

Technológia-alapú mérés-értékelés hazai és nemzetközi implementációi

A gazdaság világában akármerre nézünk, mindenhol új technológia vesz minket körül – számítógépek, nyomtatók, szkennerek, pda-k, mobiltelefonok, plusz a mindezen eszközöket összekötő hálózatok (Bennett, 2002). Mindezen eszközök azonban nemcsak mindennapi életünkben, hanem a tanulás-tanítás folyamatát tekintve is számos új lehetőséget teremtenek. A 21. század digitális korában ezért elkerülhetetlen a technológia iskolai alkalmazása.

Az, ahogyan egy adott terület ismereteit, képességeit, készségeit tanítjuk és tanuljuk, nagy mértékben meghatározza, hogyan mérhetjük azt. Ez változhat és változik is az idő folyamán. Egyrészt az új technológiai eszközök elterjedése nemcsak a tanulás és tanítás folyamatában, hanem az ahhoz szorosan kapcsolódó mérés-értékelés terén is új lehetőségeket teremt és meghatározó elemmé válik, másrészt az új generációk igénye és IKT-s jártassága is változik, egyre fejlettebbé válik, a diákok mindennapi életének szerves részét képezi.

A net-generáció, a digitális bennszülöttek attitűdjei, elvárásai szignifikánsan különböznek a korábbi generációk attitűdjeitől és elvárásaitól (Molnár, 2008). A 20. század generációi és a 21. század net-generációi között generációs szakadék ('generation lap') és már nem csak – mint a korábbi generációk között – generációs ugrás ('generation gap') van. A net-nemzedékek attitűdje (Oblinger és Oblinger, 2005), életstílusa, elvárásai, tanulási felfogása és ismeretszerzési módja hatást gyakorol az oktatással és a mérés-értékeléssel kapcsolatos elvárásaikra is. A net-generáció technológiában való jártassága és csoportorientációja új lehetőséget (Oblinger és Oblinger, 2005) teremt ismereteik, képességeik fejlettsége számszerűsítésében. A megfelelően alkalmazott, fejlett technológia fokozhatja a tanítási-tanulási és mérés-értékelési folyamat hatékonyságát: a tanulók tudásának megkonstruálását, felmérését, esetleges hiányosságai feltárását (R. Tóth, Molnár és Csapó, 2008).

A technológia mérés-értékelés folyamatába történő integrálása hatékonyabb tesztelést eredményez, mert egyrészt hasonlóan a mindennapi munkavégzéshez, a mérés-értékelés folyamatában is az információ elektronikus áramlása gyorsabb és olcsóbb (Moss és Hendry, 2002), mint ha mindazt fizikailag, manipulatív módon tennénk (Negroponte, 1995 idézi Bennett, 2002), másrészt a technológia olyan új lehetőségeket, új mérési módszereket tesz lehetővé, amelyek a hagyományos technikákkal nem megvalósíthatók. Ma már senki sem vitatja, hogy a technológiai alapú tesztelés rövid vagy hosszabb távon kiszorítja a papíralapú tesztelést, ezzel forradalmasítva a mérés-értékelés célját és lehetőségeit is.

Nem kétséges tehát, hogy a multimédiás alkalmazások új megvilágításba helyezik a tanulás és az azzal szoros kapcsolatban lévő értékelés folyamatát, és lehetőséget nyújtanak az információ, illetve a tudás új, innovatív módon történő integrálására (Molnár, 2007). A modern IKT-eszközök, webes alkalmazások új lehetőséget teremtenek az oktatás színvonalának és a mindennapi tanítási, értékelési gyakorlat minőségének növelésére.

Ezeknek a tevékenységeknek a megtanítására az oktatási rendszernek jól képzett tanároknak van szüksége, akik szakértői az információs és kommunikációs technológiáknak (Molnár, 2008), akik hatékonyan hozzájutnak információkhoz, akik ezt hatékonyan tovább tudják adni a diákoknak, és akik ismerik az új tanítási és értékelési módszereket. A tanárképzés intézményeinek kell a vezető szerepet betöltenie és a kulcsszerepet vállalnia a tanítási-tanulási-értékelési folyamat hatékonyságának növelésében az IKT-alapú tanárképzés által (R. Tóth és Molnár, 2009).

A technológia-alapú mérés-értékelés meghatározása és jelentősége

A technológia mérés-értékelésben betöltött szerepe sokrétű lehet, alkalmazásának relevanciája függ a vizsgált területtől, az adott területen a diákok technológiahasználati szokásaitól csakúgy, mint a feladatírók, adatelemzők IKT-s kompetenciáitól. Ebből adódóan nincs egységes definíció arra vonatkozólag, hogy mit jelent a technológia-alapú mérés-értékelés. Attól függően, hogy a mérés-értékelés folyamatának melyik stádiumában, szintjén vagy szintjein jelenik meg, továbbá a folyamatban részt vevők közül kik és mire használják az adott eszközt – aminek milyensége szintén definiálól erővel bírhat –, különböző definíciókat fogalmazhatunk meg.

A hagyományos papír-ceruza adatfelvétellel történő vizsgálatokban is már nélkülözhetetlen szerepet tölt be a számítógép. Általában azon történik a feladatok végső formába öntése, az adatok rögzítése, elemzése, a visszajelzések elkészítése stb. Ennek ellenére ezeket a méréseket a továbbiakban nem tekintjük technológia-alapúnak. Technológia-alapú mérés-értékelésről abban az esetben beszélünk, ha maga az adatfelvétel során is és nemcsak előtte, illetve utána kap fontos szerepet a technológia.

Az adatfelvétel tekintetében öt általános helyzetet különböztethetünk meg aszerint, hogy milyen területen, helyzetben milyen relevanciával alkalmazhatjuk a különböző eszközöket az adatfelvétel során. Ezekben belül további esetek definiálhatók annak függvényében, hogy mi az a konkrét technológia, eszköz, amit alkalmazunk, és azt hogyan, milyen tesztelési módszerrel tesszük.

Az első esetben a mérésben részt vevő diákok olyan technológiai eszközöket használnak az adatfelvétel során, amiket más esetben, az iskolában, tanórán is alkalmaznak az adott területen kapott feladat vagy probléma megoldása során. Például egy matematika teszt esetén természetes számukra a tanórán is gyakran használt számológép vagy különböző grafikai ábrázoló programok, esetleg Excel táblázatok használata, még akkor is, ha a feladatmegoldás folyamata papíralapon történik. Ezek az eszközök kötődnek számukra a matematikai jellegű problémák megoldásához, ugyanakkor általában más helyzetekben nem használják őket. Ebben az esetben, bár az adatok rögzítésében nem, de az adatfelvétel során mégis fontos szerepet játszik a technológia (Csapó, Latour, Bennett, Ainley és Law, 2009).

A második esethez olyan területek vizsgálata, olyan helyzetek tartoznak, ahol nem jellemző, vagy nagyon ritka, hogy a diákok technológiai eszközt használnak. Ilyen például az esszéírás, amit az iskolában a diákok általában papíralapon végeznek. Ebben az esetben, ha egy diák nem használ számítógépet, annak kezelése, az azon folytatott munka – a konkrét példánál maradván a mondatszerkesztés és gépelés – számára nehézkes lehet, ami negatívan befolyásolhatja eredményét. Aki többet ír billentyűzetten, mint papíron, előnyben van, nemcsak a gépelés gyorsasága, hanem a mondatformálás terén is. Ebben az esetben a papíron és számítógépen írt teszt, esszé nem összehasonlítható (Powers és Potenza, 1996). Ezekben az esetekben háromféle módon járhatunk el:

– minden egyes diákot hagyományos papír-ceruza módon tesztelünk, hogy lássuk, ott milyen hatékonyak;

– minden egyes diákot technológia-alapon tesztlünk, hogy lássuk, az adott területen milyen hatékonyak az adott eszköz használatában;

– minden diákot azon a módon tesztlünk, ahol rutinosabb, amit megszokott (*Horkay, Bennett, Allen, Kaplan és Yan, 2006, idézi Csapó, Latour, Bennett, Ainley és Law, 2009*).

A harmadik helyzet azokra a területekre jellemző, ahol nélkülözhetetlen az adott technológiai eszköz használata, anélkül értelmetlenné válna a terület mérése. Például számítógép nélkül nem lehet hatékonyan tanítani a programozást. Ezekben az esetekben a mérés-értékelés során is kézenfekvő az adott eszköz használata (*Bennett, Persky, Weiss és Jenkins, 2007; Bennett, Jenkins, Persky és Weiss, 2003*).

A negyedik típusú szituációban a vizsgálat célja lehet annak számszerűsítése, hogy a diákok képesek-e az adott feladat, probléma megoldásában jobb teljesítményt elérni, ha egy technológiai eszközt a rendelkezésükre bocsátunk, mint anélkül. Ha egy új, a diákok számára kevésbé ismert és használt eszközt adunk részükre, akkor opcióként adhatjuk azt a lehetőséget is, hogy nem veszik igénybe az adott eszközt és anélkül oldják meg a feladatot (lásd például: *Molnár, 2008*). Az e témakörbe sorolható kutatások a technológiai eszközök tanulást és tanítást befolyásoló szerepét vizsgálják (lásd például: minden diák és tanár számára biztosítanak 1–1 laptopot, lásd: *Bebell és Kay, 2010; Suhr, Hernandez, Grimes és Warschauer, 2010; Bebell és O'Dwyer, 2010*).

Nemzetközi szinten Európában Luxemburg vezeti a mezőnyt, ahol megtörtént már a papírceruza tesztekéről a számítógépalapú tesztelésre való átállás, sőt bevezettek egy országos online mérésre alkalmas rendszert. A rendszer jelenlegi verziója egyszerűre képes tesztelni a teljes populációt, majd a teszteredményekről automatikus visszajelzést készít a tanárok részére. Ezt a rendszert adaptálta többek között az OECD PISA 2009 Electronic Reading Assessment (ERA – számítógépes olvasás-felmérés), illetve az OECD PIAAC (Program for International Assessment of Adult Competencie, felnőtt kompetenciák felmérése) is.

Az ötödik esetben a technológia a diákok közötti kollaborációt és a hatékonyabb tudásépítést és tudásmegosztást támogatja. Mind a tudásépítés, mind a tudásmegosztás nélkülözhetetlen eszköze a technológia, miután az hatékonyan csak közösségi szinten történő aktív kommunikációval lehetséges (*Scardamalia, 2002; Lipponen, 2002*). A számítógéppel támogatott kollaboratív tanulás, azaz az együttműködő tanulás információs és kommunikációs technológia segítségével történő támogatása számos hazai és nemzetközi kutatást indukált (például: *Molnár P., 2009; Dorner, 2007; Molnár és Kárpáti, 2009; Neo és Neo, 2001; O'Malley,*

1995; *Macdonald, 2003; Kyllonen, 2009*).

A technológia szerepe a mérés-értékelés hatékonyságának növelésében

A mérés-értékelés egész folyamatában – a feladatírástól az eredmények visszajelzéséig – szerepet játszhat a technológia. Három esetet különböztetünk meg aszerint, hogy a technológia alkalmazása befolyásolja-e a tesztelés validitását, azaz ugyanazt mérjük-e papír-, mint technológia-alapon:

Ideális esetben (tévesztés, elírás, elütés faktorát kizárva) a tesztelés eredményét és tárgyát, a tesztelés validitását nem befolyásolja a technológia, ha azt az adatfelvétel kivételével bármely más esetben alkalmazzuk. Például

- feladatírás, tesztfejlesztés során,
- automatikus vagy félautomatikus itemgenerálásra,
- a feladatok megosztására, bírálatára, javítására (*Bejar, Lawless, Morley, Wagner, Bennett és Revuelta, 2003*), vagy
- az eredmények rögzítésére, kódolására, elemzésére,
- a visszajelzések elkészítésére és visszajuttatására, amivel nyomtatási és szállítási költséget is megtakarítunk.

Papír-ceruza tesztekhez hasonló tesztek felvétele történik technológia-alapon, azaz már az adatfelvétel, az adatok rögzítése során is alkalmazzuk a technológiát. A technológiát használhatjuk például

- a tesztek diákokhoz való eljuttatására, amivel jelentős nyomtatási, tárolási és szállítási költséget takarítunk meg. Ebbe az esetben sorolható még, ha
- digitális audió- vagy videófájlokat osztunk meg a diákokkal – gyakori tesztelési mód idegen nyelvi mérések során, amikor a diákok szövegértését méri –, amelyeket korábban magnón vagy videón keresztül hallgattak rosszabb minőségben a diákok.

Hasonlóképpen ide sorolható a diákok beszédkézségének mérése, amelyet korábban magnóval vagy videomagnóval rögzítettek (*Bennett, Goodman, Hessinger, Ligget, Marshall, Kahn és Zack, 1999*). Ennek költsége a fent nevezett okok miatt szintén magasabb, mintha a számítógép segítségével digitálisan rögzítjük a diákok beszédét. Mindkét esetben a jobb minőség alacsonyabb előállítási költséggel párosul. Előbbi esetben nem kell kazettákra másolni a hanganyagot, elegendő azt on-line eljuttatni az iskolákhoz, utóbbi esetben akár párhuzamosan is történhet az adatfelvétel és a teszteléshez nem szükséges külön eszköz fizikai szállítása.

Megfelelő tesztelési módszer (például feleletalkotó itemek alkalmazása, időmérés) alkalmazása mellett még az sem szükséges, hogy a tesztelt személyek fizikailag egy helyen legyenek, amivel szintén jelentős költségmegtakarítást érhetünk el (*Zhang, Powers, Wright és Morgan, 2003*).

A technológia lehetőséget ad a válaszok azonnali és automatikus kiértékelésére, amivel jelentős emberi munkát takarítunk meg (*Williamson, Mislevy és Bejar, 2006*), illetve felgyorsítjuk a visszajelzés folyamatát, ami szorosan összefügg a tesztelés hatékonyságával.

A harmadik esetben szintén az adatfelvétel során is szerepet játszik a technológia, azonban a mért konstruktum ebben az esetben már különbözik attól, amit papír-ceruza alapon mérhetünk, mérnénk, azaz jelentősen befolyásolja azt, amit mérünk. Ebben az esetben alkalmazhatunk olyan

interaktív szimulációkat, kísérleteket, amiket technológiai eszköz nélkül nem tudnának elvégezni a diákok, majd kérhetjük azok leírását, vagy azokkal kapcsolatos feladatok, problémák megoldását (*Bennett, Persky, Weiss és Jenkins, 2007*).

Ide sorolható az a féle tesztelés is, amikor a diákok tanulnak a tesztelés során. Ebben az esetben dinamikus a feladatadás, azaz a feladatra, problémára adott válaszaik meghatározzák a tesztelés következő problémáját, a visszajelzésből tanulnak a diákok.

Azt, hogy a technológia milyen mértékben és hogyan javítja a mérés-értékelést, meghatározza a tesztelés területe, célja, kontextusa. Az eddigi technológia-alapú mérés-értékeléssel kapcsolatos hazai és nemzetközi kutatások főképpen a nagy téttel bíró szummatív tesztekre fókuszáltak.

A technológia-alapú tesztelés típusai az alkalmazott technológia típusának és módjának függvényében

A technológia-alapú tesztelés megvalósításának számos lehetősége ismert. A közvetítő eszköz kiválasztását számos faktorcsoporthoz befolyásolja. A megfelelő eszköz kiválasztásának kulcsfontosságú szempontja, hogy minden egyes diák számára közel azonos tesztelési feltételeket biztosítson (hasonló tulajdonságú monitorok, számítógépek, internet-csatlakozás stb.), illetve alkalmas legyen az elemzéshez megfelelő minőségű adatok begyűjtésére. A továbbiakban azonosítunk néhány faktorcsoporthoz, amik pontosabban definiálják a tesztelés során alkalmazandó médiumot:

A teszt tartalma. Ha a teszt a papír-ceruza tesztekhez hasonló feleletválasztásos (vagy egyszerű feleletalkotó) kérdéseket tartalmaz, ahol a válaszok rádiógombok segítségével, azokra kattintva begyűjthetők, relatív alacsony igény fogalmazódik meg a közvetítő eszközzel szemben. Ha a teszt gazdag hanganyaggal, grafika- és videó-elemekkel rendelkezik, továbbá autentikus környezetben különböző szoftver alkalmazásokat használ, akkor a közvetítő eszközzel szemben támasztott igény már jóval magasabb.

Az iskolák internet-kapacitása. Milyen sávszélességgel, milyen minőségben, gyorsasággal tud az iskola vagy a tesztelést végző intézet kapcsolódni az internethez. Ebben az esetben figyelembe kell venni, hogy a kapcsolat a tesztelés során szimultán, és a tesztelésben részt vevő összes diák számára azonos feltételeket kell biztosítani. Ennek megvalósítását több tényező is nehezíti és befolyásolja: egyrészt az iskolán belüli egyéb internethasználat és annak típusa, másrészt az iskolán kívüli konkurens internetforgalom. Mindkettő jelentősen lassíthatja a tesztelés során az internet gyorsaságát, ezért érdemes a követelményekre szigorúbb feltételeket szabni.

A tesztelés során használt iskolai számítógépek. Az iskolai gépek száma (ha szükséges, csak azok, amelyek csatlakoznak a világhálózathoz), azok típusa, tulajdonságai (például: gyorsaság, memória, grafikus kapacitás – ha szükséges a tesztelés során), illetve a periférikus eszközök tulajdonságai (például monitor nagysága, felbontása) tartoznak ebbe a csoportba. A tesztelés sikerességét szintén befolyásolhatja a gépeken futó operációs rendszer és egyéb kiegészítő programok (Java, ActiveX) típusa és verziószáma (*Csapó, Latour, Bennett, Ainley és Law, 2009*).

A számítógép-alapú tesztelés (Computer Based Assessment – CBA) általában minden komputeres értékelést magában foglalt, kicsit tágabb értelemben használják még a technológia-alapú tesztelés (Technology Based Assessment – TBA), illetve az elektronikus tesztelés (e-Testing) kifejezéseket is. A technológia-alapú mérésen belül – elterjedtségéből fakadóan is – a legtöbb lehetőséget a számítógép-alapú értékelés rejti magában. Ebben az esetben az adatfelvétel közvetítő eszköze a számítógép, azaz a mérőeszköz (teszt, kérdőív) kérdései, feladatai, itemei a számítógép monitorján jelennek meg. A válaszok bevitele is a számítógépen keresztül történik, a billentyűzet, egér, érintőképernyő stb. segítségével, ebből adódóan azok rögtön elektronikusan rögzítésre kerülnek.

Aszerint, hogy a tesztelés platformja és mérőeszköze hogy jut el az adatfelvételt közvetlenül végző számítógépekig, a számítógépes tesztelés további típusait (helyi hálózat; internet; mobil eszköz – CD ROM, pendrive) különböztethetjük meg.

A hálózat-alapú mérés egy gyakori alkalmazása, amikor az adott hálózaton belül egyszerűen több gépen zajlik a tesztelés, a tesztelést egy külön számítógépről irányítják, ahol az adatok összegyűjtése, elemzése történik (*Csapó, Molnár és R. Tóth, 2008*). Ezt a külön, jelen esetben helyi szerverként működő gépet, notebookot, melyre a teszteléshez szükséges összes programot előzetesen installáltak, elegendő eljuttatni az iskolába, majd az iskola erőforrását, számítógépeit használva, erre csatlakozva történik a tesztelés. Ebben az esetben is két módja van a tesztelésnek. Vagy a tesztelés előtt minden egyes adatfelvételben részt vevő gépre felinstallálják a teszteléshez szükséges szoftvert, vagy a

gépek a tesztelés alatt beállított szervertől – hasonlóan az internet-alapú teszteléshez, lásd lent – csatlakozva egy böngésző segítségével érik el a szükséges programokat. Az adatok begyűjtése történhet manuálisan (CD-re, pendrive-ra írva) vagy elektronikusan (ftp-vel letölteni; *Csapó, Latour, Bennett, Ainley és Law, 2009*). A kiértékelés szoftvertől függően vagy a helyi számítógépen, vagy a központi szervertől történik (*Csapó, Molnár és R. Tóth, 2008*).

Az internet-alapú adatfelvétel irányítását egy központi szerver végzi, azaz a helyi számítógépekre nem kell felinstallálni a tesztelést végző programot. Mivel a szoftver külső gépen fut, a gépekkel szemben felállított követelmények is alacsonyabbak (*Csapó, Molnár és R. Tóth, 2008; Csapó, Latour, Bennett, Ainley és Law, 2009*), viszont megfelelő számú és minőségű internetkapcsolattal (és internetes böngészővel) rendelkező számítógépre van szükség. Az iskolában szükséges sávszélesség függ a tesztfeladatoktól és az egyszerre tesztelt diákok számától. Ha az iskolában ez nem áll rendelkezésre, a probléma megoldható vezeték nélküli (wireless) internetkapcsolat alkalmazásával is, azonban ez jelentősen megrágtja a vizsgát. Mind a kábeles, mind a vezeték nélküli megoldás esetén nélkülözhetetlen a megfelelő központi szerver alkalmazása, ami párhuzamos elérések esetén is biztosítja a tesztelés zavartalan lefolyását. A helyi hálózat, illetve internet ilyen típusú felhasználási lehetőségeiről részletesen lásd *Csapó, Molnár és R. Tóth (2008)* tanulmányát.

Ha semmilyen hálózatot nem vonunk be a tesztelés lefolytatásába, akkor a tesztelést végző programot, feladatlapot minden egyes számítógépre installálni kell, vagy egy mobil eszköz (CD-ROM, pendrive, esetleg az egész számítógép – notebook) segítségével minden egyes személyhez eljuttatni. Előbbi sok tekintetben korlátozza a tesztelést (megoldandó probléma volt például az adatok begyűjtése), utóbbi memóriatartalma azonban már lehetővé teszi, hogy az egész tesztelés lebonyolítható legyen egy pendrive segítségével. Onnan fusson a tesztelést végző program, illetve egy adatbázis-kezelő program segítségével arra mentse rá a program az adatokat. Ebben az esetben nincs szükség arra, hogy a helyi gépekre programokat installáljanak, elegendő a megfelelő USB kimenet biztosítása. A tesztelés e típusa olcsóbb, mint a papíralapú tesztelés lebonyolítása, sőt a pendrive-ok később újra felhasználhatóak (*Csapó, Latour, Bennett, Ainley és Law, 2009*).

A technológia-alapú mérés-értékelés nemzetközi implementációi

A technológia-alapú mérés-értékelésnek legnagyobb hagyománya az USA-ban van, számos nagymintás vizsgálatban alkalmaztak már számítógépet az adatfelvétel során. Általános és középiskolai szinten a legismertebb és legnagyobb online felmérések: a „Measures of Academic Progress” (MAP, Northwest Evaluation Association), a „Virginia Standards of Learning Tests” (SOL, Virginia Department of Education) és az „Oregon Assessment of Knowledge and Skills” (OAKS, Oregon Department of Education) (*Csapó, Latour, Bennett, Ainley és Law, 2009*). Mindhárom programban kizárólag feltevésválasztós itemeket alkalmaznak.

Egy hazánkban is ismert vizsga, a TOEFL (Test of English as a Foreign Language, ETS) felvétele is számítógépen zajlik. A vizsga tartalmaz esszé részt is, aminek kiértékelését egyrészt automatikusan elvégzi a számítógép (az esszé számítógépes értékeléséről lásd például: *Ben-Simon és Bennett, 2007*), másrészt egy vagy több értékelővel is javítatják. A szóbeli vizsga letételének számítógép-alapúvá tételére, annak automatizálására is találunk példát. Ebben az esetben a vizsga folyamán egy mikrofon segítségével felvesszük a vizsgázó szóbeli feleletét, majd azt különböző szoftverek segítségével elemzik.

Szintén az USA-ban nemzetileg reprezentatív és longitudinális vizsgálatosorozat a NAEP (The National Assessment of Educational Progress). A NAEP vizsgálatok keretében belül három területen történt online adatfelvétel 2001, 2002 és 2003-ban. A 2001-es

„Matematika online” (Math Online – MOL; Sandene, Horkay, Bennett, Allen, Braswell, Kaplan és Oranje, 2005) vizsgálatban 4. és 8. évfolyamos diákok vettek részt. A felsőbb évfolyamosok körében párhuzamosan papír- és számítógép-alapú adatfelvétel zajlott. Az alkalmazott papíralapú tesztek egyike azonos itemeket tartalmazott, mint a számítógép-alapú teszt, ezzel lehetővé téve a minél pontosabb összehasonlítást és a közvetítő eszköz befolyásoló hatásának vizsgálatát. A 2002-es „Az írás online” (Writing Online – WOL) teszt felvétele 8. évfolyamos diákok körében történt (Sandene, Horkay, Bennett, Allen,

Braswell, Kaplan és Oranje, 2005), ami egy részmintája a NAEP papíralapú olvasás és írás mérésben részt vett diákoknak. A papír- és számítógép-alapú feladatlapok és feldolgozandó szövegek kialakítása lehetővé tette a két médiumon nyújtott teljesítmények összehasonlítását. A „Problémamegoldás technológiailag gazdag környezetben” tesztet (Problem Solving in Technology-Rich Environments – TRE) 2003-ban oldották meg (Bennett, Persky, Weiss és Jenkins, 2007) az érintett diákok. A NAEP online felmérések újdonsága, hogy eltávolodtak a kizárólag feleletválasztós itemek számítógép-alapú alkalmazásától, és kihasználtak három terület adta lehetőségeket itemtípusok tekintetében.

Az online és számítógép-alapú vizsgáztatás az Egyesült Államokban annyira elterjedt és alkalmazott, hogy már iparágak épültek köré, akik csak ezzel, online feladatok írásával, különböző tesztelő szoftverek készítésével, fejlesztésével foglalkoznak. Ennek ellenére 2009-ben három nagy technológiai viág cég összefogott és egy nagy volumenű, nemzetközi projektet indítottak el. „A 21. századi készségek mérése és tanítása” (Assessment and Teaching of 21st Century Skills) projekt keretében négy munkacsoportban folyik a munka. A program célja egyrészt a modern társadalmakban meghatározó jelentőségű készségek azonosítása, operacionalizálása, mérhető formában való leírása, az új mérési eljárások alkalmazását támogató osztálytermi környezettel kapcsolatos vizsgálatok, új mérésmetodikai, pszichometriai modellek kidolgozása, a fel-

Összességében megállapítható, hogy a technológia használata két különböző célból jelentik meg a mérés-értékelés területén. A kutatások egyik felében azért alkalmaznak különböző technológiai eszközöket, mert csak azokkal lehet felvenni az újonnan fejlesztett mérőeszközöket. Ezek a típusú felmérések drágák, ezért kis mintán, laboratóriumi körülmények között zajlanak általában.

A kutatások másik fele a technológia nagymintás mérésekbe történő bevezetésére fókuszál, ahol a cél a papíralapú tesztelésről való áttérés. Ebből adódóan az alkalmazott technológiák terén inkább építenek a meglévő technológiai eszközökre (iskolai infrastruktúrára), vagy külön a mérés céljából szállítanak ki eszközöket (ez azonban nagymintás mérések esetén jelentősen növeli a költségeket)

használható technológiai megoldások feltérképezése és az új értékelési kultúra elterjesztéséhez szükséges oktatáspolitikai feltételek elemzése (Csapó, 2009). „A 2009 októberéig elkészült négy átfogó tanulmány (White Paper) áttekinti a technológia alapú tesztelés jelenlegi helyzetét és megfogalmazza a továbblépéshez szükséges tennivalókat. A munkacsoportok több tucat olyan területet azonosítottak, amelyeken további kutatásokra van szükség, egyrészt azért, hogy a jelenlegi papír alapú értékelést felválthassa a technológiai alapú mérés, másrészt azért, hogy sor kerülhessen az új technológiákban rejlő további

lehetőségek kiaknázására. Az egyik legfontosabb kutatási feladat azonban az új mérési technológiák validitásának elemzése.” (Csapó, 2009)

Európában kisebb hagyománya van a számítógép-alapú mérésnek, ugyanakkor a technológia-alapú mérés-értékelés bevezetése, a technológia mérés-értékelés folyamatába történő integrálása, az értékelési folyamat hatékonyabbá tétele, új generációs értékelési módszerek kidolgozása az utóbbi években számos nemzetközi kutatást indukált Európában is. Fontosságát mutatja, hogy egyrészt az Európai Unió is támogatja mielőbbi megvalósulását (workshopok támogatása – Scheuermann és Björnsson, 2009; Scheuermann és Pereira, 2008), másrészt központi helyet foglal el a három technológiai világcég (Cisco, Intel és Microsoft) támogatásával zajló, fent említett globális kutatási projektben (Csapó, Latour, Bennett, Ainley és Law, 2009).

Nemzetközi szinten Európában Luxemburg vezeti a mezőnyt, ahol megtörtént már a papír-ceruza tesztekéről a számítógép-alapú tesztesztesre való átállás, sőt bevezettek egy országos online mérésre alkalmas rendszert. A rendszer jelenlegi verziója egyszerre képes tesztelni a teljes populációt, majd a teszteredményekről automatikus visszajelzést készít a tanárok részére (Plichart, Jadoul, Vandenabeele és Latour, 2004; Plichart, Latour, Busana és Martin, 2008). Ezt a rendszert adaptálta többek között az OECD PISA 2009 Electronic Reading Assessment (ERA – számítógépes olvasás-felmérés; Haldane, 2009), illetve az OECD PIAAC (Program for International Assessment of Adult Competencie, felnőtt kompetenciák *felmérése*) is. Ennek átvételén és fejlesztésén dolgozik a DIPF (Deutsches Institute für Internationale Pädagogisches Forschung) munkacsoportja, hogy majdan ez a rendszer képezze az alapját a német nemzeti értékelési rendszernek, ha a jelenleg az IQB (Humboldt Egyetem, Berlin) által fejlesztett papír-ceruza alapú mérésekről az elektronikus mérésekre áttérnek. A rendszer mellett kidolgoztak egy itemfejlesztő programot is (Itembilder), ami alkalmas komplexebb, interaktív elemeket tartalmazó innovatív itemformátumok létrehozására is. A DIPF által vezetett TBA (Technology Bases Assessment) projekt másik lényeges eleme a komplex problémamegoldás számítógép-alapú mérése (Wirth és Klieme, 2003, Wirth és Funke, 2005). A MicroDYN megközelítés fényében teljes mértékben kihasználják a technológia adta lehetőségeket, és egy dinamikusan változó környezettel szembesítik a tesztelt személyeket (Blech és Funke, 2005; Greiff és Funke, 2008). A TAO rendszer szerepel az EU ajánlásaiban, ezért a fejlesztő munkák koordinálására az Unió az Isprában működő Joint Research Center keretében elindított egy nemzetközi együttműködési programot.

Technológia-alapú mérés-értékelés szerepe az IEA- és az OECD-mérésekben

A legnagyobb nemzetközi empirikus adatokra támaszkodó mérési-értékelési programokat a mai napig az IEA és az OECD (Organization for Economic Co-operation and Development) bonyolítja le. Az utóbbi években mind a TIMSS, mind a PIRLS esetében egy-egy tanároknak és iskoláknak szóló kérdőív erejéig már tesztelték a számítógép-alapú felmérést, de diákok körében – pilot vizsgálatok kivételével – még nem alkalmazták. A számítógép-alapú vizsgálatok bevezetéséhez szükséges adatok begyűjtését, teljes körű helyzetfelmérést a SITES (Law, Pelgrum és Plomp, 2008) vizsgálatok valósították meg. Az eredmények és a nemzetközi trendek hatására az IEA 2011-re tervezi a PIRLS választható moduljaként a web-alapú olvasás mérését, illetve 2013-ra tervezi a számítógép alapú mérés tágabb bevezetését az ICILS (Computer and Information Literacy) projekt keretein belül. Utóbbiban mind iskolák, mind tanárok, mind diákok érintettek lesznek.

Az OECD PISA (Programme for International Student Assessment) felmérések 2000-ben indultak és háromévenkénti ciklusban ismétlődnek (OECD, 2000, 2004, 2006, 2007). A mintát 15 éves diákok alkotják. Három fő területet (matematikai, természettudományos műveltség, illetve olvasáskultúra) és egy minor területet vizsgálnak. A felmé-

rések 2006-ig kizárólag papíralapúak voltak. A technológia-alapú mérés bevezetését papíralapú kérdőívek előzték meg. Ezek többek között rákérdeztek a tanulók számítógéphasználati szokásaira, hozzáférhetőségére, számítógéppel kapcsolatos attitűdjükre, bizonyos számítógépes műveletek végzésének gyakoriságára, az abban való magabiztosságra (OECD, 2005). A háttérkérdőív elemzésével többek között a tanulók informatikai szokásai, képességei, attitűdjei minősíthetők (Dancsó, 2007). A technológia tesztelés során történő alkalmazása 2006-ban a „Természettudományi tudás számítógépes felmérése” (Computer-Based Assessment of Science, CBAS; OECD, 2009) nemzetközi opció meghirdetésével (Csapó, 2009) valósult meg. A részt vevő iskolákba hat, a tesztelés lebonyolításához szükséges programokkal ellátott laptopot szállítottak, amelyek közül egy a teszt-adminisztrátori funkciókat látta el (Csapó, Latour, Bennett, Ainley és Law, 2009). A diákok válaszaik mind az adott laptopon, mind az adminisztrátori laptopon rögzítésre lettek. A CBAS-ban való részvételre 13 ország jelentkezett, de a végén csak 3 vett részt (az eredményeket lásd: Martin, 2009; Halldórsson, McKelvie és Björnsson, 2009; Lee, 2009; Sørensen és Møller Andersen, 2009).

Az elektronikus tesztelés folytatásaként, elterjesztésének következő lépcsőfokaként a 2009-es mérési ciklusban az „Elektronikus szövegek olvasása” (Electronic Reading Assessment, ERA) volt nemzetközi opció. 2009-ben már több, mint 20 ország vállalta az elektronikus tesztelésben való részvételt. Az adatfelvételt a TAO platform segítségével bonyolították le. A tesztelést nem az interneten keresztül hajtották végre, hanem pendrive-okat szállítottak ki az iskolákba, amelyek tartalmazták a tesztelés lebonyolításához szükséges programokat, illetve a diákok válaszaikat is a pendrive-on lévő adatbázis-kezelő program kezelte, raktározta el.

A tervek szerint 2012-ben az „Általános problémamegoldás” (General Problem Solving) számítógépes felmérésére kerül sor (Csapó, 2009). A megvalósítási lehetőségek között felmerült a dinamikus problémamegoldás (lásd: Blech és Funke, 2005; Wirth és Funke, 2005; Greiff és Funke, 2008), melynek során a diákoknak a számítógéppel való interakció révén azonosítania kell egy véges állapotú automata viselkedési szabályait, majd az így megszerzett tudást különböző feladatok megoldásában kell hasznosítani (Csapó, 2009).

Az OECD szervezésében jelenleg előkészületi fázisban van a felnőttkori készségek vizsgálatára irányuló, 2011-re tervezett felmérés, a PIAAC (Programme for the International Assessment for Adult Competencies). A mérésben 27 ország 16 és 65 év közötti polgárai vesznek részt, minimum 5000 ember országonként (OECD, 2008). Három fő műveltségi területen vizsgálják a felnőttek képességeinek fejlettségi szintjét: szövegértés és -használat, számolás és technológiailag gazdag környezetben végzett problémamegoldás. A szóbeli interjúkat kérdezőbiztosok végzik, a tesztek a TAO platform közvetíti.

Összességében megállapítható, hogy a technológia használata két különböző célból jelentik meg a mérés-értékelés területén. A kutatások egyik felében azért alkalmaznak különböző technológiai eszközöket, mert csak azokkal lehet felvenni az újonnan fejlesztett mérőeszközöket. Ezek a típusú felmérések drágák, ezért kis mintán, laboratóriumi körülmények között zajlanak általában.

A kutatások másik fele a technológia nagymintás mérésekbe történő bevezetésére fókuszál, ahol a cél a papíralapú tesztelésről való áttérés. Ebből adódóan az alkalmazott technológiák terén inkább építenek a meglévő technológiai eszközökre (iskolai infrastruktúrára), vagy külön a mérés céljából szállítanak ki eszközöket (ez azonban nagymintás mérések esetén jelentősen növeli a költségeket).

A technológia-alapú mérés-értékelés hazai helyzete

Hazánkban az első nagyobb volumenű online tesztelés 2008 májusában zajlott. Az adatfelvételben 5. évfolyamos diákok vettek részt (n=843) 24 település, 34 iskolájából. A minta kiválasztása során nem volt cél a reprezentativitás. Ennek ellenére egy azonos korú reprezentatív mintának és a jelen kutatás mintájának szülők iskolai végzettsége szerinti eloszlásában ($\chi^2=7,13$, $p>0,05$) nincs szignifikáns különbség.

A papíralapú teszt számítógépre adaptálása során a kutatók igyekeztek a teszt minél több tulajdonságát megőrizni (lásd: *Csapó, Molnár és R. Tóth, 2009; R. Tóth, Molnár és Csapó, 2009; Csapó, Molnár, Pap-Szigeti és R. Tóth, 2009*). A kognitív teszt (induktív gondolkodás teszt) mellett mind a tesztelés előtt, mind utána egy-egy háttérkérdőívet (IKT kérdőív és szocio-ökonómiai háttér adatokra kérdező kérdőív) is kitöltötték a diákok, illetve a tesztelés során közreműködő tanárok is.

Minden egyes tanuló először papíron, majd néhány hét különbséggel számítógépen is megoldotta az induktív gondolkodás fejlettségét mérő tesztet. A teszt megoldására mindkét formátumban 35 perc állt a tanulók rendelkezésére. A számítógépes adatfelvétel interneten zajlott a TAO (Testing Assisté par Ordinateur – számítógépes tesztelés) platformon keresztül (*Plichart, Jadoul, Vandenabeele és Latour, 2004; Latour és Farcot, 2008; Latour és Martin, 2007*). A kutatás eredményeit lásd *Csapó, Molnár és R. Tóth (2009)*, illetve *Csapó, Molnár, Pap-Szigeti és R. Tóth (2009)* tanulmányaiban.

Ez a típusú kutatás, amikor minden egyes diák mindkét médiumon megoldja az alkalmazott tesztet, nemzetközi viszonylatban is egyedi. Ebben az esetben diákszinten összehasonlítható az alkalmazott médium befolyásoló hatása.

A fokozatos eltávolodás jegyében 2009 tavaszán az SZTE Oktatásméleti Kutatócsoport online tesztelés bevezetéséért felelős munkacsoportja folytatta a papír-ceruza és számítógép-alapú tesztelés összehasonlító vizsgálatait. Három területen (matematika, olvasás és problémamegoldás) és két évfolyamon (2. és 6. évfolyam) hajtott végre párhuzamosan papír- és számítógép-alapú adatfelvételt. A minta és a mérőeszközök kiválasztása, illetve az alkalmazott itemtípusok különböztek az egyes területek esetében, így az elemzések során lehetővé válik a kutatás tervezéséből, az alkalmazott itemek különbözőségéből adódó eltérések pontosabb leírása, azonosítása (*Molnár, 2009; R. Tóth, 2009; Hódi és R. Tóth, 2009*).

2010-ben folytatódtak a fent említett kutatások és a TAO platform egyik modulja segítségével, a TAO CAPI kérdőívmodullal megtörtént egy tanároknak szóló nem lineáris kérdőív felvétele (*Tóth, 2009*).

A kutatócsoport az elkövetkező néhány évben több területen számos hasonló mérést tervez, fokozatosan eltávolodva a papíralapú mérés korlátaitól és áttérve a számítógép-alapú mérésre, értékelésre.

A fent ismertetett kutatás és hasonló nemzetközi kutatások eredményei valószínűsítik, hogy az áttérés folyamán még jó ideig párhuzamosan fog zajlani a papír- és számítógép-alapon történő adatfelvétel. Míg szükséges a korábbi eredményekkel történő összehasonlítás, addig a mérés hasonló típusú tervezése a papíralapú tesztelésétől való fokozatos eltávolodás fényében nélkülözhetetlen.

Irodalom

- Bebell, D. és Kay, R. (2010): One to One Computing: A Summary of the Quantitative Results from the Berkshire Wireless Learning Initiative. *The Journal of Technology, Learning, and Assessment*, 9. 2. sz. <http://escholarship.bc.edu/jtla/vol9/2/>
- Bebell, D. és O'Dwyer, L. (2010): Educational Outcomes and Research from 1:1 Computing Settings. *The Journal of Technology, Learning, and Assessment*, 9. 1. sz. <http://escholarship.bc.edu/jtla/vol9/1/>
- Bejar, I. I., Lawless, R. R., Morley, M. E., Wagner, M. E., Bennett, R. E. és Revuelta, J. (2003): A feasibility study of on-the-fly item generation in adaptive testing. *Journal of Technology, Learning and Assessment*, 2. 3. sz. <http://www.bc.edu/research/intasc/jtla/journal/v2n3.shtml>
- Bennett, R. E. (2002): Inexorable and Inevitable: The Continuing Story of Technology and Assessment. *Journal of Technology, Learning and Assessment*, 1. 1. sz. <http://escholarship.bc.edu/jtla/vol1/1/>
- Bennett, R. E., Goodman, M., Hessinger, J., Liggett, J., Marshall, G., Kahn, H. és Zack, J. (1999): Using multimedia in large-scale computer-based testing programs. *Computers in Human Behavior*, 15. sz. 283–294.
- Bennett, R. E., Jenkins, F., Persky, H. és Weiss, A. (2003): Assessing complex problem-solving performances. *Assessment in Education*, 10. sz. 347–359. <http://www.ets.org/Media/Research/pdf/RM-03-03-Bennett.pdf>
- Bennett, R. E., Persky, H., Weiss, A. R. és Jenkins, F. (2007): Problem solving in technology-rich environments: A report from the NAEP Technology-Based Assessment Project (NCES 2007-466). National Center for Education Statistics, US Department of Education, Washington, DC. <http://nces.ed.gov/pubsearch/pubsinfo.asp?pubid=2007466>
- Ben-Simon, A. és Bennett, R. E. (2007): Toward more substantively meaningful automated essay scoring. *Journal of Technology, Learning and Assessment*, 6. 1. sz. <http://escholarship.bc.edu/jtla/vol6/1/>
- Blech, C. és Funke, J. (2005): Dynamis review: An overview about applications of the Dynamis approach in cognitive psychology. Deutsches Institut für Erwachsenenbildung, Bonn. http://www.die-bonn.de/espid/dokumente/doc-2005/blech05_01.pdf
- Csapó Benő (2009): Nemzetközi kutatási-fejlesztési programok a technológiai alapú értékelés elősegítésére. IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19–21. 94.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papír alapú teszektől a számítógépes adaptív tesztelésig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 3–4. sz. 3–16.
- Csapó Benő, Molnár Gyöngyvér, Pap-Szigeti Róbert és R. Tóth Krisztina (2009): A mérés-értékelés új tendenciái: a papír- és számítógép-alapú tesztelés összehasonlító vizsgálatai általános iskolás, illetve főiskolás diákok körében. In: Kozma Tamás és Perjés István (szerk.): *Új kutatások a neveléstudományokban*. MTA Pedagógiai Bizottság, Budapest. 99–108.
- Csapó, B., Latour, T., Bennett, R., Ainley, J. és Law, N. (2009): *Technological Issues of Computer-Based Assessment of 21st Century Skills*. Kézirat.
- Csapó, B., Molnár, G. és R. Tóth, K. (2009): Comparing paper-and-pencil and online assessment of reasoning skills. A pilot study for introducing electronic testing 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.
- Dancsó Tünde (2007): A PISA 2003 eredményeinek hazai vonatkozású elemzése az IKT-eszközök használatáról. In: Korom Erzsébet (szerk.): *V. Pedagógiai Értékelési Konferencia. Szeged, 2007. április 12-14. Program. Tartalmi összefoglalók*. Szegedi Tudományegyetem, Szeged. 88.
- Dorner Helga (2007): *Kollaboratív tudásépítés számítógéppel segített tanulási környezetben – a tudásépítő interakciók elemzése*. MultiMédia az Oktatásban 2007 konferencia, Budapesti Műszaki Főiskola, 2007. augusztus 23–24.
- Greiff, S. és Funke, J. (2008): *Measuring complex problem solving: The MicroDYN approach*. Kézirat. Universität Heidelberg, Heidelberg. http://www.psychologie.uni-heidelberg.de/ae/allg/forschun/dfg_komp/Greiff&Funke_2008_MicroDYN.pdf
- Halldórsson, A., McKelvie, P. és Björnsson, J. (2009): Are Icelandic boys really better on computerized tests than conventional ones: Interaction between gender test modality and test performance. 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. 178–193.
- Hódi Ágnes és R. Tóth Krisztina (2009): Olvasási képesség mérése számítógépes környezetben. IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19–21. 96.
- Horkay, N., Bennett, R. E., Allen, N., Kaplan, B. és Yan, F. (2006): Does it matter if I take my writing test on computer? An empirical study of mode effects in NAEP. *Journal of Technology, Learning and Assessment*, 5. 2. sz. <http://escholarship.bc.edu/jtla/vol5/2/>
- Kyllonen, P. (2009): New constructs, methods and directions for computer-based assessment. 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. 151–156.

Latour, T. és Farcot, M. (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.

Latour, T. és Martin, R. (2007): *TAO, An Open and Versatile Computer-Based Assessment Platform Based on Semantic Web Technology*. ECRIM. <https://www.tao.lu/downloads/publications/TAO-ErcimNews71-Oct2007.pdf>

Law, N., Pelgrum, W. J. és Plomp, T. (2008, szerk.): *Pedagogy and ICT use in schools around the world: Findings from the IEA SITES 2006 study*. CERC-Springer, Hong Kong.

Lee, M-K. (2009): CBAS in Korea: Experiences, Results and Challenges. 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. 187–193.

Lipponen, L. (2002): Exploring foundations for computer-supported collaborative learning. In: Stahl, G. (szerk.): *Proceedings of Computer Supported Collaborative Learning*. Boulder, CO. 72–81.

Macdonald, J. (2003): Assessing online collaborative learning: process and product. *Computers & Education*, **40**. 4. sz. 377–391.

Martin, R. (2009): Utilising the Potential of Computer Delivered Surveys in Assessing Scientific Literacy. 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. 165–170.

Molnár Gyöngyvér (megjelenés alatt): Új ICT eszközök alkalmazása az iskolai gyakorlatban. In: Korom Erzsébet (szerk.): *Kihívások a XXI. század iskolájában*.

Molnár Gyöngyvér (2008): A Rasch modell kiterjesztése nem dichotóm adatok elemzésére: a rangskálás és parciális kredit modell. *Iskolakultúra*, **18**. 1. sz. 66–77.

Molnár Gyöngyvér (2009): *Papír- és számítógépalapú tesztelés összehasonlító vizsgálata problémamegoldó környezetben*. IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19–21. 95.

Molnár Pál (2009): Számítógéppel támogatott együttműködő tanulás online közösségi hálózatos környezetben. *Magyar Pedagógia*, 3. sz.

Molnár Pál és Kárpáti Andrea (2009): Az együttműködő tanulás támogatása az oktatási informatika eszközeivel: MapIt vitatérkép. *Új Pedagógiai Szemle*, **59**. 2. sz. 48–60.

Moss, J. és Hendry, G. (2002): Use of electronic surveys in course evaluation. *British Journal of Educational Technology*, **33**. 5. sz. 583–592.

Negroponte, N. (1995): *Being digital*. Vintage, New York.

Neo, M. és Neo, K. T. K. (2001): Innovative teaching: Using multimedia in a problem-based learning environment. *Educational Technology & Society*, **4**. 4. sz. 19–31.

Oblinger, D. G. és Oblinger, J. L. (2005): Is It Age or IT: First Steps Toward Understanding the Net Generation. In: Oblinger, D. G. and Oblinger, J. L. (szerk.): *Educating the Net Generation*. EDUCAUSE. 2.1-2.20.

OECD (2000): *OECD Education at a Glance – OECD indicators*. OECD, Paris.

OECD (2004): *Learning for Tomorrow's. First results from PISA 2003*. OECD, Párizs.

OECD (2005): *Are Students Ready for a Technology-Rich World? What PISA Studies Tell us?* OECD, Paris. <http://www.oecd.org/dataoecd/28/4/35995145.pdf>

OECD (2006): *Assessing Scientific, Reading and Mathematical Literacy. A Framework for PISA 2006*. OECD, Paris. <http://www.oecd.org/dataoecd/63/35/37464175.pdf>

OECD (2007): *PISA – The OECD Programme for International Student Assessment*. OECD, Paris. <http://www.oecd.org/dataoecd/51/27/37474503.pdf>

OECD (2009). *PISA CBAS analysis and results – Science performance on paper and pencil and electronic tests*. OECD, Paris.

O'Malley, C. E. (1995, szerk.): *Computer-Supported Collaborative Learning*. Springer-Verlag, Berlin.

Plichart, P., Jadoul, R., Vandenabeele, L. és Latour, T. (2004): *TAO, a Collective distributed computer-based assessment framework built on semantic web standards*. In: Proceedings of the International Conference on Advances in Intelligent Systems – Theory and Application AISTA2004, In cooperation with IEEE Computer Society, November 15–18, 2004. Luxembourg, Luxembourg.

Plichart, P., Latour, T., Busana, G. és Martin, R. (2008): Computer based school system monitoring with feedback to teachers. In: *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008*. AACE, Chesapeake, VA. 5065–5070.

Powers, D. E. és Potenza, M. T. (1996): *Comparability of testing using laptop and desktop computers (RR-96-15)*. Educational Testing Service, Princeton,

- NJ. 2009. 10. 05-I megtekintés, http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/14/dc/69.pdf
- R. Tóth Krisztina (2009): A számítógépes és papírceruza tesztek eredményeinek összehasonlító vizsgálata háttérváltozók alapján. IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19–21. 97.
- R. Tóth Krisztina és Molnár Gyöngyvér (2009): A jövő tanárainak IKT ismeretei és elvárásai. *Pedagógusképzés*, 7. 1. sz.
- R. Tóth Krisztina, Molnár Gyöngyvér és Csapó Benő (2008): *A számítógépes tesztelés lehetőségei*. VI. Pedagógiai Értékelési Konferencia, Szeged, 2008. április 11–12. 84.
- R. Tóth, K., Molnár, G. és Csapó, B. (2009): *Online assessment of reasoning skills*. Paper presented at the 13th European Conference for the Research on Learning and Instruction. Amsterdam, 2009. 08. 25–29.
- Sandene, B., Horkay, N., Bennett, R., Allen, N., Braswell, J., Kaplan, B. és Oranje, A. (2005): *Online Assessment in Mathematics and Writing: Reports From the NAEP Technology-Based Assessment Project, Research and Development Series (NCES 2005–457)*. U.S. Department of Education, National Center for Education Statistics, U.S. Government Printing Office, Washington, DC.
- Scardamalia, M. (2002): Collective cognitive responsibility for the advancement of knowledge. In: Smith, B. (szerk.): *Liberal education in a knowledge society*. Open Court, Chicago, IL. 67–98.
- Scheuermann, F. és Björnsson, J. (2009, szerk.): *New approaches to skills assessment and implications for large-scale testing. The transition to computer-based assessment*. Office for Official Publications of the European Communities, Luxembourg.
- Scheuermann, F. és Pereira, G. A. (2008, szerk.): *Towards a research agenda on computer-based assessment*. Office for Official Publications of the European Communities, Luxembourg.
- Sørensen, H. és Andersen, A. M (2009): How did Danish Students solve the PISA CBAS items? Right and Wrong Answers from a Gender Perspective. 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. 194–207.
- Suhr, K. A., Hernandez, D. A., Grimes, D. és Warschauer, M. (2010): Laptops and Fourth Grade Literacy: Assisting the Jump over the Fourth-Grade Slump. *The Journal of Technology, Learning, and Assessment*, 9. 5. sz. <http://escholarship.bc.edu/jtla/vol9/5/>
- Tóth Edit (2009): A tanárok értékeléssel kapcsolatos attitűdjei, nézetei (tematikus előadás). In: Bárdos Jenő és Sebestyén József (szerk.): *IX. Országos Neveléstudományi Konferencia: Neveléstudomány – Integritás és integrálhatóság. Program és tartalmi összefoglalók*. Magyar Tudományos Akadémia, Pedagógiai Bizottság, Veszprém. 206–207.
- Williamson, D. M., Mislevy, R. J. és Bejar, I. I. (2006, szerk.): Automated scoring of complex tasks in computer-based testing. Erlbaum, Mahwah, NJ. 403–412.
- Wirth, J. és Funke, J. (2005): Dynamisches Problemlösen: Entwicklung und Evaluation eines neuen Messverfahrens zum Steuern komplexer Systeme. In: Klieme, E., Leutner, D. és Wirth, J. (szerk.): *Problemlösekompetenz von Schülerinnen und Schülern*. VS Verlag für Sozialwissenschaften, Wiesbaden. 55–72.
- Wirth, J. és Klieme, E. (2003): Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, 10. 3. sz. 329–345.
- Zhang, Y., Powers, D. E., Wright, W. és Morgan, R. (2003): *Applying the Online Scoring Network (OSN) to Advanced Placement Program (AP) tests (RM-03-12)*. Educational Testing Service, Princeton, NJ. <http://www.ets.org/research/researcher/RR-03-12.html>.
- A kutatás a TÁMOP 3.1.9 program, az Oktatásméleti Kutatócsoport és az MTA-SZTE Képességkutató Csoport keretében zajlott. Molnár Gyöngyvér a kutatás során Bolyai János Kutatási Ösztöndíjban részesült.*