános iskola első osztályától egészen az egyetemi felvételig. Az oregoni oktatási mérési skála mind terjedelmét, mind mélységét tekintve limitált. Kritikusai megegyeznek abban, hogy ez a rendszer túlzottan a felidézésre koncentrál, és nem veszi eléggé figyelembe a tanulás általánosabb értékeit és eredményeit. Az oregoni rendszer, ami az általa mért tudástípusokat illeti, felszínesnek is tűnhet, annak ellenére, hogy megalkotóik hosszú ideje elkötelezettek a valós oktatási mérés mellett.

Az OECD nemzetközi mérési rendszerét, a PISA-t azonban már nem lehet kritikai szavakkal illetni amiatt, hogy pusztán felszínes oktatási kimeneteket felmérését végezné. A mérés tartalmát a kognitív tudományok legkorszerűbb eredményei alapján tervezték meg. Ugyanakkor a mi nézőpontunkból a PISA mérési rendszerét ugyanaz a minden részletre kiterjedő figyelem jellemzi, amellyel a valós oktatási mérés lineáris skálázási rendszerét megalkották. Ez lehetővé teszi a különálló tantárgyi skálák alkalmazását, és e skálák egymáshoz való viszonyítását is. A matematika, az olvasás-szövegértés és természettudományok terén elért eredményeket – és azok változásait – egymással is össze lehet hasonlítani. Ez lehetővé teszi, hogy a döntéshozók megalapozottabban értelmezzék, miben is áll az oktatás eredményessége.

A PISA-tól természetesen többet várunk el, mint egy országos oktatási mérési rendszertől: feltétel a valós mérési rendszer, mely lehetővé teszi a különböző tárgyakban elért eredmények pontos értelmezését iskolák, nyelvek, kultúrák és országok között. Az összehasonlításoknak az évek folyamán állandó kalibrációkon kell alapulniuk. A testmagasság példájánál maradva elmondhatjuk, hogy centi használatával bárhol a világon meg tudjuk mérni az emberek magasságát. Mivel mindenütt ugyanazt az egységet használjuk, nem kell ugyanazt a fizikai eszközt használnunk, és nem kell ugyanarra a helyre összehozni az embereket. Sőt, még arra sincs szükség, hogy a világ összes országát ugyanaz a csapat látogassa meg. Az amerikaiak nyugodtan használhatják saját eszközüket, mely lábban és hüvelykben mér. Még ennél is fontosabb, hogy értelmes megállapításokat tehetünk az amerikai gyermekek magasságáról akkor is, ha sosem jártunk az országban. Az állandó, nemzetközileg elfogadott mérési skála kidolgozásának igénye teszi a PISA-t és a hozzá hasonló projektumokat a többiek fölött állóvá.

Hasonlóképp, egy ország vizsgarendszerét, egyetemi és főiskolai felvételi vizsgáit is lehet hosszú távon egységesen az ország sajátos oktatási értékeinek és érdekeinek megfelelően működtetni. Ma már rendelkezésre állnak a feltételek ahhoz, hogy hosszabb távon is egységesen, ugyanazzal a skálával mérjük a tanulók teljesítményeit, mégpedig úgy, hogy a különböző tudásterületek skálái is egységesek legyenek, és így az azokon elért eredményeket is összehasonlíthatjuk egymással. A tanulóknak is nagy szükségük lenne erre, hiszen így ők maguk is megtudhatnák, hogy melyik területeken jobbak, mik az erősségeik. Ma ugyanis nem kapnak objektív visszajelzéseket saját képességeiket illetően, és nem biztos, hogy továbbtanulásuk tekintetében mindig jó döntéseket hoznak. Nem egyszerű feladat az oktatás eredményeinek mérése és értelmezése – így miért is várnánk el a tanulóktól, hogy még az értékelőknél is okosabbak legyenek, és eldöntsék, melyik felvételi vizsgára jelentkezzenek? Ámha két különböző tudásterületet felmérő tesztet egyetlen mérési skálához kötnénk, lehetővé válna, hogy a diákok reálisan ítéljék meg saját tudásukat. Így azokat is, akik a saját készségeiket esetleg alacsonyabbra értékelték, azonos feltételekkel felmérve hozzásegíthetnénk a megalapozottabb döntéshez. Ez egyszerű mérési feladat, amelynek megoldása elősegíti, hogy egy ország oktatási erőforrásait célirányosan, az azokból valószínűleg legtöbbet profitálók javára használják fel.

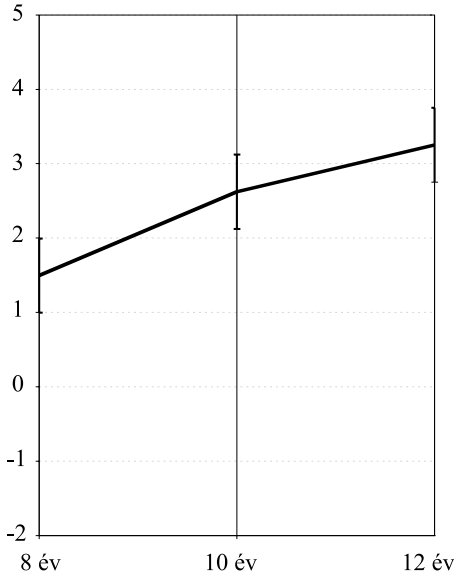
A mérések értelmezése

Hogyan használhatjuk fel értő módon az oktatási értékelési rendszerből származó mérési eredményeket? Például az sem egyértelmű, hogy az országos tesztelési rendszerre épülő angliai iskolai rangsorok táblázatai mi módon lehetnek az iskolák közötti összevetés megfelelő alapjai. Hiszen nincs szó semmiféle valós mérésről, hiányzik az egységes skála, ami nélkül nehéz az eredményeket egységesen értelmezni. Ezzel szemben viszont felvethető, hogy az oregoni iskolák összehasonlítása sem eléggé hasznos, mivel nem az oktatás eredményessége szempontjából lényeges tényezőket mérik, annak ellenére, hogy messzemenően figyelembe veszik a mérés szigorú követelményeit. Ez utóbbi esetben – az értékelés alapmodelljét megtartva – a mérések tartalmát korszerűsítve lehet a rendszert tökéletesíteni. Magyarországon *Csapó Benő* munkássága hívta fel a figyelmet arra, hogy a kérdések milyen széles körére lehet szükség a pedagógiai mérések folyamán. Például kutatócsoportjával számos olyan felmérést végzett, amellyel felszínre hozta a különböző iskolák diákjai eredményessége közti szisztematikus különbségeket, és ezeket az iskolák eltérő feltételeivel vetette össze. (*Csapó*, 2001 augusztus; 2002) Különös jelentőséggel bír, hogy Magyarországon sokkal nagyobbak az egyes iskolai osztályok közti különbségek, mint az osztályokon belüliek. Ezek olyan eredmények, amelyeket utóbb a PISA felmérések is megerősítettek. Egy demokratikus oktatási rendszer keretei között el kell gondolkoznunk azon, hogy még az azonos városban lévő egyes iskolák közt is jelentősek a különbségek – hiszen ezeket az egyenlő oktatási esélyekkel kapcsolatos elképzelések tükrében kell értelmeznünk.

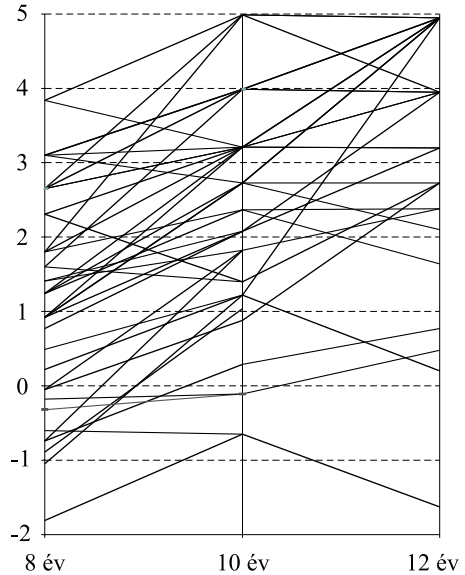
Nézzünk most egy példát a saját kutatásunkból, amelyet a James Cook Egyetemen munkacsoportjában végeztünk. (*Bond*, 1998; *Endler – Bond*, 2001; 2002) Megvizsgáltuk, hogyan fejlődik a fiatalok gondolkodása egy időszak alatt. Azt találtuk, hogy Endlernek a középiskolai diákok kognitív fejlődésére vonatkozó megállapítása összecseng azzal, amit *Piaget* várt volna el, bár *Piaget* nem ilyen jellegű vizsgálati technikát használt. A 3. ábrán látható, hogy a középiskola kezdetén (8. évfolyam) a diákok gondolkodása általában eléri a fejlett konkrét műveleti szakaszt (IIB). Két évre rá, 14 éves korban, a teljesítményszintek középpértéke a korai formális műveleti gondolkodás jegyeit mutatja (IIIA). Mire a fiatalok elvégzik az iskolát, átlagos képességük a fejlett formális műveleti szinten áll (IIIB). Az ábra akár kézikönyv-illusztráció is lehetne – és még nagyobb a jelentősége azáltal, hogy valós mérési skála az alapja.

A 4. ábra nem az osztályok átlageredményeit mutatja be, hanem az egyéni teljesítmények összességét. Az adatok Queenslandből származnak, egy, a középiskola öt éve alatt végbemenő kognitív fejlődés felméréséből. Az ábra e jellemzőt illetően az egyes mérési periódusok során az iskolai populáció nagymértékű különbségeit mutatja. Figyelembe véve a tanulók közötti különbségeket, valószínűleg nagyon nehéz lehet ennek az osztálynak a középiskolában természettudományos tárgyakat tanítani. A James Cook Egyetemen végzett kutatás középpontjában (*Bond*, 2001; 2002b; *Grobecker – Bond*, 1999; *Noelting – Coudé – Rousseau és mtsai.*, 2000) az értelmi fejlődés és az oktatás eredményessége közötti összefüggések feltérképezése állt. Az egyik fő célja az volt, hogy a mérés segítségével megmutassa a tanárnak, milyen szinten állnak a tanulói, mit lehet nekik megtanítani, és mi az, aminek a befogadására még nincsenek felkészülve. Az ilyen jellegű felmérés közvetlen segítséget nyújthat az órai munka tervezéséhez, az egyénre szabott tanári munkához, hiszen a tanároknak épp az a szerepük, hogy nagyobb szakmai tudásukkal és a tanulókra vonatkozó gazdagabb információk alapján segítsék a diákok tanulását.

Az üzenet minden bizonnyal világos azok számára, akik az oktatási mérést azért végzik, hogy a tanároknak nyújtsanak segítséget. Valami olyasmit mondanak, ami a tanárokat abban segíti, hogy jobban megértsék tanítványaik tanulási gondjait, és így a diákoknak pontosabb, célirányosabb segítséget nyújthassanak, jól megválaszthassák az alkalma-



3. ábra. Átlagos kognitív fejlődésszintek a középiskolában



4. ábra. A kognitív fejlődés egyéni profiljai középiskolában

zott fejlesztési hatásokat. Gondoljuk azonban meg, mi mindenre van szükség ahhoz, hogy a pedagógiai mérések valóban a gyakorló tanárok érdekeit szolgálják. Például ha csak a következő tanévben kapják meg az eredményeket a tanárok, az már keveset segít a konkrét munkában. Ha az eredmények vajmi keveset árulnak el arról, mit kellene a tanárnak tennie annak érdekében, hogy a tanulók előmenetelét javítsa, akkor nem sokra jutunk vele. E cél érdekében az oktatási mérésnek diagnosztikusnak is kell lennie, hogy a tanárok olyan információhoz jussanak az értékelés által, amelyet felhasználhatnak tanításuk tervezésében – a következő hét, hónap tervezésében, nemcsak a következő tanévben. Nem kétséges, hogy meg kell adni ezt az információt a tanároknak, hogy az jobban áthathassa a napi gyakorlatot. Ha egy félév közepén kiadott teszt eredményei nem jutnak el az iskolába a tanév végére, az értékelés semmilyen hatással nem lehet a tanórai munkára.

Ha a tesztelés valós mérésen alapul, és nem merül ki pusztán számításban és statisztikában, ha a tesztelés valóban a dolgok megértésére és alkalmazására, nemcsak a felidézésre összpontosít, olyan visszajelzést adhatunk a tanároknak, amelyet már a következő heti munkában felhasználhatnak. Az 5. ábra a diagnosztikai visszajelzésre ad példát. Az 1. számú tanuló 40 százalékot ért el a teszten, amely a negyedik osztályos tanulók matematikai fogalomértését mérte. A tanulói térkép bal oldalán azok a tesztitemek (az itemek sorszámai) láthatók, melyekre a tanuló helyesen válaszolt, a térkép jobb oldalán pedig azok, melyeken hibázott. Az egyes itemek nehézségi szintjét az elhelyezkedésük jelzi a logit intervallum-skálán a térkép közepén. A tanuló e teszttel mért készségét az „xxx” hely jelöli ki a térképen. Az értékek elhelyezkedését kiegészíti felette és alatta a hibásáv, a becsléssel kapcsolatos mérési hiba függvényében. A tanuló matematikai teszteredménye jó példáját adja fejlettségi szintjének. Minden olyan itemet helyesen oldott meg, amely könnyebb volt, mint a saját készségszintje, de egyetlen olyat sem tudott megoldani, amely nehezebb volt, mint a becsült készségszintje ezen a teszten. A tanuló teljesítménye (a teszt itemeire vetítve) szinte tökéletes vonalban reprezentálja a mérési elvnek való megfelelést. Tanárát nem aggasztanák a kisebb eltérések ott, ahol a tanuló képességeihez közeli két item váratlanul helyes (10.3) illetve hibás (7.2).

lódési hiányosság megnyilvánulásának vagyunk itt tanúi, amelynek messzemenő következményei lehetnek a jövőben, amikor a tanulónak ezeket a korábbi ismereteket kellene alkalmaznia bonyolultabb feladatok megoldásában. (Bond – Fox, 2001, 179–183.) A viszajelzésnek ez a formája nagy haszonnal járhat a tanár számára, és kijelölheti azt a módot, amellyel a tanítási munka során segíthet a probléma megoldásában. Ha a tesztelés nem jár ilyen haszonnal a diák tanulása számára, miért is fordítunk mi, az oktatási mérés elkötelezettjei olyan sokat rá?

A tantervfejlesztés segítése

Mi következik mindebből? Az, hogy a választott megközelítés lehetővé teszi, hogy az oktatási mérés végső és valószínűleg legfontosabb kérdésére választ adjunk. A mai gyakorlat azt mutatja, hogy amikor az iskolákkal tudatják, hogy diákjaik képességei az államilag meghatározott szint alatt maradnak, a tipikus válaszok általában a következők: „Pocsék a tanítás!”, „Pocsék a tanulás!”, „Érdektelenek a szülők!” és „Alacsony színvonalú az iskola!” Alig hallani olyan reakciót, hogy „Nem lehet, hogy az államilag jóváhagyott tantervvel van valami gond?”, illetve „Lehet, hogy a tanterv nem állja meg a helyét?” Megkérdőjelezzük-e valaha a legnagyobb hatalommal rendelkező oktatási hivatalok által kiadott tantervi dokumentumok érvényességét? Felteszük-e a kérdést, hogy egy tantervi változtatást milyen konkrét felmérés igazol? Úgy tűnik, a tanulók, a tanárok és az iskolák értékelése a lényeges, és szinte sosem merül fel a tanterv egyes elemeinek értékelése. Az itt bemutatott mérési elvek felhasználásával képessé válhatunk arra, hogy a tanterven magán is elgondolkodjunk, különösen, ami annak tartalmát és sorrendiségét illeti.

Az Oregonban kidolgozott mérési rendszer eredeti célja, legalábbis részben, az volt, hogy a közoktatási rendszer tantervi követelményeinek vizsgálatát szolgálja. Sajnálatos, hogy a hivatalos tesztelésre vonatkozó állami és szövetségi követelmények a projektum fókuszává a tanulók tesztelését tette, lemondva a tantervfejlesztésről. Amikor Oregonban évente nyilvánosságra hozzák az állami mérések eredményeit, ilyen kérdéseket tesznek fel: Melyek a legjobb iskolák? Mi okozta a csökkenést egyes iskolákban az államilag megállapított szintekhez képest? Ritkán, nagyon elvétve lehet csak olyanokat hallani, hogy a tantervi követelmények megfelelőek-e terjedelmüket és sorrendiségüket illetően.

Ausztráliában, Queenslandben szerencsénk volt: az általános iskolai matematika-tanterv a fejlődési elvekre épült már attól az időtől fogva, hogy Jean Piaget munkássága közvetlen hatást gyakorolt a nyugati világ osztálytermeire. Igaz, a tantervfejlesztés eredeti céljai a későbbi generációk számára már a múlt homályába vesznek, a matematikai tantervi követelmények általános fejlődési elvei alapvetően ma is helytállóak. Ha ezt a megközelítést komolyan vesszük, mint ahogy ebben a tanulmányban is, és felméri a tanulók kognitív fejlődését, továbbá ennek eredményét a közvetlenül a matematikai tanterve épülő teszteredményekkel vetik össze, megállapítható, hogy a queenslandi általános iskolai matematika-tanterv nagyrészt a megfelelő sorrendet követi. Gyakran találkozunk azonban azzal a problémával, hogy számos iskolában korábbra teszik a tanterv bizonyos elemeit. Így például az első években a formális matematika több követelményének nem tud megfelelni az iskolát kezdő tanulók jelentős száma – mivel nem érték el az azokhoz szükséges kognitív fejlettséget. A sorrendiséget illetően tehát helyes úton járunk, de további munkát igényel a tartalom finomítása. (Bond, 2001)

A diák fejlődési szakasza és a tantervi követelmények közti különbségek ezzel természetesen nem merülnek ki. Ami a matematikai és természettudományos teljesítményt illeti – mely területek kognitív fejlődési kutatásunk elsődleges fókuszának számítanak (Bond, 2002a; 2002b) – e különbségek végigkísérik a középiskolai oktatást a tanulók jelentős hányadánál. Az eredményekből látható, hogy az oktatási következmény ezen tanulók számára jobb esetben a magolásból származó szerény siker lehet, vagy rossz esetben

a meg nem értés miatti sikertelenség – ami elfogadhatatlan kimenet bármely olyan rendszer számára, mely komolyan veszi a társadalmi igazságosságba és az oktatás demokrázizmusába vetett hitet.

Összegezve

Elgondolkodhatunk most azon, mi teszi jóvá a tesztelést, és hogy mi az oktatás lényege, hogy lássuk, fel tudjuk-e vállalni a jelenleg a kettő között meglévő szakadék áthidalását. Az oktatási mérést manapság az jellemzi, hogy képes-e egy adott napon a tanuló a megállapított standard elérésére – és ez által a tanulókat az „elég jók” és a „még nem elég jók” csoportjaiba sorolja: a sikeres és sikertelen vizsgázókéra. Az oktatás azonban inkább hasonlítható egy lépcső megmászásához, a tanárral az oldalunkon, aki minden egyes tanulót segít a fejlődésben. A mérésnek egy kérdésre kell válaszolnia: „Mennyit haladt felfelé a tanuló a lépcsőn?” Az oktatás dinamikus folyamat, sosem lehet statikus. Ebben az értelemben a tesztelésnek az oktatást segítő értékelési folyamat részévé kell válnia. A mérést arra kell használnunk, hogy kiteljesítse az oktatás körét. Most a tanulóról adunk információkat. Arra is szükség van azonban, hogy ne csak a tanuló teljesítményének megismerésére használjuk a tesztelést, hanem hogy az eredményeket a tanterv értékelésében is kamatoztassuk. Jelenleg a tanterv az, amelynek alapján a tanulók képességeit felmérő tesztek kidolgozzák. A kör kiteljesítéséhez elengedhetetlen, hogy a tanulókról szerzett információkat a tantervre alkalmazzuk. E tanulmány végső üzenete az, hogy ezt csak jó teszteléssel érhetjük el – olyannal, amely valós mérési elvekre épül.

Irodalom

- Bond, T. G. (1998): Fifty years of formal operational research: The empirical evidence. *Archives de Psychologie*, 66. 217–234.
- Bond, T. G. (2001): Ready for school? Ready for learning? An empirical contribution to a perennial debate. *The Australian Educational and Developmental Psychologist*, 18. 1. 77–80.
- Bond, T. G. (2002a): *Changing minds and empirical evidence: Can we measure change*. XI. International Thinking Conference, Harrogate, Egyesült Királyság.
- Bond, T. G. (2002b): *Relationships between cognitive development and school achievement: A Rasch measurement approach*. Annual Conference of the Australian Association for Research in Education, Brisbane, Ausztrália.
- Bond, T. G. – Fox, C. M. (2001): *Applying the Rasch model: Fundamental measurement in the human sciences*. Mahwah, N.J.: Erlbaum.
- Csapó B. (2001): *Efficiency of schools in developing students' cognitive competencies*. IX. European Conference for Research on Learning and Instruction, Fribourg, Svájc.
- Csapó, B. (2002): Az osztályok közötti különbségek és a pedagógiai hozzáadott érték. In: Csapó, B. (Ed.): *Az iskolai műveltség*. Osiris Kiadó, Budapest. 269–297.
- Endler, L. C. – Bond, T. G. (2001): Cognitive development in a secondary science setting. *Research in Science Education*, 30. 4. 403–416.
- Endler, L. C. – Bond, T. G. (2002): *Measuring changed minds: Empirical evidence from Thinking Science interventions in Australia and USA*. XI. International Thinking Conference, Harrogate, Egyesült Királyság.
- Grobecker, B. – Bond, T. (1999): Children's construction of addition, *Archives de Psychologie*, 67. 95–122.
- Noelting, G. – Coudé, G. – Rousseau, J. P. – Bond, T. – Brunel, M. L. (2000): Can qualitative stage characteristics be revealed quantitatively? *Archives de psychologie*, 68, 259–275.