

# R. Tóth Krisztina

## PAPÍR-CERUZA ÉS SZÁMÍTÓGÉPES TESZTEK EREDMÉNYEINEK ÖSSZEHASONLÍTÓ VIZSGÁLATA

---

*Achievement differences in traditional and computer-based testing*

The transition from paper-and-pencil (PP) to computer-based (CB) assessment has gained a growing interest in the last decades. This shift may lead to achievement differences. Several analyses did not find significant differences between students' PP and CB performance, however, a number of studies report substantial differences across media. This paper provides an overview about the nature and importance of CB testing. Then, the most important results of the first Hungarian pilot study exploring the medium effect are presented. The PP inductive reasoning test created by Csapó (1994) was administered to a sample of 5<sup>th</sup> graders (N=843). Few weeks later, the same items along with questionnaires for background variables in an online form were administered by means of the TAO platform. Analysis of the data showed that students performed significantly better at the PP test compared to the CB one. On the other hand, the subtest-level analysis highlighted various achievement differences e.g. the multiple choice items indicated better results in a CB environment and the subtests with open-ended items seemed easier in PP format. This paper presents the achievement differences based on background variables between genders and the relationship between computer use and standardized test scores.

### BEVEZETÉS

Az elmúlt évtizedekben számos rendszerszintű pedagógiai vizsgálatot (pl.: PIRLS, TIMSS, OECD-PISA, Országos Kompetenciamérés stb.) végeztek mind nemzetközi, mind hazai szinten, melyek adatai pontos képet adtak a tanulók tudásáról. E rendszer-, illetve iskolaszintű kutatási eredmények oktatáspolitikai döntések alapjául szolgálhatnak, de csak kis mértékben támogatják a tanárok mindennapi munkáját. Ezért kerültek az osztálytermi folyamatok, az oktatás mikroszintje az utóbbi években a pedagógiai értékeléssel foglalkozó kutatások figyelmének középpontjába (Csapó, 2008).

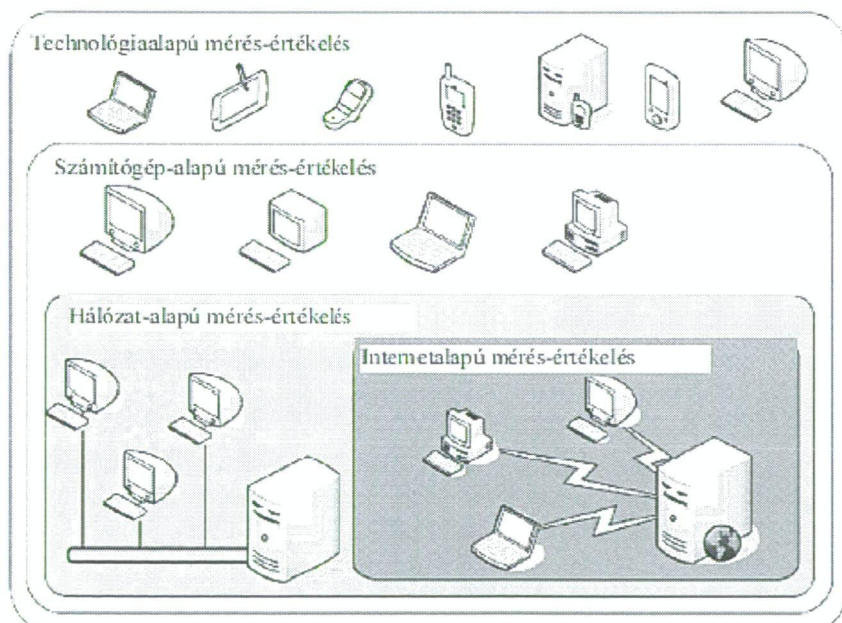
A tanítási-tanulási folyamat hatékonyságának növelését, a tanulók egyéni fejlődését leginkább a diagnosztikus és formatív értékelésből származó információk visszacsatolása segítheti, amely a tanulók évenként többszöri felmérését követeli meg. A gyakori tesztelés megvalósításához azonban a papír-ceruza (angolul: paper-and-pencil – PP) tesztek már nem nyújtanak megfelelő keretet, mert a mérés-értékelés-visszajelentés folyamata idő- és költségigényes. Ezért a tesztelési folyamat új tesztmédiára adaptálása elkerülhetetlennek tűnik, és ebben az információs és kommunikációs technológiák (IKT), a gyors, nagy kapacitású számítógépek központi szerephez jutnak. Az IKT eszközök nyújtotta előnyök mellett az Internet adta le-

hetőségek hasznosítása is új lehetőségeket teremt a pedagógiai mérés-értékelés területén, a tanulók tudásának felmérésben, hiányosságaik feltárásában, ezáltal a tanítási-tanulási folyamat hatékonyabbá tételében (R. Tóth, Molnár és Csapó, 2008).

## SZÁMÍTÓGÉPES TESZTELÉS ÉS JELENTŐSÉGE<sup>1</sup>

### A számítógépes tesztelés

Az IKT eszközök pedagógiai mérési-értékelési folyamatba integrálásával megvalósul a *technológialapú* (vagy más néven elektronikus) *mérés-értékelés* (angolul: technology-based assessment, e-testing/assessment), ahol a tesztelés mobil eszközökön (pl.: PDA, szavazógység stb.) vagy számítógépek különböző fajtáin keresztül zajlik. Az elektronikus tesztmédiумok közül a számítógép a jelenleg legelterjedtebb IKT eszköz. A *számítógépes* vagy *számítógép-alapú tesztelés* során (angolul: computerised, com-



1. ábra. Az elektronikus tesztelési formák hierarchikus viszonya (idézte Csapó, Molnár és R. Tóth, 2008)

<sup>1</sup> A fejezet alapját a Csapó, Molnár és R. Tóth (2008) tanulmány alkotja.

puter-based – CB – testing) a feladatlap vagy megoldandó item a számítógép monitorán jelenik meg (on-screen presentation), és a tesztelt személy például billentyűzet vagy egér használatával adja meg válaszait (Jurecka és Hartig, 2007).

A számítógépek különálló (stand alone) munkaállomásokként, offline módon, vagy hálózati kapcsolaton keresztül (bálozat-alapú mérés-értékelés) is megvalósíthatják a tesztek kiosztását, a tanulói válaszok rögzítését, kiértékelését és az eredmények visszacsatolását. A számítógépek közötti kapcsolatot egy helyi hálózat, illetve a világháló (Internet-alapú vagy online tesztelés, angolul: Internet-based, web-based assessment, online testing) is megteremtheti. Ezen elektronikus tesztelési formák hierarchikus viszonyát mutatja be az I. ábra.

## A számítógépes tesztelés előnyei

A mérési-értékelési folyamatok számítógépre adaptálását elsősorban az alábbi technológia nyújtotta előnyök ösztönözték:

1. *Tanulói teljesítmények azonnali visszacsatolása:* lehetővé vált több ezer tanuló adott időben történő felmérése, eredményeik azonnali visszajelentése.
2. *Információgazdag értékelés megvalósítása:* a CB tesztelés bővebb információval szolgál a tesztelés folyamatáról, mint a tradicionális tesztelés: mérhetjük az egyes itemek megoldásához szükséges időt, feltérképezhetjük hányszor javított a diák a megoldásán stb., így minden tesztelt személyről egyénileg megállapíthatjuk, mely feladat(ok) megoldása okozott leginkább nehézséget számára.
3. *Innovatív itemtípusok alkalmazása:* az új tesztmédiium lehetővé teszi az innovatív – multimédiás elemekkel, hiperlinkekkel gazdagított – itemtípusok alkalmazását, melyekkel bővül a pedagógiai tesztekkel mérhető területek száma. Például az elektronikus szövegek hatékony olvasásához a tanulónak képesnek kell lennie explicit és beágyazott hiperlinkek alkalmazására, és jártasnak kell lennie a navigációs eszközök (görgetősáv, menük) használatában (Mendelovits, Lumley és McCrae, 2007 alapján), melyek ismerete csak CB környezetben mérhető.
4. *Bővül a tesztelésbe bevonható tanulók köre:* a CB teszteléssel a „látási, hallási és a kézírás készségével kapcsolatos problémák jó része kiküszöbölhető” (Kárpáti, 2002. 8. o.), így a fogyatékkal élő tanulókat is bevonhatjuk a tesztelésbe. A kézírás problémáit megszüntethetjük egyrészt a fejegér (head tracker) alkalmazásával, „amely révén a képernyőn lévő egérkurzort pusztán fejmozgatással lehet irányítani” (HVG, 2008), másrészt a szemegér (eye tracker) használatával, „amely pusztán a szemmozgás érzékelésével és követésével váltja ki az egérmozgást” (HVG, 2008).

5. Gyorsabb, pontosabb, ugyanakkor költséghatékonyabb tesztvégrehajtás valósulhat meg a hagyományos, PP tesztelés mellett, illetve helyett. *Rose, Hess, Hörhold, Brähler és Klapp* (1999) eredményei alapján a CB teszteléssel a tesztelés dokumentációs költségének 2/3-át meg lehet takarítani. Az eredmények kiértékelése a teszt befejezését követő percekben megtörténik, és minimalizálható az emberi figyelmetlenség miatt bekövetkező kiértékelési hiba, mely Bucher (1987. 17. o., idézi *Becker*, 2004) alapján az esetek 10%-ában fordul elő.
6. Személyre szóló tesztelés: a számítógép lehetőséget kínál az egyéni teljesítményen alapuló, különböző személyre szóló tesztelési formák megvalósítására, melyek PP-alapú tesztekkel egyáltalán nem, vagy csak nagyon költségesen kivitelezhetők. Ilyenek például az adaptív tesztelés, a többlépcsős (multistage) tesztelés, továbbá a pedagógiai játékok (educational serious games) használatával történő tudás-, illetve képességmérés (l. *R. Tóth*, 2009).

## Számítógépes tesztelés nemzetközi viszonylatban

A CB tesztelés előnyeit, a technológia nyújtotta lehetőségeket kiaknázva számos oktató-kutató intézet/szervezet bevezette, vagy tervezi bevezetni a számítógépes tesztelést a PP tesztek helyett, illetve mellett (*R. Tóth, Molnár és Csapó*, 2009). Az OECD-PISA 2006-ban a számítógép-alapú természettudományi tudás (Computer-Based Assessment of Science) és 2009-ben az elektronikus szövegolvasás (Electronic Reading Assessment) részteszttel megkezdte a PP tesztekéről az elektronikus tesztelésre történő áttéréssel kapcsolatos próbaméréseket.

Az Egyesült Államok több állama elkezdte átültetni értékelési rendszerének komponenseit online formátumra (*Bennett*, 2003). Az Educational Testing Service (ETS) a PP alapú *Test of English as a Foreign Language* (TOEFL) nemzetközi nyelvvizsga teszt mellett 2005 szeptemberében elindította a nyelvvizsga internetes változatát, a *TOEFL Internet-based Test*-et (*ETS*, 2007).

Európában, ugyancsak elektronikus környezetben tesztelnek Norvégiában, ahol az OECD-PISA 2000 mérésen nyújtott eredmények hatására bevezetett országos tesztelést az oktatáskutatóknak lehetőség szerint CB formában kellett megoldaniuk (*Moe*, 2008). A holland CITO<sup>2</sup> is fejleszt számítógépes és webes tesztek különböző célcsoportok számára. Luxemburgban, 2007-ben egy iskolamonиторozási (school monitoring) program keretében 2700–3500 tanulót teszteltek le 7 nap alatt

---

<sup>2</sup> [http://www.cito.com/research\\_and\\_development/computer\\_webbased\\_testing.aspx](http://www.cito.com/research_and_development/computer_webbased_testing.aspx), 2008. 07. 28-i megtekintés

(Martin, Latour, Burton, Busana, Keller, Reichert, Plichart, Jadoul és Swietlik, 2007) a TAO<sup>3</sup> (Testing Assisté par Ordinateur – Computer-Based Testing) platformmal.

Magyarországon szintén számítógépen történik a vizsgáztatás a jogosítvány megszerzéséhez szükséges KRESZ vizsgán (*Informatika és Tudomány*, 2004), valamint a Szegedi Tudományegyetem Oktatásméлет Kutatócsoportja által 2008 tavaszán végzett próbaméréssel elkezdődött a magyarországi, nagymintás, elektronikus tesztelés bevezetése is.

## A TESZTMÉDIUM HATÁSA A TESZTEREDMÉNYEKRE

Amennyiben a PP tesztelésről egy másik tesztmédiumba (pl.: számítógépre) térünk át, empirikus adatokkal kell alátámasztani, hogy a médium megváltoztatása nem gyakorol hatást a teszteredményre. Ezért össze kell hasonlítani a két médiumon (PP és CB) elért eredményeket, amit több nemzetközi ajánlás is megfogalmaz. Például a *Standards for Educational and Psychological Testing* (1999) empirikus megerősítés elkészítését javasolja, hogy minden kétséget kizáró bizonyíték álljon rendelkezésre a tesztelés különböző formáiból nyert eredmények felcserélhetőségére (*Fitzpatrick és Triscari*, 2005).

Ha az új tesztmédiumba a mérés-értékelés eredményeire mégis hatást gyakorol, meg kell határozni: (a) a két tesztmédiumba elért teljesítmények különbségének mértékét, (b) az eltérés irányát, (c) az előnyösen vagy hátrányosan érintett tanulócsoporthoz (k)at, illetve (d) azokat a faktorokat (függő változókat), amelyek az eltérésért felelősek. Kutatásmódszertani szempontból a lehetséges függő változók három dimenzió mentén csoportosíthatók: a minta, a teszt és a tesztmédiumba (számítógép) jellemzői mentén. Ezek közül a tesztelési folyamat számítógépre adaptálásakor leggyakrabban vizsgált tényezők a tesztet kitöltő diákok számítógép-használatbeli jártassága, a monitorról történő olvasás sebessége, illetve a billentyűzet használata.

A különböző változók mentén végzett összehasonlító vizsgálatok eltérő eredményeket állapítanak meg. Egyes elemzések szerint nincs, vagy nem minden esetben van jelentős különbség a két médiumon elért teljesítmények között (*Hargreaves, Shorrocks-Taylor, Swinnerton, Tait és Threllfall*, 2004; *Bodmann és Robinson*, 2004). Mások viszont statisztikailag szignifikáns eltérést mutatnak ki a PP és CB teszteredmények között. Léteznek tanulmányok (*Pommerich*, 2004; *Way, Davis és Fitzpatrick*, 2006), melyek a diákok tradicionális teszteken elért jobb teljesítményét

<sup>3</sup> www.tao.lu

támasztják alá, ezzel szemben egyéb publikációk (*Clariana és Wallace, 2002; Pomplun, Frey és Becker, 2002*) eredményei szerint az elektronikus tesztek jelentősen könnyebbnek bizonyulnak.

## TESZTEREDMÉNYEK EMPIRIKUS VIZSGÁLATA<sup>4</sup>

### Kutatási kérdések és a minta

Egy, az SZTE Oktatásméleti Kutatócsoport által alkalmazni kívánt elektronikus tesztelőrendszer bevezetéséhez – a nemzetközi ajánlásokat követve – meg kell vizsgálnunk, hogy ugyanazon teszt PP tesztelésből származó és CB formátumban felvett eredményei megegyeznek-e. Ennek vizsgálatára 2008 tavaszán próbamérést végeztünk, mely során a következőkre kerestük a választ:

- (1) megállapítható-e szignifikáns különbség a tanulók induktív gondolkodását mérő teszt PP és CB környezetben elért teljesítménye között,
- (2) amennyiben a kimutatható különbség realizálódik, milyen mértékű és irányú eltérés mutatkozik a két tesztmediumon az összpontszám és a részeredmények tekintetében, illetve,
- (3) a kutatásban vizsgált háttérváltozók milyen hatást gyakorolnak a teszteredményekre. Jelen tanulmányban a nem és az IKT használat változókat vontuk be az adatok elemzésébe.

A mintát ( $n=843$ ) a Kutatócsoport által végzett longitudinális vizsgálatban résztvevő 7 régió, 24 település, 26 iskolájának 5. osztályos tanulóiból képeztük. A minta 53%-át fiúk alkották.

### Mérőeszközök és adatfelvétel

A mérés fókuszában a *Csapó Benő* által készített, 58 itemből álló induktív gondolkodás teszt (*Csapó, 1994*) állt. A teszt időkorlátos volt, 35 perc állt a tanulók rendelkezésére a feladatok megválaszolására. A teszt három résztesztet tartalmaz: a számok analógiája, szóbeli analógiák, és számsorok résztesztet. A számok analógiája és a számsorok részteszt nyílt végű, feleletalkotó feladatokból, a szóbeli analógiák részteszt pedig zárt, feleletválasztó itemekből állt. A PP formátumban bemért tesztet digitalizáltuk, ügyelve arra, hogy a tesztkérdések formája ne változ-

---

<sup>4</sup> E fejezet a *Csapó, Molnár és R. Tóth (2009)* publikáció eredményei alapján készült.

zon. A 2. ábra példát mutat ugyanazon item megjelentésére tradicionális és online környezetben.

Papír-ceruza itemek	Számítógépes itemek
a) 20→32 :: 8→20 :: 11→_____	a) 20 → 32 :: 8 → 20 :: 11 → _____
a) SZÉK : BÜTOR = KUTYA : ?	SZÉK : BÜTOR = KUTYA : ?
1 MACSKA	<input type="radio"/> MACSKA
2 ÁLLAT	<input type="radio"/> ÁLLAT
3 TACSKÓ	<input type="radio"/> TACSKÓ
4 RÓKA	<input type="radio"/> RÓKA
5 KUTYAÓL	<input type="radio"/> KUTYAÓL

2. ábra. Itemek papír-ceruza és számítógépes formátumban  
(idézve Csapó, Molnár és R. Tóth, 2009)

A 2. ábra illusztrálja, hogy ugyanazon item megoldása más-más tevékenységet várt el a tanulóktól a két közvetítő médiumon. Míg papíron karikázni kellett a helyesnek vélt megoldást, addig online kattintani kellett. A nyílt végű feladatoknál papíron egy számot írtak a diákok a vonalra, számítógépen a billentyűzet segítségével adták meg válaszaikat.

A komparatív vizsgálatban minden tanuló kétszer írta meg az induktív gondolkodás tesztet. Először a longitudinális mérés keretében (2008 május) a papír-ceruza feladatlapot töltötték ki. Majd néhány héttel ezután került sor az online adatfelvételre (2008 június). Az elektronikus tesztelés a TAO szoftveren (Farcot és Latour, 2008) keresztül valósult meg.

A képességmérő teszten kívül kitöltöttek a diákok két elektronikus háttérkérdőívet is, melyekkel IKT tapasztalatukat, számítógép-használati szokásaikat, illetve a számítógépes-teszteléssel kapcsolatos attitűdjüket térképeztük fel. Szintén kérdőíves vizsgálat keretében számoltak be a tesztelés lebonyolításával kapcsolatos észrevételeikről az elektronikus tesztelést levezető tanárok.

## Eredmények

Az induktív gondolkodás teszt eredményeinek empirikus vizsgálata alapján megállapíthatjuk, hogy az alkalmazott mérőeszköz papír-alapú (Cronbach- $\alpha_{PP}=0,91$ ) és elektronikus (Cronbach- $\alpha_{CB}=0,9$ ) környezetben is megbízhatóan mér.

A komparatív elemzések magasabb átlagos teljesítményt igazoltak a hagyományos teszten ( $x_{PP}=27,2\%$ ,  $sd_{PP}=14,9\%$ ), mint a számítógépes adatfelvétel ( $x_{CB}=26,0\%$ ,  $sd_{CB}=14,0\%$ ) során. Ez a mindössze 1,2%-os eltérés szignifikáns ( $t=3,57$ ,  $p<0,01$ ).

Az összpontszámbeli eltéréseket előidéző faktorok azonosításához azonban mélyebb elemzések szükségesek, ezért a kapott eredményeket résztesztekre bontva tovább vizsgáltuk. Az 1. táblázat tartalmazza a három részteszten PP és CB alapon elért, százalékban kifejezett eredményeket, illetve zárójelben a hozzájuk tartozó szórásokat. A táblázat negyedik oszlopában feltüntetett, a két tesztmédiumon megmutatózó teljesítménykülönbségeket szignifikanciavizsgálatnak vetettük alá, mely alapján jelentős különbséget igazoltunk mindhárom részteszten. A legjelentősebb eltérés (4,38%) a számsorozatok részteszt esetén mutatkozik.

1. táblázat. A papír-ceruza és számítógépes tesztek eredményeinek részteszt szintű összehasonlítása

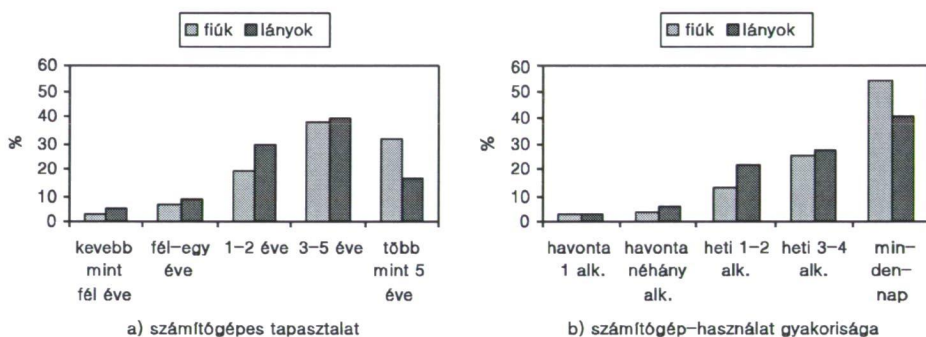
	PP teljesítmény %-ban (sd)	CB teljesítmény %-ban (sd)	Átlagok eltérése	t	p
Számanalógiák	27,0% (21,6%)	25,4% (20,3%)	-1,62	2,68	0,007
Szóanalógiák	40,3% (21,6%)	42,8% (22,2%)	2,5	-5,27	0,000
Számsorozatok	14,3% (11,5%)	9,9% (9,4%)	-4,38	10,71	0,000

Az eltérések irányát tekintve a nyílt és zárt végű feladatok esetén különböző tendencia rajzolódik ki. A nyíltvégű, feleletalkotó feladatokat tartalmazó számanalógia és számsorozat részteszteken a hagyományos teszteredmények átlagosan magasabbak, míg a zártvégű, többszörös választásos itemek megválaszolása online környezetben könnyebbnek bizonyult. Ennek oka lehet például, hogy a számsorozatok hiányzó tagjainak meghatározása, illetve a számanalógiás feladatokban a számpárok megállapításánál segítséget nyújthatott a tradicionális tesztelés során a papír, melyen közvetlenül a feladatok mellett segédszámításokat tudtak végezni a diákok. Noha az elektronikus tesztek kitöltésekor is használhattak volna psziko-



zatpapírt a mellékszámítások elvégzéséhez, a tesztelést vezető tanárok válasza alapján mégsem tették.

A háttérváltozókkal végzett elemzések közül e tanulmányban a nemek közötti eltérésekre fókuszáltunk. Az eredmények alátámasztják, hogy a lányok a tradicionális tesztelés alkalmával jobb eredményt ( $x_{pp}=27,66\%$ ,  $sd_{pp}=14,33\%$ ) értek el, mint CB-alapon ( $x_{CB}=26,04\%$ ,  $sd_{CB}=13,89\%$ ). Ugyanakkor a fiúk teljesítménye nem függött az alkalmazott tesztmédiától ( $x_{pp}=26,73\%$ ;  $x_{CB}=25,99\%$ ). A nemek közti különbség magyarázatára IKT háttérváltozókat vontunk be az elemzésbe: a számítógép-használat gyakoriságát és a számítógépes tapasztalatot.



3. ábra. A számítógép-használat és számítógépes tapasztalat gyakorisági eloszlása (R. Tóth, 2008 alapján)

A háttérkérdőív kérdéseire (pl.: „Milyen gyakran használsz számítógépet?”; „Mióta használsz számítógépet?”) adott válaszok alapján megállapítható, hogy a fiúk gyakorlottabb számítógép használók, mint a lányok (3/a ábra), hiszen a több mint öt éve computert használók fiúk (31,6%) közel kétszer annyian vannak, mint a lányok (16,6%). További eltérés mutatkozik a számítógép-használat gyakorisága között (3/b ábra), ahol a nemek közötti legjelentősebb eltérés a heti 1–2 alkalommal számítógépezők és a mindennapos felhasználók között mutatkozik meg. A mindennap számítógépezők között 13,4%-kal több fiú, míg a heti 1–2 nap számítógépezők között 8,8%-kal több lány van.

Ugyanakkor ezek a háttérváltozó valószínűleg önmagukban nem magyarázzák a nemek közötti teljesítménykülönbségeket, mert a részteszt szintű elemzések megmutatták, hogy a szóbeli analógiák és számsorozatok részteszten is ugyanaz a tendencia rajzolódik ki a két nem között. Mindkét nem a szóbeli analógiák rész-

teszten online jobb teljesítményeket mutatott, ezzel szemben a számsorozatok összefüggéseinek felismerése, és a megkezdett sorozatok hiányzó elemeinek meghatározása papíron szignifikánsan jobb teljesítményt eredményezett. A számanalógiák részteszten a fiúk teljesítménye médiafüggetlen volt, míg a lányok PP-alapon eredményesebbnek bizonyultak. A bemutatott eltérések okának felderítésére további vizsgálatok, egyéb háttérváltozók bevonása szükséges.

## KONKLÚZIÓ ÉS JÖVŐBELI TERVEK

A bemutatott pilot mérés adatai alapján megállapíthatjuk, hogy 1,2%-os különbség van ugyanazon teszt PP és CB környezetben felvett eredményei között, mely a kialakított mintán szignifikánsnak bizonyult. A vizsgálat a tanulói összteljesítményeket tekintve a papír-alapú tesztet könnyebbnek igazolta. A résztesztekre irányuló elemzések a nyílt végű feladatokat tartalmazó részteszteken magasabb PP, a zárt feladatokat tartalmazó szóbeli analógiák részteszten pedig magasabb CB eredményeket támasztottak alá. Ezek a részteszt szintű vizsgálatok valószínűsítik, hogy a teszteredmények közötti különbségek függenek az alkalmazott item sajátosságaitól: a megoldáshoz szükséges tevékenységtől és az item típusától.

A háttérelmzések rámutattak, hogy a teljes teszt szintjén a fiúk teljesítménye médiafüggetlen volt, míg a lányok a hagyományos teszten szignifikánsan jobban teljesítenek. Az eltérés okának feltárására bevont IKT háttérváltozók nem adtak egyértelmű magyarázatot a nemek közötti különbségekre, így a két tesztmédiumon elért eredménykülönbségek interpretálására vonatkozóan a jövőben további feladatok állnak előttünk. Törekednünk kell arra, hogy megismerjük azokat a változókat, amelyek a teszteredmények különbségeiért felelősek.

### Megjegyzés

A tanulmány TÁMOP 3.1.9/08/01 kutatási program, az Oktatásméleti Kutatócsoport és az MTA-SZTE Képességkutató Csoport keretében készült.

## IRODALOM

- Bennett, R. E. (2003): *Online Assessment and the Comparability of Score Meaning*. (Research Memorandum). Educational Testing Service, Princeton, NJ.
- Becker, J. (2004): *Computergestütztes Adaptives Testen (CAT) von Angst entwickelt auf der Grundlage der Item Response Theorie (IRT)*. Digitális disszertáció. Freie Universität, Berlin.
- Bodmann, S. M. és Robinson, D. H. (2004): Speed and performance differences among computer-based and paper-pencil tests. *Journal of Educational Computing Research*, 31. 1. sz. 51–60.

- Csapó, B. (2008): Integrating recent developments in educational evaluation: formative, longitudinal and online assessments. Keynote Lecture. The European Conference on Educational Research. Gothenburg, Sweden. September 8–9.
- Csapó Benő (1994): Az induktív gondolkodás fejlődése. *Magyar Pedagógia*, 94. 1–2. sz. 53–80.
- Csapó, B., Molnár, Gy. és R. Tóth, K. (2009): Comparing Paper-and-Pencil and Online Assessment of Reasoning Skills: A Pilot Study for Introducing TAO in Large-scale Assessment in Hungary. In: Scheuermann, F. és Björnsson, J. (szerk.): *The Transition to Computer-Based Assessment – New Approaches to Skills Assessment and Implications for Large-scale Testing*. Office for Official Publications of the European Communities, Luxembourg. 113–118.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papíralapú teszteléstől a számítógépes adaptív tesztelésig. A pedagógiai mérés-értékelés technikájának fejlődési tendenciái: *Iskolakultúra*, 3–4. sz. 3–16.
- Clariana, R. és Wallace, P. (2002): Paper-based versus computer-based assessment: key factors associated with the test mode effect. *British Journal of Educational Technology*, 33. 5. sz. 593–602.
- ETS (2007): Test and Score Data Summary for TOEFL Internet-Based Test. 2009. 07. 28-i megtekintés, Educational Testing Service, <http://www.ets.org/Media/Research/pdf/TOEFL-SUM-0506-iBT.pdf>
- Farcot, M. és Latour, T. (2008): An Open Source and Large-Scale Computer Based Assessment Platform: A real Winner. In: Scheuermann, F. és Pereira, A. G. (szerk.): *Towards a Research Agenda on Computer-Based Assessment. Challenges and needs for European Educational Measurement*. Office for Official Publications of the European Communities, Luxembourg. 64–67.
- Fitzpatrick, S. és Triscari, R. (2005): Comparability Studies of the Virginia Computer-Delivered Tests. Paper presented at the annual meeting of the American Educational Research Association. Montreal, Canada. April 11–15.
- Hargreaves, M., Shorrocks-Taylor, D., Swinnerton, B., Tait, K. és Threllfall, J. (2004). Computer or paper? That is the question: does the medium in which assessment questions are presented affect children's performance in mathematics? *Educational Research*, 46. 1. sz. 29–42.
- HVG (2008): HVG.hu. Világot látni egy roboton keresztül. 2009. 07. 28-i megtekintés, [http://hvg.hu/Tudomany/20080320\\_elte\\_sajtoklub\\_kornyezeti\\_intelligencia.aspx](http://hvg.hu/Tudomany/20080320_elte_sajtoklub_kornyezeti_intelligencia.aspx)
- Jurecka, A. és Hartig, J. (2007): Computer- und netzwerkbasieretes Assessment. In: Hartig, J. és Klieme, E. (szerk.): *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik*. Bundesministerium für Bildung und Forschung, Bonn, Berlin. 37–48.
- Kárpáti Andrea (2002): Informatikai „keresztanterv” – A számítógéppel segített tanítás és tanulás új paradigmája. 2009. 07. 28-i megtekintés, [www.isze.hu/download/tananyagok/Tanulmany.pdf](http://www.isze.hu/download/tananyagok/Tanulmany.pdf)
- Martin, R., Latour, T., Burton, R., Busana, G., Keller, U., Reichert, R., Plichart, P., Jadoul R., és Swietlik, J. (2007): TAO: use cases and item templates; TAO: architecture and use cases of a collaborative, internet-based platform for computer-assisted testing. Poster presentation at 12<sup>th</sup> European Association for Research on Learning and Instruction, Budapest, Hungary, August 28.–September 1.
- Mendelovits, J., Lumley, T. és McCrae, B. (2007): Constructing and scoring tasks for an electronic reading assessment. Paper presented at 2007 PMRT National Roundtable Conference on Assessment and Reporting. Adelaide, Australia, November 19–20.
- Moe, Eli (2008): Introducing large scale computer-based testing of English – experiences and future challenges. Presented at International Research Workshop, The Transition to Computer-Based Assessment, Lessons learned from the PISA 2006 Computer Based Assessment of Science (CBAS) and implications for large scale testing, Reykjavik, Iceland, September 29.–October 1.

- Pommerich, M. (2004): Developing Computerized Versions of Paper-and-Pencil Tests: Mode Effects for Passage-Based Tests. *Journal of Technology, Learning, and Assessment*, 2. 6. sz. 1–44. 2009. 07. 28-i megtekintés, *Journal of Technology, Learning, and Assessment* [on-line] <http://escholarship.bc.edu/jtla/vol2/6/>
- Pomplun, M., Frey, S., és Becker, D. F. (2002): The Score Equivalence of Paper-and-Pencil and Computerized Versions of a Speeded Test of Reading Comprehension. *Educational and Psychological Measurement*, 62. 2. sz. 337–354.
- R. Tóth Krisztina (2009): Az elektronikus tesztelés és tesztvégrehajtás fajtái. Előadás: VII. Pedagógiai Értékelési Konferencia, Szeged, 2009. április 24–25.
- R. Tóth Krisztina (2008): Az online teszteléssel kapcsolatos attitűdök és eredmények a háttérváltozók rükrében. Előadás: VIII. Országos Neveléstudományi Konferencia, Budapest, 2008. november 13–15.
- R. Tóth Krisztina, Molnár Gyöngyvér és Csapó Benő (2009): Online Assessment of Reasoning Skills. Paper presented at the 13th Biennial Conference of the European Association for Research on Learning and Instruction Amsterdam, the Netherlands, August 25–29.
- R. Tóth Krisztina, Molnár Gyöngyvér és Csapó Benő (2008): A számítógépes tesztelés lehetőségei. Előadás: VI. Pedagógiai Értékelési Konferencia. Szeged, 2008. április 11–12.
- Rose, M., Hess, V., Hörhold, M., Brähler, E. és Klapp, B. F. (1999): Mobile computergestützte psychometrische Diagnostik. Ökonomische Vorteile und Ergebnisse zur Teststabilität. *Psychotherapie Psychosomatik Medizinische Psychologie*, 49. 202–207.
- Sg.hu Informatika & Tudomány (2004): Bevezették a számítógépes KRESZ vizsgát. 2009. 07. 27-i megtekintés, Sg.hu Informatika & Tudomány, <http://www.sg.hu/cikk.php?cid=33919>
- Way, W. D., Davis, L. L., és Fitzpatrick, S. (2006): Score comparability of online and paper administrations of the Texas Assessment of Knowledge and Skills. Paper presented at the Annual Meeting of the National Council on Measurement in Education. San Francisco, CA, April. 8–10.