



A PhD-értekezés tézisei

TÓTH KRISZTINA

**FELSŐ TAGOZATOS DIÁKOK PAPÍR-CERUZA ÉS SZÁMÍTÓGÉP
ALAPÚ TESZTEREDMÉNYEINEK ÖSSZEHASONLÍTÓ VIZSGÁLATA
SZÖVEGÉRTÉS, INDUKTÍV GONDOLKODÁS ÉS
PROBLÉMAMEGOLDÁS TERÉN**

Témavezető:

Dr. Molnár Gyöngyvér
habilitált egyetemi docens

Szegedi Tudományegyetem
Bölcsészettudományi Kar
Neveléstudományi Doktori Iskola
Információs és kommunikációs technológiák az oktatásban
doktori program

Szeged
2015

BEVEZETÉS

Az elmúlt négy évtizedben számos pedagógiai mérés-értékeléssel foglalkozó intézet és szervezet kezdte meg papír alapú mérőeszközeinek digitalizálását, számítógépes adaptációját (Csapó, Ainley, Bennett, Latour és Law, 2012). A papír-ceruza tesztek számítógépes változatának elkészítését számos tényező ösztönzi (Bennett, 2003; Bodmann és Robinson, 2004; Choi és Tinkler, 2002; Csapó, Molnár és Tóth, 2008; Horne, 2007), például az adminisztráció leegyszerűsödése, a hosszú távú költséghatékonyság, a számítógépes mérés-értékelés kínált lehetőségek, mint az új, papír-alapon nem alkalmazható itemtípusok (Tóth, Molnár, Wüstenberg, Greiff és Csapó, 2011), tesztalgoritmusok használata (Scalise és Gifford, 2006; Schröders és Wilhelm, 2011), illetve a log fájl elemzés (Goldhammer, Naumann, Stelter, Tóth, Rölke és Klieme, 2014; Tóth, Rölke és Goldhammer, 2012a, 2012b; Tóth, Rölke, Greiff és Wüstenberg, 2014; Tóth, Rölke, Naumann és Goldhammer, 2012; Tóth, Wüstenberg, Rölke és Greiff, 2012) nyújtotta előnyök.

Noha a számítógépes tesztelés már több évtizede számos formában jelen van (Asuni, 2009) és egyre szélesebb körben alkalmazzák, még ma is számos kihívással (Paek, 2005) szembesíti a neveléstudományi kutatókat már azon a szinten is, amikor tradicionális, papír alapú tesztjeiket számítógépes környezetben prezentálják. Egy új tesztkörnyezet (tesztet közvetítő eszköz) bevezetésével kapcsolatban felmerül a kérdés, hogy a közvetítőeszköz megváltozása hatást gyakorol-e a diákok teljesítményére, és biztosítva van-e, hogy a diák nem részesül előnyben, illetve nem kerül hátrányba azáltal, hogy számítógépes tesztek alkalmazunk papír-ceruza mérőeszközök helyett (Paek, 2005). Erre a kutatási kérdésre nincs egyértelmű válasz a szakirodalomban, mert (1) a technológia folyamatosan fejlődik, ezáltal változnak a technológiai és tesztinterfész-paraméterek is, valamint (2) a mérésekben résztvevő diákok, (3) a mérési terület és az alkalmazott mérőeszköz jellemzőinek sokszínűsége szintén befolyásolhatja a tanulói teljesítményeket. Ezért olyan pedagógiai/pszichológiai vizsgálatok esetén, ahol két különböző tesztkörnyezetből származó adattal dolgozunk, biztosítanunk kell, hogy a két tesztmédiumon felvett adat felcserélhető legyen, illetve ha nem ekvivalensek a pontszámok, akkor minimalizálnunk kell a tesztkörnyezet megváltozásának hatását az érvényes tudományos következtetések levonásához.

ELMÉLETI HÁTTÉR

A papír-ceruza és az elektronikus formában felvett teszteken elért eredmények ekvivalenciájának kérdésével számos nemzetközi ajánlás foglalkozik (l. pl. Bartram és Coyne, 2005), melyek alapján a papír-ceruza és számítógépes környezetben elért teljesítmények egyezősége empirikus úton, összehasonlító vizsgálatok segítségével igazolható (Paek, 2005). Amennyiben a két környezetből származó tesztpontszám statisztikai egyenlősége nem bizonyított, mert például ugyanazon teszt papíron vagy számítógépes környezetben felvéve nem eredményez ekvivalens tanulói teljesítményeket (Leeson, 2006), a közvetítőeszköz befolyásoló hatása igazolódik, melyet a továbbiakban médiahatásnak (mode effect; Clariana és Wallace, 2002) nevezünk. Amennyiben a teszteredmények ekvivalenciája biztosított, az elért teljesítmény médiafüggetlen.

A tesztkörnyezet teljesítményre gyakorolt hatásának vizsgálata már az 1970-es évektől kezdve jelen van (*Gaskill és Marshall, 2007*). Így az elmúlt évtizedekben már több, mint 300 nemzetközi tanulmány központi témája a tesztkörnyezet megváltozásából adódó teljesítménykülönbségek vizsgálata és az eltérések okainak feltárása (*Wang, Jiao, Young, Brooks és Olson, 2008*). A médiahatás-vizsgálatok egy része nem igazolt szignifikáns különbséget a tradicionális és a számítógépes környezetben elért teszteredmények között (*Horkay, Bennett, Allen, Kaplan és Yan, 2006; Johnson és Green, 2006; Puhan, Boughton és Kim, 2007; Wang és mtsai, 2007, 2008*). Ezzel szemben más médiahatás-kutatások szignifikáns különbséget igazoltak a két tesztkörnyezetben elért tanulói teljesítmények között (*Bennett, Braswell, Oranje, Sandene, Kaplan és Yan, 2008; Csapó, Molnár és Tóth, 2009; Higgins, Russell és Hoffmann, 2005; Jeong, 2014; Molnár, Tóth és Csapó, 2011; Tóth és Hódi, 2011*).

A kutatási eredmények inkonzisztenciájának hátterében az áll, hogy a médiahatás-vizsgálatok eltérő mérési területen, különböző típusú mérőeszközökkel, másfajta szoftverekkel és szoftverbeállításokkal készültek. Gyakran a diákok különböző populációjára fókuszáltak, és eltérő adatfelvételi és adatelemzési módszereket alkalmaztak.

AZ EMPIRIKUS VIZSGÁLAT CÉLJAI, KUTATÁSI KÉRDÉSEI

A disszertációban bemutatott vizsgálatok célja a médiahatás mértékének vizsgálata induktív gondolkodás, szövegértés és problémamegoldás terén. A 2008-2014 közötti adatfelvételekből származó empirikus adatok elemzését több szempont szerint is elvégeztük: a mérési terület, a minta- és az itemjellemzők, valamint a médiahatás mértékének összefüggéseit elemeztük. Az iskolákban elérhető infrastrukturális tényezők hatásának vizsgálatára nem térünk ki, mert ez az infrastruktúra folyamatosan változik, megújul, valamint az online tesztek felvétele során biztosítottuk, hogy az iskolákban elérhető legkisebb és legnagyobb méretű monitoron is minden információ egyformán jelenjen meg.

A mérési terület jellemzőinek vizsgálata során a nemzetközi ajánlásokban (l. pl. *Bartram és Coyne, 2005*) foglaltakat követve összehasonlítjuk a tesztpontszámok statisztikai egyenlőségét a teljes teszt és a résztesztek szintjén. Megvizsgáljuk, igazolódik-e médiahatás a három vizsgált területen, mert a médiahatás-vizsgálatok (l. pl. *Bennett és mtsai, 2008; Hülber és Molnár, 2013; Wang és mtsai, 2007, 2008*) inkonzisztens eredményei alapján nem zárható ki a médiahatás jelenléte. A komparatív elemzéseinkben kitérünk a médiahatás mérési területek szerinti vizsgálatára, összehasonlítjuk a papír-ceruza és számítógépes tesztek reliabilitás-mutatóit is, azért, hogy igazoljuk, a diákok teljesítménye azonos megbízhatósággal mérhető a különböző tesztmédiákon (*Lottridge, Nicewander, Schulz és Mitzel, 2008*).

A minta jellemzői mentén végzett vizsgálatok demográfiai háttérváltozók (pl. kor, nem, településtípus, szocio-ökonómiai státusz) és számítógép használatához kapcsolódó változók (pl. számítógépes tapasztalat, számítógépes attitúd és számítógépes ismeret), valamint a papír alapon elért teljesítmények segítségével elemzik a médiahatás mértékét és irányát. A médiahatás irányának vizsgálata során arra kerestük a választ, hogy papír alapon vagy számítógépen igazolódik magasabb teljesítmény.

Függetlenül attól, hogy igazolódik-e médiahatás a papír-ceruza és online átlagteljesítmények között, a teljes teszt, illetve résztesztszintű vizsgálatokat kiegészítjük itemszintű elemzésekkel. Ennek okai: (a) amennyiben teszt- vagy résztesztszintű médiahatás

igazolódik, az itemszintű vizsgálatok feltárhatják a médiahatás okát; (b) a tanulói átlagteljesítmények médiafüggetlensége esetén is előfordulhat, hogy egyes itemek működése eltér a két tesztkörnyezetben (l. pl. *Gu, Drake és Wolfe, 2006; Poggio, Glasnapp, Yang és Poggio, 2005; Sandene, Horkay, Bennett, Allen, Braswell, Kaplan és Oranje, 2005*).

AZ EMPIRIKUS VIZSGÁLATOK MÓDSZEREI

Mérőeszközök

A médiahatás-vizsgálat elvégzéséhez alkalmazott tesztek a Szegedi Longitudinális Program keretében is kiközvetített mérőeszközök voltak. A mérőeszközök további közös jellemzője, hogy korábban papír alapon széles körben használt, *jól működő tesztek* voltak, illetve a tesztek feladatai, itemei jelentős változtatás nélkül *digitalizálhatóak* voltak. A kutatásba bevont mérőeszközök kiválasztása során további fontos szempont volt, hogy a tesztek különböző képességterület mérésére fókuszáljanak, ezzel kiküszöbölve az eredmények és levont következtetések esetleges konstruktumfüggőségét.

Az első médiahatás kutatásban *Csapó (1994)* széles körben alkalmazott és validált induktív gondolkodás tesztjére esett a választás. A második empirikus vizsgálatban *Molnár Éva* által készített szövegértésteztet és *Molnár (2010)* problémamegoldás-tesztjét vontuk be a médiahatás-elemzésekbe. A képességtesztek mellett a médiahatás háttérváltozóinak vizsgálatára háttérkérdőívet készítettük.

Az induktív gondolkodás teszt

Az induktív gondolkodás teszt (58 item) három résztesztből állt: a számanalógiák (14 item), szóanalógiák (28 item), és számsorok (16 item) résztesztből (*Csapó, 1994*). A számanalógiák és a számsorok részteszt számolást igénylő, feleletalkotó feladatokat, a szóanalógiák részteszt pedig rövid, feleletválasztó itemeket tartalmazott.

A szövegértéstezt

A szövegértési képesség vizsgálatát két, egymással párhuzamos tesztváltozattal végeztük. A párhuzamos tesztváltozatok ugyanolyan instrukciókat és (kizárólag feleletválasztó) itemtípusokat tartalmaztak, valamint mindkét teszt 36 (feleletválasztó) itemből állt (l. *Tóth és Hódi, 2011*). A papír-ceruza és számítógépes tesztek felépítése megegyezett, mindkét teszt két résztesztből állt. Az első részteszt egy bekezdésekre tagolt, folyamatos szöveget tartalmazott, a második részteszt nem folyamatos szövegekből (táblázat, térképek és diagram) állt.

A problémamegoldás-teszt

A problémamegoldás-teszt (28 item) egy realiztikus helyzetbe ágyazott történetet mutatott be, melynek feladatai egy öttagú család nyaralása során felmerült problémák voltak. A teszt a feladatok típusát tekintve feleletválasztó és feleletalkotó itemeket tartalmazott. Minden feladat megoldásához számolásra volt szükség.

A háttérkérdőív

A háttérváltozók vizsgálatához egy hosszabb terjedelmű kérdőívet (64 item) készítettünk, melyben feltártuk a tanulók demográfiai jellemzőit (pl. nem, szülők iskolai végzettsége, könyvek száma), számítógépes tapasztalatát, attitűdjét és ismereteit. A kérdőív felépítését tekintve két fő komponensre tagolódott: háttér adatok (9 item) és a számítógép használatához kapcsolódó tényezők (56 item).

A számítógépes tapasztalat mérésére 21 kérdőívtételből álló részkérdőívet készítettünk. Azt kívántuk feltárni (Jones és Clarke, 1995 alapján), hogy a tanulók mióta, milyen gyakran használnak számítógépet, naponta hány órán át számítógépeznek, interneteznek és milyen tevékenységeket végeznek számítógépezés közben (pl. szövegszerkesztés, levelezés, zenehallgatás).

A számítógépes attitűd vizsgálata során a számítógép fontossága, a számítógépes szorongás és a számítógépezés élvezete faktorokra fókuszáltunk. A számítógépes attitűd méréséhez Knezek, Christensen, Miyashita és Ropp (2000) a disszertáció mintájának megfelelő korú diákokon bemért, jól működő Számítógépes attitűd kérdőív (*Computer Attitude Questionnaire*) kérdőív tételait adaptáltuk (18 item).

A számítógépes ismeretek feltárására készített részkérdőív 17 itemből állt. Ebben a részkérdőívben a különböző számítógépes tevékenységek elsajátítási szintjére kérdeztünk rá.

Minta

Az empirikus kutatások mintáját 5–7. évfolyamos tanulók alkották, melynek összetételét és a mintakialakítás módszerét az 1. táblázatban foglaltuk össze. Az elemzésekben felhasznált papír alapú mérőeszközzel végzett kutatások 2008 és 2014 között, a Szegedi Tudományegyetem Oktatáselméleti Kutatócsoportja által vezetett követővizsgálat (Csapó, 2007) keretében zajlottak.

1. táblázat. A vizsgálat sorozat mintái

Mérési terület	Évfolyam	Adatfelvétel		Minta	N
		PP	CB		
Induktív gondolkodás	5.	2008	2008	azonos	843
Induktív gondolkodás	5.	2008	2010	illesztett	702
Induktív gondolkodás	6.	2005	2010	illesztett	262
Induktív gondolkodás	7.	2010	2014	illesztett	410
Szövegértés	6.	2009	2009	azonos	449
Problémamegoldás	6.	2010	2010	illesztett	342

Megjegyzés: PP: papír-ceruza adatfelvétel; CB: számítógépes adatfelvétel

Az online teszteredmények papír-ceruza eredményekkel történő összevetését két különböző mintavételi eljárás segítségével végeztük: (a) ugyanaz a tanuló megoldotta

ugyanannak a tesztnek a papír-ceruza és az online változatát (a táblázatban azonos minta névvel jelöltük); vagy (b) korábbi, longitudinális vizsgálatban résztvevő tanulók eredményeihez hasonlítottuk a számítógépes környezetben elért teljesítményeket (illesztett mintavételezéssel).

Adatfelvétel

A papír-ceruza adatfelvétel az SZTE Oktatáselméleti Kutatócsoport longitudinális vizsgálatának keretein belül, 2008 és 2014 között zajlott. A számítógépes adatfelvételt a TAO-platfornon (Testing Assisté par Ordinateur – számítógépes tesztelés; *Farcot és Latour*, 2008; *Tóth, Molnár, Latour és Csapó*, 2011) keresztül végeztük. A háttérkérdőívet szintén elektronikus formában juttattuk el a diákokhoz a TAO-platforn kérdőívszerkesztő moduljának (*Devosa és Vizvári*, 2011) segítségével.

Elemzési módszerek

A disszertációban klasszikus és valószínűségi tesztelmélethez tartozó elemzési módszereket alkalmaztunk, továbbá a mesterséges intelligencia területéről származó, adatbányászati eljárással is kiegészítettük komparatív vizsgálatainkat. Az elemzések elvégzéséhez SPSS és Conquest adatelemző szoftvert, R-programcsomagot, valamint a WEKA adatbányászati szoftvert alkalmaztunk.

A DISSZERTÁCIÓ FŐBB EREDMÉNYEI

Az empirikus vizsgálat eredményeit a mérési terület, a minta és az itemjellemzők mentén összegezzük.

A médiahatás vizsgálata a mérési terület jellemzői alapján

A három mérési területen végzett papír-ceruza és számítógépes tesztek megbízhatósági mutatóinak összehasonító vizsgálata, *Lundy* (2008) eredményeivel egyezően, jellemzően nem igazolt jelentős reliabilitásbeli eltéréseket a teljes teszt szintjén (l. *Molnár, Tóth és Csapó*, 2011). Az induktív gondolkodás teszt részteszt szintű elemzéseinek eredményei szerint, a feleletválasztó feladatokat tartalmazó szóbeli analógiák részteszt feladatai azonos megbízhatósággal mértek a két tesztmédiumon. Ezzel szemben a feleletalkotó feladatokat tartalmazó részteszteken jelentős különbség adódott a papír-ceruza és számítógépes Cronbach- α értékek között; a papír-ceruza teszteken szignifikánsan magasabb reliabilitás igazolódott, mint a számítógépes tesztváltozatokon (l. *Tóth*, 2014).

A szövegértéstezt részteszt szintű vizsgálatának eredményei alapján a nem folyamatos szövegeken statisztikailag egyenlő reliabilitásmutató igazolódott a két tesztkörnyezetben, ugyanakkor a folyamatos szöveghez tartozó részteszten, papír-ceruza környezetben magasabb volt a Cronbach- α érték, mint online (l. *Tóth és Hódi*, 2010). A problémamegoldás-teszt nem osztható résztesztekre, ezért ezen nem végeztünk részteszt szintű elemzéseket.

A papír-ceruza és számítógépes környezetben adódó teljesítménykülönbségek vizsgálata minden mérési területeken kis- és közepes mértékű médiahatást igazolt. *Choi és Tinkler* (2002), valamint *Pomplun, Ritchie és Custer* (2006) eredményeivel egyezően minden mérési területen szignifikánsan magasabb tesztpontszámot értek el a diákok papír alapon, mint a számítógépen (l. *Csapó, Molnár, Pap-Szigeti és Tóth*, 2009; *Hódi és Tóth*, 2010; *Molnár és Tóth*, 2008; *Tóth*, 2010; *Tóth, Molnár és Csapó*, 2009). Az elemzések eredménye alapján megállapítottuk, hogy a tesztet közvetítő környezet hatást gyakorolt a diákok átlagteljesítményére.

Az induktív gondolkodás teszt eredményeinek résztesztszintű vizsgálata igazolta, hogy a számolást igénylő, feleletalkotó feladatokon, papír alapon magasabb teljesítményt értek el a diákok, mint számítógépen. A feleletválasztó feladatokat tartalmazó részteszten pedig jellemző volt a médiahatás-függetlenség, illetve a kismértékű médiahatás.

A szövegértéstezsten a különböző típusú szövegeken adódó médiahatás mértékét modern tesztelméleti módszerrel megvizsgálva megállapítottuk, hogy a diákok mindkét szövegtípus esetén, papír-ceruza környezetben magasabb tesztpontszámot értek el, mint online. Továbbá a folyamatos szövegek megértése során jelentősebb médiahatás adódott, mint a nem folyamatos szövegek esetén (l. *Tóth és Hódi*, 2011).

A minta jellemzőinek és a médiahatás mértékének kapcsolata

A három évfolyamra kiterjedő, illesztett mintavétellel végzett hosszmetzeti vizsgálat eredményei alapján – *Kingston* (2009) konklúziójával egyezően – az évfolyam (kor) előrehaladtával nem csökken a médiahatás mértéke sem a teljes teszt, sem a résztesztek szintjén. *Poggio* és munkatársai (2005), *Sandene* és munkatársai (2005), illetve *Horkay* és munkatársai (2006) kutatási eredményeivel egyezően, a szocio-ökonómiai státusz, valamint az induktív gondolkodás teszten adódó papír-ceruza és számítógépes teljesítménykülönbség mértéke között nem igazolódott szignifikáns összefüggés. A problémamegoldó gondolkodás vizsgálata során a szocio-ökonómiai státusz kismértékű kapcsolatot mutatott a médiahatás mértékével.

A nemek szerinti bontásban végzett médiahatás-vizsgálatok eredményei szerint a papír-ceruza mérőeszközök digitalizálása és online alkalmazása hatást gyakorolt mind a fiúk, mind a lányok átlagteljesítményére. A fiúk és a lányok egyaránt magasabb eredményt értek el papír-ceruza formátumban, mint online.

A nemek közötti különbségek feltárása során igazolódott, hogy minden mérési területen a fiúk és lányok papír-ceruza és számítógépes átlagteljesítmény-különbsége megegyezett a teljes teszt és a résztesztek szintjén is. Ezáltal egymáshoz képes sem a fiúk, sem a lányok nem kerültek előnybe vagy hátrányba a számítógépes tesztelés bevezetése által.

A számítógéphez kapcsolódó változók hatásvizsgálata során a naponta számítógépezéssel töltött órák száma gyenge kapcsolatot mutatott a médiahatás mértékével a teljes teszt és a résztesztek szintjén is. Azonban ebből nem következett, hogy akinek van otthon számítógépe – ezáltal több lehetősége a számítógépezésre – az előnybe került volna a számítógépes tesztelés során az otthoni számítógéppel nem rendelkező családok diákjaival szemben. A további, számítógépes tapasztalathoz tartozó változók (pl. a tanuló mióta, milyen gyakran használnak számítógépet és milyen tevékenységeket végeznek számítógépezés

közben) nem függték össze a diákok papír-ceruza és számítógépes teljesítmény-különbségével (médiahatás), így az eltérő számítógépes tapasztalattal rendelkező diákok egyformán reagáltak a tesztmédiium megváltozására.

Feltételezésünk igazolódott (*Lissitz, Jiao, Xie, Li és Kang, 2011-es munkájával egyezően*), miszerint a számítógépes attitűd három alsóskálájának (a számítógép fontossága, a számítógépezés élvezete és a számítógépes szorongás) nincs hatása a médiahatás mértékére egyik mérési területen sem. A számítógépes attitűdtől eltérően a számítógépes ismeret kis mértékben befolyásolta a médiahatás mértékét problémamegoldás és induktív gondolkodás terén.

A teljesítmény alapú mintafelosztással végzett médiahatás-vizsgálat eredményei (*Pomplun és mtsai, 2006 konklúziójával egyezően*) igazolták, hogy a különböző képességszintű tanulók, eltérő módon reagálnak a tesztmédiium megváltozására. Ugyanakkor várakozásainkkal, illetve *Pomplun és munkatársai (2006)* eredményeivel ellentétben a papír alapon alacsonyabb teljesítményt nyújtó diákok számítógépes környezetben magasabb teljesítményt értek el mindhárom mérési területen (induktív gondolkodás, szövegértés és problémamegoldás), mint papíron. A nyomtatott teszten közel átlagos teljesítményt nyújtó diákok teljesítménye részben médiafüggetlennek bizonyult, részben kismértékű médiahatást mutatott. A médiahatás irányát tekintve az átlagos képességű diákok kismértékben jobban teljesítettek tradicionális környezetben, mint számítógépen. Ugyanakkor a jó képességű diákok hátrányba kerültek a papír alapú tesztek digitalizált változatának felvétele során mindhárom mérési területen (*Tóth, Molnár, Hódi és Csapó, 2011*).

Az itemjellemzők és a médiahatás összefüggései

Az itemjellemzők közül megvizsgáltuk, hogy az item típusa, az item nehézsége és az item megoldásához elvárt tevékenység hogyan függ össze a médiahatás mértékével. A különböző típusú itemek eltérő működésének vizsgálata során igazolódott, hogy mind a feleletalkotó, mind a feleletválasztó itemeken szignifikáns itemszintű médiahatás adódott; mindkét feladattípus megoldása papír-ceruza környezetben könnyebb, mint számítógépen. Továbbá a feleletválasztó és feleletalkotó itemek átlagos itemszintű médiahatás-mértékei között nem volt szignifikáns különbség. A feladattípus és médiahatás összefüggés-vizsgálatának eredményei rávilágítottak, hogy a rövid szöveges válaszok begépelése, ezáltal a billentyűzet használatában való jártasság, önmagában nem okozott médiahatást a disszertációban alkalmazott teszteken (*Russel, 1999 és Russel és Haney, 1997* eredményeivel ellentétben), hiszen mind a feleletválasztó, mind a feleletalkotó feladatokon jelentős mértékű, közel azonos médiahatás igazolódott.

Az itemnehézségi mutatókkal végzett elemzések eredményei szerint a papír-ceruza formában felvett teszt itemnehézségi mutatói összefüggtek az itemszintű médiahatás mértékével mindhárom mérési területen, így az itemek nehézsége a vizsgált területtől függetlenül magyarázóerővel bír az eltérő itemműködés szempontjából.

Az itemeket aszerint osztályoztuk, hogy megoldásukhoz szükséges-e számolás, majd megvizsgáltuk, hogy az item megoldásához elvárt tevékenység befolyásolja-e az itemszintű médiahatás mértékét. A három mérési terület eredményeit felhasználó elemzések alapján a

számolást nem igénylő feladatok médiafüggetlen működése igazolódott. Ugyanakkor a számolást igénylő feladatokon jelentős volt a médiahatás.

A teststimulus vizsgálata során a problémamegoldás- és a szövegértésteztek különböző dokumentumtípusain (pl. táblázat, ábra) adódó átlagos teljesítménykülönbségeket elemeztük. Eredményeink szerint a különböző dokumentumtípusokhoz tartozó feladatokon eltérő mértékű volt a médiahatás. Szövegértés terén a legkisebb médiahatás a táblázatban prezentált nem folyamatos szövegeken adódott. Ezzel szemben a problémamegoldó gondolkodás vizsgálatában a táblázatban foglalt információkhoz kapcsolódó feladatokon volt a legnagyobb mértékű a médiahatás. E két eredmény eltérését magyarázhatja, hogy míg szövegértés esetén a diákoknak információ-visszakeresés volt a feladata, addig problémamegoldás terén a diákok először táblázatból kerestek ki információkat (számokat), majd számolniuk kellett a megtalált adatokkal. Az eredmények alapján arra a megállapításra jutottunk, hogy a 6. osztályos diákok közel hasonló hatékonysággal kerestek vissza táblázatba foglalt információkat online és papír-ceruza környezetben, ugyanakkor a feladatmegoldáshoz elvárt tevékenység (pl. számolás elvégzése) nehezítette a feladatmegoldást és magasabb médiahatást eredményezett. A szövegértés és problémamegoldás különböző dokumentumtípusain tapasztalt eltérő mértékű médiahatás arra enged következtetni, hogy a dokumentum típusa önmagában nem határozza meg a médiahatás mértékét.

Összességében megállapítottuk, hogy az item típusa és a dokumentumtípus változó önmagában nem áll kapcsolatban a médiahatás mértékével, ugyanakkor az itemnehézség és a feladatmegoldáshoz elvárt tevékenység (számolást igényel/számolást nem igényel) magyarázóerővel bír a médiahatás mértékét tekintve. Annak meghatározására, hogy e változók közül melyik a leginkább domináns, C4.5 adatbányászati algoritmust használtunk, és az itemek 68 százalékáról meg tudtuk mondani a vizsgálatba bevont változók (a mérési terület, a mintakialakítás módja, a teszt- és itemjellemzői) alapján, hogy a két tesztkörnyezetben eltérően működik-e vagy sem. A legdominánsabb változó a mérési terület. Az itemnehézség mind problémamegoldás, mind induktív gondolkodás terén meghatározó szerepet töltött be. Továbbá az induktív gondolkodás vizsgálatokban a mintakialakítás módja is releváns volt az eltérő itemműködés meghatározásában.

AZ EMPIRIKUS VIZSGÁLAT JELENTŐSÉGE, TOVÁBBI KUTATÁSI LEHETŐSÉGEK

Az empirikus adatokon nyugvó elemzések egyik legfontosabb megállapítása, hogy a papír-ceruza tesztek és azok számítógépes adaptációjából származó adatok egyforma megbízhatósággal alkalmazhatók a pedagógiai következtetések levonására. Komparatív vizsgálataink eredményei megerősítik számos nemzetközi vizsgálat konklúzióját (*Choi és Tinkler, 2002; Pomplun és mtsai, 2006; Schöders és Wilhelm, 2011*), mely szerint a különböző mérési területeken szignifikáns médiahatás igazolódik a teszt különböző szintjein (teljes teszt, részteszt és item). Ezért, ha áttérünk online tesztelésre és használni szeretnénk a korábban papír-ceruza formában gyűjtött adatainkat (pl. a Szegedi Longitudinális Programban), akkor szükség van egy olyan lineáris transzformációra, amellyel megfeleltethetjük az egyik tesztkörnyezetben felvett adatainkat egy másik tesztkörnyezetben felvett adatokkal, hogy minimalizáljuk a tesztkörnyezet megváltozásából fakadó teljesítménybeli eltéréseket.

Empirikus vizsgálatunk legnagyobb jelentőségét abban látjuk, hogy a lineáris transzformációhoz a disszertáció három dimenzió (minta, mérési terület jellemzői és itemszintű vizsgálatok) kérdéskörét járta körbe és fejtette ki az egyes dimenziók komponenseinek hatását a papír-ceruza és számítógépes teljesítménykülönbségek tekintetében.

Összességében a disszertációban bemutatott vizsgálatok arra törekedtek, hogy a diákok médiafüggő teljesítményének több mérési területre érvényes okait feltárják. Ezért olyan szempontokat vontunk be elemzéseinkbe, amelyek több vagy akár minden mérési területen megjelennek. Ennek előnye, hogy az eredmények/hasonlóságok általánosíthatók, kiindulási alapot szolgáltatnak további vizsgálatokhoz. Hátránya azonban, hogy az egyes mérési területek elméleti kérdéseire nem ad választ. Ezért számos további kutatási feladat megfogalmazható. Például a problémamegoldó gondolkodás vizsgálata során, a problémák elméleti modell alapján történő csoportosítása után, érdemes lenne megvizsgálni, hogy a különböző típusú problémák megoldásában egyező mértékű médiahatás adódik-e. A vizsgálatok módszertani jellemzőire fókuszálva további feladatok lehetnének például, hogy meghatározzuk fókuszcsoportos interjú módszerrel, mi áll annak hátterében, hogy az eltérő képességű tanulók különböző módon reagálnak a tesztmédiium megváltozására. A médiahatás-vizsgálatot érdemes lenne szélesebb életkori intervallumban is elvégezni, hogy lássunk a különböző korú diákok teljesítményére hogyan hat a számítógépes tesztelés bevezetése. Ezek a megfogalmazott további vizsgálatok gazdagíthatják a médiahatás-vizsgálatok hazai és nemzetközi eredményeit.

Irodalom

- Asuni, N. (2008): Quality Features of TCEExam, an Open-Source Computer-Based Assessment Software. 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, Luxemburg, 58–63.
- Bartram, D. és Coyne, L. (2005): *International Guidelines on Computer-Based and Internet Delivered Testing*. International Test Commission, Punta Gorda.
- Bennett, R.E. (2003). *Online assessment and the comparability of score meaning*. Educational Testing Service, Princeton. 2015.02.15-i megtekintés, <http://www.ets.org/>
- Bennett, R. E., Braswell, J., Oranje, A., Sandene, B., Kaplan, B. és Yan, F. (2008): Does it matter if I take my mathematics test on computer? A second empirical study of mode effects in NAEP. *Journal of Technology, Learning, and Assessment*, 6. 9. sz. 2015.02.15-i megtekintés, <http://www.jtla.org>
- 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.
- Choi, S. W. és Tinkler, T. (2002): Evaluating comparability of paper-and-pencil and computer-based assessment in a K-12 setting. 2015.02.15-i megtekintés, <http://www.docstoc.com/docs/102661762/Evaluating-comparability-of-paper-and-pencil-and-computer--based>

- 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.
- Csapó Benő (2007): Hosszmetszeti felmérések iskolai kontextusban – az első átfogó magyar iskolai longitudinális kutatási program elméleti és módszertani keretei. *Magyar Pedagógia*, **107**. 4. sz. 321–355.
- Csapó Benő (1994): Az induktív gondolkodás fejlődése. *Magyar Pedagógia*, **94**. 1-2. sz. 53–80.
- Csapó, B., Ainley, J., Bennett, R., Latour, T. és Law, N. (2012): Technological issues of computer-based assessment of 21st century skills. In: Griffin, P., McGaw, B. és Care, E. (szerk.): *Assessment and teaching of 21st century skills*. Springer, New York. 143–230.
- Csapó Benő, Molnár Gyöngyvér, Pap-szigeti Róbert és Tóth Krisztina (2009): A mérés-értékelés új tendenciái, a papír alapú teszteléstől az online tesztelésig. In: Perjés István és Kozma Tamás (szerk.): *Új kutatások a neveléstudományokban. Hatékony tudomány, pedagógiai kultúra, sikeres iskola*. Magyar Tudományos Akadémia, Budapest. 99–108.
- Csapó Benő, Molnár Gyöngyvér és Tóth Krisztina (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, Luxemburg. 113–118.
- Csapó Benő, Molnár Gyöngyvér és Tóth Krisztina (2008): A papíralapú teszteléstől a számítógépes adaptív tesztelésig. A pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 3-4. sz. 3–16.
- Devosa Iván és Vízvári Gergely (2011): Recent Developments of TAO CAPI. Előadás: 3rd Szeged Workshop on Educational Evaluation. Szeged, 2011. április 27–28.
- 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*. European Commission Joint Research Centre, Ispra. 64–67.
- Gaskill, J. és Marshall, M. (2007): *Comparisons between paper and computer-based tests: Foundation Skills Assessment 2001–2006 data*. Society for the Advancement of Excellence in Education, Kelowna. 2015.02.15-i megtekintés, <http://www.sae.ca/pdfs/038.pdf>.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. és Klieme, E. (2014): The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, **106**. 3. sz. 608–626.
- Gu, L., Drake, S. és Wolfe, E. W. (2006): Differential item functioning of GRE mathematics items across computerized and paper-and-pencil testing media. *Journal of Technology, Learning, and Assessment*, **5**. 4. sz. 2015.02.15-i megtekintés, Journal of Technology, Learning, and Assessment [on-line] <http://www.jtla.org>
- Higgins, J., Russell, M. és Hoffmann, T. (2005): Examining the effect of computer-based passage presentation on reading test performance. *Journal of Technology, Learning, and Assessment*, **3**. 4. sz. 2015.02.15-i megtekintés, <http://www.jtla.org>
- Horne, J. (2007): Gender differences in computerised and conventional educational tests. *Journal of Computer Assisted Learning*, **23**. 1. sz. 47–55.

- 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. 2015.02.15-i megtekintés, <http://www.jtla.org>
- Hódi Ágnes és Tóth Krisztina (2010): Assessing Reading Skills in Printed and Computerized Environment. Előadás: Junior Researchers of the European Association for Research on Learning and Instruction (JURE) Conference. Előadás. Frankfurt am Main, Németország, 2010. július 19-22.
- Hülber László és Molnár Gyöngyvér (2013): Papír és számítógép alapú tesztelés nagymintás összehasonlító vizsgálata matematika területén, 1-6. évfolyamon. *Magyar Pedagógia*, **113**. 4. sz. 243–263.
- Jeong, H. (2014): A comparative study of scores on computer-based tests and paper-based tests. *Behaviour and Information Technology*, **33**. 4. sz. 410–422.
- Jones, T. és Clarke, V. A (1995): Diversity as a determinant of attitudes: a possible explanation of the apparent advantage of single-sex settings. *Journal of Educational Computing Research*, **12**. 1. sz. 51–64.
- Johnson, M., és Green, S. (2006): On-line mathematics assessment: The impact of mode on performance and question answering strategies. *Journal of Technology, Learning, and Assessment*, **4**. 5. sz. 2015.02.15-i megtekintés, <http://www.jtla.org>
- Kingston, N. M. (2009): Comparability of computer- and paper-administered multiple-choice tests for K-12 Populations: A Synthesis. *Applied Measurement in Education*, **22**. 1. sz. 22–37.
- Knezek, G. A., Christensen, R. W., Miyashita, K. T. és Ropp, M. M. (2000): *Instruments for Assessing Educator Progress in Technology Integration*. Institute for the Integration of Technology into Teaching and Learning University of North Texas, Denton, Texas, USA. 2015.03.15-i megtekintés, <http://www.iittl.unt.edu/pt3II/book1.htm>
- Leeson, H. V. (2006): The mode effect: A literature review of human and technological issues in computerized testing. *International Journal of Testing*, **6**. 1. sz. 1–24.
- Lissitz, H. J., Chao X., Ming L., Yoon J. K. (2011): Student Characteristics and CBT Performance: An Overview of the Literature by MARCES. 2015.02.15-i megtekintés, MARCES, <http://marces.org/current/Student%20Characteristics%20and%20CBT%20Performance.pdf>
- Lottridge, S. M, Nicewander, W. A., Schulz, E. M. és Mitzel, H. C. (2008): *Comparability of Paper-based and Computer-based Tests: A Review of the Methodology*. Pacific Metrics Corporation, Monterey. 2015.02.16-i megtekintés, Pacific Metric, Corporation, <http://www.pacificmetrics.com/research/white-papers/>
- Lundy, J. J. (2008): *Assessing psychometric equivalence of paper-and-pencil and interactive voice response (IVR) modes of administration for the EQ-5D and the QLQ-C30*. 2015.02.18-i megtekintés, The University of Arizona, http://arizona.openrepository.com/arizona/bitstream/10150/193902/1/azu_etd_10040_sip1_m.pdf
- Molnár Gyöngyvér (2010): Papír- és számítógép alapú tesztelés összehasonlító vizsgálata problémamegoldó környezetben. In: Perjés István és Kozma Tamás: *Új Kutatások a Neveléstudományokban*. Aula Kiadó, Corvinus Egyetem, Budapest. 135–144.
- Molnár Gyöngyvér, Tóth Krisztina és Csapó Benő (2011): The relationship between item type, students' characteristics and media-effect in CBA. Előadás: 14th European

- Conference for the Research on Learning and Instruction. Exeter, United Kingdom, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 4–5.)
- Paek, P. (2005): *Recent trends in comparability studies*. (Pearson Educational Measurement Research Report 05-05). Pearson Educational Measurement, Upper Sadle River.
- Poggio, J., Glasnapp, D. R., Yang, X. és Poggio, A. J. (2005): A comparative evaluation of score results from computerized and paper and pencil mathematics testing in a large scale state assessment program. *Journal of Technology, Learning, and Assessment*, **3**. 6. sz. 2015.02.15-i megtekintés, <http://www.jtla.org>
- Pomplun, M., Ritchie, T. és Custer, M. (2006): Factors in paper-and-pencil and computer reading score differences at the primary grades. *Educational Assessment*, **11**. 2. sz. 127–143.
- Puhan, P., Boughton, K. és Kim, S. (2007): Examining differences in examinee performance in paper and pencil and computerized testing. *Journal of Technology, Learning, and Assessment*, **6**. 3. sz. 2011.03.15-i megtekintés, <http://ejournals.bc.edu/ojs/index.php/jtla/issue/view/172>
- Russell, M. (1999): Testing on computers: A follow-up study comparing performance on computer and on paper. *Education Policy Analysis Archives*, **7**. 20. sz. 2013.03.02-i megtekintés, <http://epaa.asu.edu/epaa/v7n20>.
- Russell, M. és Haney, W. (1997). Testing writing on computers: An experiment comparing student performance on tests conducted via computer and via paper-and-pencil. *Education Policy Analysis Archives*, **5**. 3. sz. Paper 186.
- 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*. (NCES 2005-457). U.S. Department of Education, National Center for Education Statistics, Washington.
- Scalise, K. és Gifford, B. (2006): Computer-based assessment in e-learning: A framework for constructing “Intermediate Constraint” questions and tasks for technology platforms. *The Journal of Technology, Learning, and Assessment*, **4**. 6. sz. 3–44.
- Schroeders, U. és Wilhelm, O. (2011): Equivalence of reading and listening comprehension across test media. *Educational and Psychological Measurement*, **71**. 5. sz. 849–869.
- Tóth Krisztina (2014): Papír-ceruza és számítógépes induktív gondolkodás-tesztek reliabilitásának összehasonlító vizsgálata. Előadás: XII. Pedagógiai Értékelési Konferencia. Szeged, 2014. május 1-3. (Tartalmi összefoglalók: 107.)
- Tóth Krisztina (2010): Papír-ceruza és számítógépes tesztek eredményeinek összehasonlító vizsgálata. In: Vajda Zoltán (szerk.): *Bölcsész-műhely 2009*. JATEPress, Szeged. 125–136.
- Tóth Krisztina és Hódi Ágnes (2011): Comparing Students Reading Comprehension Achievement along Different Text-types in Paper-based and Computerized Environment. Előadás: American Educational Research Association - Annual Meeting. New Orleans, 2011. április 8-12.
- Tóth Krisztina és Hódi Ágnes (2010): Olvasási képesség mérése számítógépes környezetben. In: Perjés István és Kozma Tamás (szerk.): *Új Kutatások a Neveléstudományokban*. Aula Kiadó, Corvinus Egyetem, Budapest. 145–155.
- Tóth Krisztina, Molnár Gyöngyér és Csapó Benő (2009): Online Assessment of Reasoning Skills. Előadás: 13th Biennial European Conference for the Research on Learning and Instruction. Amsterdam, Hollandia, 2009. augusztus 25-19.
- Tóth Krisztina Molnár Gyöngyér, Hódi Ágnes és Csapó Benő (2011): Examining Media-Effect among Subgroups of Students with Different Ability Levels. Előadás: 14th

- European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 630-632.)
- Tóth Krisztina, Molnár Gyöngyvér, Thibaud Latour és Csapó Benő (2011): Az online tesztelés lehetőségei és a TAO platform alkalmazása. *Új Pedagógiai Szemle*, **61**. 1–2–3–4–5. sz. 8–22.
- Tóth Krisztina, Molnár Gyöngyvér, Sascha Wüstenberg, Samuel Greiff és Csapó Benő (2011): Measuring adults' dynamic problem solving competency. Előadás: 14th European Conference for the Research on Learning and Instruction. Exeter, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 1421–1422.)
- Tóth, K., Rölke, H. és Goldhammer, F. (2012a): Investigating Test-taking behaviour in simulation-based assessment – Visual data exploration. Előadás: 4th International Conference on Education and New Learning Technologies, Barcelona, 2012. július 2-4.
- Tóth, K., Rölke, H. és Goldhammer, F. (2012b): Educational Process Mining- Clustering Students' Test-taking Behaviour in Internet-based Simulations. Előadás: X. Pedagógiai Értékelési Konferencia, Szeged, 2012. április 26-28. (Tartalmi összefoglalók: 44.)
- Tóth K., Rölke, H., Greiff, S. és Wüstenberg, S. (2014): Discovering Students' Complex Problem Solving Strategies in Educational Assessment. In: Stamper, J., Pardos, Z., Mavrikis, M. és McLaren, B. M. (szerk.): *Proceedings of the 7th International Conference on Educational Data Mining*. 225–228.
- Tóth, K., Rölke, H., Naumann, J. és Goldhammer, F. (2012): Analyse des Problemlöseverhaltens in simulierten Hypertext-Umgebungen. Előadás: 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Bielefeld, 2012. szeptember 22-27.
- Tóth, K., Wüstenberg, S., Rölke, H. és Greiff, S. (2012): Prediction of students' performance on test taking processes in Complex Problem Solving. Előadás: 30th International Congress of Psychology. Cape Town, South Africa, 2012. július 22-27.
- Wang, S., Jiao, H., Young, M., Brooks, T. és Olson, J. (2008): Comparability of computer-based and paperand-pencil testing in K-12 reading assessments: A meta-analysis of testing mode effects. *Educational and Psychological Measurement*, **68**. 1. sz. 5–24.
- Wang, S., Jiao, H., Young, M., Brooks, T. és Olson, J. (2007): A meta-analysis of testing mode effects in grade K-12 mathematics tests. *Educational and Psychological Measurement*, **67**. 2. sz. 219–238.

A disszertáció témaköréhez kapcsolódó publikációk

- Csapó Benő, Molnár Gyöngyvér, Pap-szigeti Róbert és Tóth Krisztina (2009): A mérés-értékelés új tendenciái, a papír alapú teszteléstől az online tesztelésig. In: Perjés István és Kozma Tamás (szerk.): *Új kutatások a neveléstudományokban. Hatékony tudomány, pedagógiai kultúra, sikeres iskola.* Magyar Tudományos Akadémia, Budapest. 99–108.
- Csapó Benő, Molnár Gyöngyvér és Tóth Krisztina (2010): Implementing an Online Formative Assessment System: From Paper-Based to Computer-Based Testing. Előadás: American Educational Research Association - Annual Meeting, Denver, USA, 2010. április 29. - május 4.
- Csapó Benő, Molnár Gyöngyvér és Tóth Krisztina (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, Luxemburg. 113–118.
- Csapó Benő, Molnár Gyöngyvér és Tóth Krisztina (2008): A papíralapú teszteléstől a számítógépes adaptív tesztelésig. A pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 3-4. sz. 3–16.
- Goldhammer, F., Naumann, J., Stelter, A., Tóth, K., Rölke, H. és Klieme, E. (2014): The time on task effect in reading and problem solving is moderated by task difficulty and skill: Insights from a computer-based large-scale assessment. *Journal of Educational Psychology*, **106**. 3. sz. 608–626.
- Hódi Ágnes és Tóth Krisztina (2010): Assessing Reading Skills in Printed and Computerized Environment. Előadás: Junior Researchers of the European Association for Research on Learning and Instruction (JURE) Conference. Frankfurt am Main, Németország, 2010. július 19-22.
- Hódi Ágnes és Tóth Krisztina (2009): Olvasási képesség mérése számítógépes környezetben. Előadás: IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19-21. (Tartalmi összefoglalók: 96–97.)
- Molnár Gyöngyvér és Tóth Krisztina (2008): A számítógép- és papír-alapú tesztelés eredményeinek összehasonlító vizsgálata 5. évfolyamon. Előadás: VIII. Országos Neveléstudományi Konferencia, Budapest, 2008. november 13–15. (Tartalmi összefoglalók: 131.)
- Molnár Gyöngyvér, Tóth Krisztina és Csapó Benő (2011): Comparing Paper-Based and Computer-Based Testing in the First Grade. Előadás: American Educational Research Association - Annual Meeting, New Orleans, Louisiana, USA, 2011. április 8-12.
- Molnár Gyöngyvér, Tóth Krisztina és Csapó Benő (2011): The relationship between item type, students' characteristics and media-effect in CBA. Előadás: 14th European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 4–5.)
- Molnár Gyöngyvér, Tóth Krisztina és Tóth Edit (2010): Developing Online Diagnostic Assessment - Experiences of a Large-scale National Case Study in Public Education in Hungary. Előadás: EDEN Workshop, Budapest, 2010. október 24-27.
- Tóth Krisztina (2014): Papír-ceruza és számítógépes induktív gondolkodás-tesztek reliabilitásának összehasonlító vizsgálata. Előadás: XII. Pedagógiai Értékelési Konferencia. Szeged, 2014. május 1-3. (Tartalmi összefoglalók: 107.)

- Tóth Krisztina (2010a): Adatbányászat a neveléstudomány területén (Educational Data Mining). Előadás: VIII. Pedagógiai Értékelési Konferencia. Szeged, 2010. április 16-17. (Tartalmi összefoglalók: 93.)
- Tóth Krisztina (2010b): Comparing Paper- and Computer-based Test-Taking Processes. Előadás: 2nd Szeged Workshop on Educational Evaluation. Szeged, 2010. április 14-15.
- Tóth Krisztina (2010c): Papír-ceruza és számítógépes tesztek eredményeinek összehasonlító vizsgálata. In: Vajda Zoltán (szerk.): *Bölcsészmuhely 2009*. JATEPress, Szeged. 16–27.
- Tóth Krisztina (2009a): The Szeged Experience. Előadás: Szeged Workshop on Educational Evaluation. Szeged, 2009. április 27-28.
- Tóth Krisztina (2009b): Számítógépes és papír-ceruza tesztek eredményeinek összehasonlító vizsgálata háttérváltozók alapján. Előadás: IX. Országos Neveléstudományi Konferencia, Veszprém, 2009. november 19-21. (Tartalmi összefoglalók: 97–98.)
- Tóth Krisztina (2009c): Az elektronikus tesztelés és tesztvégrehajtás fajtái. Előadás: VII. Pedagógiai Értékelési Konferencia, Szeged, 2009. április 24-25. (Tartalmi összefoglalók: 46.)
- Tóth Krisztina (2008): Az online teszteléssel kapcsolatos attitűdök és eredmények a háttérváltozók tükrében. Előadás: VIII. Országos Neveléstudományi Konferencia, Budapest, 2008. november 13–15.
- Tóth Krisztina és Hódi Ágnes (2013): A mérőeszköz-bővítéstől a tesztelési folyamat vizsgálatáig: számítógépes tesztelés nagymintás nemzetközi vizsgálatokban. *Iskolakultúra*, 9. sz. 75–88.
- Tóth Krisztina és Hódi Ágnes (2011): Comparing Students Reading Comprehension Achievement along Different Text-types in Paper-based and Computerized Environment. Előadás: American Educational Research Association - Annual Meeting. New Orleans, 2011. április 8-12.
- Tóth Krisztina és Hódi Ágnes (2010a): Olvasási képesség mérése számítógépes környezetben. In: Perjés István és Kozma Tamás (szerk.): *Új Kutatások a Neveléstudományokban*. Aula Kiadó, Corvinus Egyetem, Budapest. 145–155.
- Tóth Krisztina és Hódi Ágnes (2010b): Olvasási képesség fejlettségének mérése online környezetben kisiskolás tanulók körében. Előadás: VIII. Pedagógiai Értékelési Konferencia. Szeged, 2010. április 16-17. (Tartalmi összefoglalók: 121.)
- Tóth Krisztina, Molnár Gyöngyvér és Csapó Benő (2009): Online Assessment of Reasoning Skills. Előadás: 13th Biennial European Conference for the Research on Learning and Instruction. Amsterdam, Hollandia, 2009. augusztus 25-19.
- Tóth Krisztina, Molnár Gyöngyvér és Csapó Benő (2008): A számítógépes tesztelés lehetőségei. Előadás: VI. Pedagógiai Értékelési Konferencia, Szeged, 2008. április 11-12. (Tartalmi összefoglalók: 84.)
- Tóth Krisztina Molnár Gyöngyvér, Hódi Ágnes és Csapó Benő (2011): Examining Media-Effect among Subgroups of Students with Different Ability Levels. Előadás: 14th European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 630-632.)
- Tóth Krisztina, Molnár Gyöngyvér, Thibaud Latour és Csapó Benő (2011): Az online tesztelés lehetőségei és a TAO platform alkalmazása. *Új Pedagógiai Szemle*, 61. 1–2–3–4–5. sz. 8–22.
- Tóth Krisztina, Molnár Gyöngyvér, Sascha Wüstenberg, Samuel Greiff és Csapó Benő (2011): Measuring adults' dynamic problem solving competency. Előadás: 14th European

- Conference for the Research on Learning and Instruction. Exeter, 2011. augusztus 30-szeptember 3. (Tartalmi összefoglalók: 1421–1422.)
- Tóth, K., Rölke, H. és Goldhammer, F. (2012a): Investigating Test-taking behaviour in simulation-based assessment – Visual data exploration. Előadás: 4th International Conference on Education and New Learning Technologies, Barcelona, 2012. július 2-4.
- Tóth, K., Rölke, H. és Goldhammer, F. (2012b): Educational Process Mining- Clustering Students' Test-taking Behaviour in Internet-based Simulations. Előadás: X. Pedagógiai Értékelési Konferencia, Szeged, 2012. április 26-28. (Tartalmi összefoglalók: 44.)
- Tóth K., Rölke, H., Greiff, S. és Wüstenberg, S. (2014): Discovering Students' Complex Problem Solving Strategies in Educational Assessment. In: Stamper, J., Pardos, Z., Mavrikis, M. és McLaren, B. M. (szerk.): *Proceedings of the 7th International Conference on Educational Data Mining*. 225–228.
- Tóth, K., Rölke, H., Naumann, J. és Goldhammer, F. (2012): Analyse des Problemlöseverhaltens in simulierten Hypertext-Umgebungen. Előadás: 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Bielefeld, 2012. szeptember 22-27.
- Tóth, K., Wüstenberg, S., Rölke, H. és Greiff, S. (2012): Prediction of students' performance on test taking processes in Complex Problem Solving. Előadás: 30th International Congress of Psychology. Cape Town, South Africa, 2012. július 22-27.