

R. Tóth Krisztina<sup>1</sup> – Molnár Gyöngyvér<sup>2</sup> – Thibaud Latour<sup>3</sup> – Csapó Benő<sup>2</sup>

## Az online tesztelés lehetőségei és a TAO platform alkalmazása

*Az elmúlt évtizedben látványosan fejlődött a pedagógiai értékelés gyakorlata, amit elsősorban a nagyléptékű nemzetközi felmérések (PISA, TIMSS, PIRLSS) rendszeressé válása és a nemzeti értékelési rendszerek (nálunk az országos kompetenciamérések) kiépülése fémjelzett. A papíralapú felmérések fejlődési lehetőségei azonban nagyrészt kimerültek. A mérőeszközök előállításának és az adatok rögzítésének rendszeres költségei, a logisztikai problémák, az eredmények visszajelzésének ideje szoros korlátokat jelentenek. Éppen ezért világszerte kiterjedt kutatások folynak a technológiaalapú mérés-értékelés megvalósítása, és az új generációs értékelési módszerek kidolgozása érdekében.*

A kérdés fontosságát mutatja, hogy az Európai Unió kiemelten támogatja a tesztelési folyamat számítógépes alapokra helyezését (például workshopok szervezésével, lásd SCHEUERMANN és BJÖRNSSON 2009; SCHEUERMANN és PEREIRA 2008). Hasonló célokkal indított három technológiai világcég (Cisco, Intel és Microsoft) összefogásával egy globális kutatási projektet (Assessment and Teaching of 21st Century Skills – A 21. századi készségek mérése és tanítása – lásd KOZMA 2009), melynek első fázisa 2010-ben zárult (MCGAW és GRIFFIN 2011).

Ma már a számítógépes online tesztelésnek minden egyes eleme rendelkezésre áll, és számos kísérleti program van folyamatban; a rendszerszerűen, nagymintás felmérésekre használható megoldások kidolgozásához még nagyon sok kutató-fejlesztő munkára van szükség. Egy széles körben alkalmazható rendszer kiépítéséhez, feladatokkal való feltöltéséhez általában öt-hét éves munkára van szükség, és a technológia még nem áll elég hosszú ideje rendelkezésre ahhoz, hogy ilyen rendszerek kifejlesztésére sor kerülhetett volna. Az online tesztrendszerek fejlesztése speciális szakértelmet és jelentős forrásokat igényel, ami ugyancsak korlátozza megvalósíthatóságukat. Ugyanakkor az online tesztelés annyi előnyt kínál, ami indokolja e fejlesztési folyamat elindítását, illetve felgyorsítását.

A technológiaalapú mérés-értékelés fontosabb hazai és nemzetközi implementációival, a technológiaalapú tesztelés egyes típusaival, továbbá a technológiaalapú tesztelésnek a mérési-értékelési folyamat fejlesztésében játszott szerepével korábbi tanulmányainkban

1 Deutsches Institut für Internationale Pädagogische Forschung

2 Szegedi Tudományegyetem Neveléstudományi Intézet

3 Centre de Recherche Public Henri Tudor

részletesebben foglalkoztunk (lásd MOLNÁR 2010a; CSAPÓ, MOLNÁR és R. TÓTH 2008; illetve CSAPÓ, MOLNÁR, PAP-SZIGETI és R. TÓTH 2009). Ebben a tanulmányban a technikai megvalósítás egy konkrét platformját, a TAO-rendszert mutatjuk be, és felvázoljuk eddigi hazai és fontosabb nemzetközi alkalmazásait.

## AZ ONLINE TESZTELÉS ALKALMAZÁSAI – LEHETŐSÉGEK ÉS KORLÁTOK

A technológiaalapú tesztelésnek számos különböző megoldása létezik. A technológiák csoportosításának egyik szempontja az, miként jut el a teszt maga a tesztet megoldó személyekhez. Ezt az angol „*delivery methods*” kifejezés analógiájára *eljuttatási módszernek* nevezzük. Az eljuttatási módszer és a tesztelés funkciója között szoros kapcsolat van: a különböző funkciójú adatgyűjtés más-más technológiai megoldást igényel, illetve az egyes eljuttatási technológiák korlátai különböző alkalmazásokat tesznek lehetővé (CSAPÓ, LATOUR, BENNETT, AINLEY és LAW 2011). A jelenleg alkalmazható megoldásokat érdemes részletesebben áttekintenünk, mert az egyes funkcióra jól működő rendszerek téves várakozásokat kelthetnek más alkalmazásokkal kapcsolatban, vagy megfordítva, bizonyos célokra (ma még) alkalmatlan eljuttatási módszerek kudarcai elzárhatják a fejlesztés lehetőségeit az adekvát felhasználásoktól is.

A leggyakoribb eljuttatási megoldások közé tartozik a *vizsgaközpontba telepített számítógépes rendszer*. A tesztelt személyeknek egy megfelelően biztosított helyszínen kell eljutniuk, ahol megoldják a teszteket. Ezzel a módszerrel egyszerűbben lehet védeni mind a titkosított feladatokat, mind pedig a személyes adatokat, továbbá biztosítani a tesztelés standard feltételeit. Erre a megoldásra van szükség, ha a tesztnek komoly tétje van (pl. különböző vizsgák), és (ma még) akkor is, ha a teszt megoldása közben nagy tömegű információt kell mozgatni. Az internet sávszélessége ugyanis korlátozza nagy tömegű (kép, hang, esetleg videó) információ egyidejű és gyors eljuttatását sok felhasználóhoz. A központosított tesztelő rendszereken lehet a legtöbb korszerű multimédiás alkalmazást megvalósítani, ugyanakkor némi korlátozással hasonló célokra lehet használni az iskolákban telepített, egységesen berendezett számítógépes laboratóriumokat, ilyen esetben azonban a fenntartás és működtetés tetemes költséget, képzett személyzetet igényel.

Egy másik gyakran alkalmazott eljuttatási megoldás a tesztek hordozható médiumon (CD, DVD, USB drive, memóriakártya stb.) való továbbítása a tesztelt személyekhez. Ebben az esetben az adatbiztonság már csak korlátozott mértékben érvényesül, így főleg akkor alkalmazható, ha egyedi adatfelvételre kerül sor, és a tesztet megoldó személyek számára kisebb a tesztelés tétje. Ezt a megoldást alkalmazta korábban a PISA is a 2006-os számítógépes természettudomány tesztelés és a 2009-es számítógépes szövegértés vizsgálat esetében. Hordozható médium alkalmazása esetén a felmérés helyszínén csak számítógépekre van szükség, ami lehetővé teszi nagyobb információigényű feladatok közvetítését is. Elvileg így nagyobb információigényű feladatokat is közvetíteni lehet.

Logisztikai szempontból nem különbözik ettől lényegesen a hordozható számítógépek (laptop, notebook, netbook, tablet PC stb.) alkalmazása. Az OECD PIAAC (Program for International Assessment of Adult Competencies) ezt a módszert alkalmazta felnőttek informatikai lehetőségekben gazdag környezetben történő problémamegoldó gondolkodásának mérése során. Mindegyik megoldásnál ki kell szállítani és vissza kell gyűjteni a mérőeszközöket (hordozható médiumot, számítógépet), így a logisztika ugyanolyan, vagy (értékes eszközökről lévén szó) bonyolultabb annál, mint amit a papíralapú tesztelés igényel. Kétségtelen előny a gazdagabb feladatformátum és az adatrögzítési költségek kiküszöbölése. A tetemes költségek és a szervezési feladatok miatt azonban rendszeres és tömeges felmérésekre ezt a megoldást sem lehet alkalmazni.

Az internet-alapú online tesztelés esetében a felmért személyek elvileg bárhol lehetnek, a tesztek (itemek, feladatok) egy központi szerverről érkeznek, és a megoldások ugyanoda jutnak vissza. Ebben az esetben alapkövetelmény, hogy a felmért személyek (közel) azonos számítógépeket használjanak. A tesztek megoldása történhet erre a célra készített célszoftver (kliens program) vagy egy általános célú böngésző segítségével. Ennek az eljuttatási módszernek kritikus pontja a felhasználó és a szerver közötti adatforgalom sebessége. Itt minden késlekedés elvesz a tesztet megoldó idejéből, akár már néhány másodperces késés is zavaró. Hosszabb késés már frusztráló hatású lehet, a bizonytalanság pedig a megbízható felmérést lehetetlenné teszi. Alapkövetelmény tehát, hogy az internetkapcsolat megfelelően gyors legyen. Sok felhasználó (pl. néhány ezer) tanuló egyidejű kiszolgálása gazdag multimédia tartalmú feladatokkal komoly szerverkapacitást igényelhet. Az internet használata – ha a tesztnek jelentős tétje van – a feladatok közvetítése révén különböző adatbiztonsági megoldásokat igényel, ami tovább növeli a költségeket. Ugyanakkor az egy központi szerver és az azt fenntartó fejlesztő munkacsoport, illetve technikai személyzet hosszabb távon gazdaságossá teheti az online tesztelést, különösen, ha az gyakori és sok felhasználót kiszolgál.

A felméréseket osztályozhatjuk két további szempontból is, melyekre az előzőekben már utaltunk. Az egyik a teszt jelentősége, az eredmény súlya a felmért személy számára. Ebből a szempontból megkülönböztethetjük a kicsi (pl. egy iskolai önértékelő feladatlap) vagy a nagy (pl. egy záróvizsga) fontossággal bíró (low-stake, high-stake) tesztek. Ebben a felosztásban az egyén számára nem sorsdöntő jelentőségű tesztelés megvalósítására alkalmasabb, mert kevesebb követelményt támaszt az online eljuttatási módszer.

A másik szempont az, hogy milyen gyakori, és milyen részletes felmérésre kerül sor. Az összegző-lezáró (szummatív) tesztek nagyobb tananyagmennyiséget, szélesebb körű tudást mérnek fel. Alkalmazásukra ritkábban kerül sor, és a megfelelő tesztek a felmért tudásból csak mintát vesznek. A segítő-formáló (formatív, diagnosztikus) célú értékelés gyakori felmérést igényel, részletes és az egyes konkrét tudáselemekre irányul. Az online tesztelés mindkét esetben alkalmazható, azonban a részletes, gyakori mérés papíralapon gyakorlatilag szinte megvalósíthatatlan, a másik eljuttatási módszer pedig idő- és szervezési igénye miatt nem felel meg a gyakori alkalmazás igényeinek.

Összegezve elmondhatjuk, hogy az eljuttatási módszerek tekintetében az online tesztelés jelenti a legkényelmesebb, és hosszabb távon a leggazdaságosabb technológiai megoldást. Ugyanakkor az internet és az információtechnológia jelenlegi fejlettségi szintje mellett elsősorban a kisebb tétellel járó, gyakori, formatív és diagnosztikai célú felmérések megvalósítására alkalmas. Az online eljuttatási módszer korlátokat jelenthet a feladatok információtartalma, így a multimédiás megoldások alkalmazása tekintetében is.

## ONLINE ADATGYŰJTŐ RENDSZEREK ÉS A TESZTELŐ PLATFORMOK

A teszteléshez hasonló adatgyűjtés online megvalósítására számos megoldás áll rendelkezésre, melyek egy része széles körben, akár ingyenesen is hozzáférhető. A továbbiakban röviden áttekintjük ezeket, egyrészt annak illusztrálására, milyen könnyen hozzáférhető ez a technológia, és milyen mértékben terjed egyszerű alkalmazása, másrészt annak bemutatására, mennyivel több követelményt támaszt egy rendszeres felmérésre használható tesztelő platform. Ezen eszközök az általános célú adatgyűjtésen túl a kérdőívek készítésére és egyszerű tesztek szerkesztésére is alkalmasak.

A széles körben elterjedt online oktatási platformok is tartalmaznak valamilyen értékelő, tesztelő modult. Ezek a platformok (educational platforms, pl.: Moodle, Coospace) azért jöttek létre, hogy a tanulás-tanítás különböző, elektronikus környezetben végezhető folyamatait megkönnyítsék, mint például a tanárok, diákok kollaboratív együttműködésén túl tananyagok kiosztását, házi feladatok beszedését teszik lehetővé. Egyik további funkciójuk segítségével számítógépes környezetben végezhetjük a tudás-, illetve képességmérést.

A kérdőívszerkesztő alkalmazások (survey software, pl.: OpenSurveyPilot, SurveyMonkey, GoogleSurvey) kizárólag adatgyűjtésre használhatóak, teljesítménymérés során a tanulói válaszok rögzítését, kérdőíves vizsgálat alkalmával háttéradatok felvételét könnyítik meg. A digitális kérdőív elkészítéséhez a kérdőív szerkesztője regisztrálja magát az adott online alkalmazásba. Ezután rögzíti kérdéseit, melyek lehetnek zárt- (rádiógombos, jelölőnégyzetes és legördülő menüt tartalmazó kérdések) és nyíltvégűek (rövid, illetve hosszú szöveges választ váróak) is. A kérdőív rögzítését követően elérhetőségét (mely egy URL cím, jelszóval védett kérdőívek esetén egy azonosító és jelszó páros) kiküldi a kérdőív kitöltőinek. Miután a megkérdezettek megválaszolták a kérdéseket, a válaszokat táblázatos formában letöltheti az adott weboldalról. A kérdőívek esetében nem lehet a „jó válasz” fogalmát értelmezni, ezért a kérdőívszerkesztők a tesztelő szoftverekkel ellentétben nem jelzik vissza a teljesítményeket, legfeljebb zárt kérdések alkalmazásakor a válaszok eloszlásának ismertetésével találkozhatunk.

Az online tesztszerkesztő szoftverek (test creation tools, pl.: Hot Potatoes, WebQuiz, Questiontools) többnyire olyan alkalmazások, melyek internetes böngésző segítségével érhetők el. A teszt készítőjének ebben az esetben is regisztrálnia kell a kiválasztott rendszerbe, majd a belépést követően elkészítheti feladatait. A feladatok felviteléhez segítséget nyújtanak a beépített feladatsablonok. E rendszerek képesek a teszt közvetítésére, diákokhoz való eljuttatására is, azaz a tesztek kiosztására, a tanulói válaszok rögzítésére, a feladatok

automatikus javítására (többnyire zárt itemtípusok esetén), valamint a tanulói teljesítmények azonnali visszajelentésére.

A mérés-értékelési platformok (assessment platforms, ilyen például a TAO is) az online tesztelő szoftverek és oktatási platformok elektronikus mérés-értékelési funkcióját látják el, rendszeres, professzionális alkalmazásra készülnek. A szoftvernek nagyszámú feladatot kell tárolnia, melyek hosszabb fejlesztő munka eredményeként készülnek el és jelentős szellemi értéket képviselnek. Ezért ezek a platformok az előzőekben említetteken túl számos egyéb funkcióval is kiegészülnek. Nagyszámú itemtípus, különböző tesztvégrehajtó algoritmus (pl.: adaptív, multi-stage – többlépcsős tesztalgoritmus) kezelését célozzák meg, valamint a személyre szóló tesztelés alapját, a parametrizált itembankok létrehozását segítik. Ezek a rendszerek támogatják a kollaboratív fejlesztő munkát, amikor például sok munkatárs dolgozik a feladatok fejlesztésén.

Az értékelési platformok többnyire modul rendszerben készülnek, az egyes modulok egymással kapcsolatban állnak, közöttük megoldott a szabványosított adatforgalom. A modulok együttesen a pedagógiai értékelés teljes folyamatát átfogják, a feladatírástól, tesztek szerkesztésétől a tesztelés megvalósításán keresztül a visszajelzésekig, statisztikai elemzésekig és jelentéskészítésig.

A bemutatott számítógépes alkalmazások között kereskedelmi forgalomban kapható és ingyenesen elérhető, nyílt forráskódú szoftverek is megtalálhatók. A kereskedelmi forgalomban kapható szoftverek előnye, hogy többnyire egy felhasználóbarát, jól működő alkalmazáshoz jutunk a program megvásárlásával. Ugyanakkor ezek meglehetősen sokba kerülnek, és mivel garantáltan sokféle feltétel mellett működniük kell, viszonylag szerény képességekkel rendelkeznek, tipikusan csak néhány feladattípus kezelésére alkalmasak. Ha igényeinkhez szeretnénk adaptálni a rendszert, vagy számunkra nélkülözhetetlen funkcióval kívánjuk bővíteni, akkor a szoftver fejlesztőit kell felkérni (és megfizetni) a programozói munka elvégzésére – magyarországi alkalmazásra gondolva mindenképp a magyar nyelvi adaptáció elkészítésére. Mivel ezek a rendszerek zártak, nagyobb, folyamatos fejlesztést feltételező projektek lebonyolítására nem alkalmasak. A saját választásunkat meghatározta az is, hogy a diagnosztikus értékelés specifikus igényeit kielégítő platform a projekt elindulásakor egyáltalán nem volt kereskedelmi forgalomban.

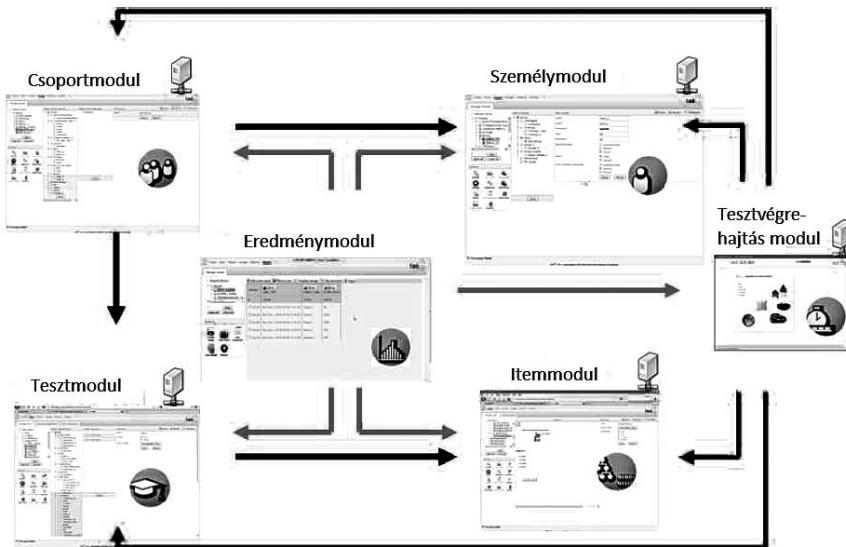
A nyílt forráskódú szoftverek rendkívüli előnye, hogy megfelelnek a bővíthetőség követelményének, hiszen a számítógépes program forráskódja elérhető, ezáltal átláthatóvá válik a teljes rendszer. Az átláthatóság lehetővé teszi, hogy a szoftver használói akár a szoftver fejlesztőivé is váljanak. A program ingyenes hozzáférése bátoríthatja a szélesebb körű felhasználást, így a folyamatos evolúciót is. Egy megfelelő fejlettségű, nyílt forráskódú mérés-értékelési platform a bővíthetőség és adaptálhatóság miatt nemzeti vagy több országon átívelő mérés-értékelési program, mint az OECD-PISA vagy PIAAC vizsgálatok alapjául szolgálhat. A nyílt forráskódú szoftverek alkalmazásának hátránya ugyanakkor, hogy nagyobb szakértelmet vár el a felhasználóktól, azoknak a program minden részletét ismerniük kell, és minden módosítás után maguknak kell a biztonságos működésért felelősséget vállalniuk.

## A TAO RENDSZER

A TAO (Testing Assisté par Ordinateur<sup>4</sup> – számítógépes tesztelés; LATOUR ÉS MARTIN 2007; LATOUR ÉS FARCOT 2008; MARTIN 2009; MARTIN, LATOUR, BURTON, BUSANA, KELLER, REICHERT, PLICHART, JADOUŁ ÉS SWIETLIK 2007a, 2007b) egy általános, nyílt forráskódú mérés-értékelési platform, amely egyaránt lehetőséget biztosít számítógéppel támogatott tesztfejlesztésre és internetalapú tesztelés levezetésére is. A szoftvert a Luxemburgi Egyetem és a Centre de Recherche Public Henri Tudor Intézet tervezte és implementálta. A TAO program segítségével tesztfeladatot szerkeszthetünk, parametrizált elektronikus itembankokat hozhatunk létre, elektronikus tesztet állíthatunk össze, különböző tesztalgoritmusok alkalmazásával online tesztelést vezethetünk le, valamint a teszt megoldását követően a tesztek azonnali kiértékelését, a tanulói teljesítmények azonnali visszajelentését valósíthatjuk meg.

A szoftver felépítését tekintve moduláris, azaz a mérés-értékelés-visszajelentés folyamatát egymással összeköttetésben álló komponensek végzik. Ez a decentralizált struktúra lehetővé teszi felhasználók időtől és helytől független együttműködését, mely támogatja a tesztszerkesztők kollaboratív tevékenységét és a munka felosztását. A rendszer fő moduljait (item, teszt, személy, csoport és tesztvégrehajtás modul) és azok kapcsolatát az 1. ábra mutatja be.

1. ÁBRA: A keretrendszer moduljainak egymáshoz való viszonya



Forrás: Ras, Plichart és Latour, 2010

Az itemmodul a tesztitemek és -feladatok létrehozásáért felelős modul. A rendszerben meglévő gyakorlatok szerkesztése és az új itemek fejlesztése közben előnézet segítségével megtekinthetők a szerkesztés alatt álló feladatok. A szöveges információkon túl multimédiás elemeket (kép, ábrák, hang, videó) és hiperhivatkozásokat is használhatunk a feladatok létrehozása során. A tesztfeladatok digitalizálását beépített sablonok segítik (MARTIN és mtsai, 2007a, 2007b). A rendszerbe felvitt feladatok megoldási idejét korlátozhatjuk (időkorlát megadásával), valamint minden itemhez megjegyzéseket fűzhetünk, melyeket megjeleníthetünk a tesztelés során.

Az elkészített itemekből összeállíthatjuk tesztjeinket a teszt modulban. Ez a modul határozza meg a teszt végső megjelenítési formáját. Itt van lehetőségünk különböző – korábban általunk meghatározott – itemparaméterek megadására (pl.: feladat súlya, itemnehézség, diszkriminációs index, találgatás valószínűsége), a teszt időkorlátjának beállítására, a tesztvégrehajtási folyamat szabályozására (pl.: engedélyezett-e a visszalépés a korábbi feladatokra, hol helyezkedjenek el a navigációs gombok, az egyes itemekhez tartozó megjegyzések megjelenítéséről is dönthetünk), és tesztvégrehajtási algoritmusok (pl.: szekvenciális, véletlenszerű itemkiválasztás) definiálására.

A tesztelés során a diákok internetes böngésző segítségével érik el a tesztfelületet, ahol azonosítójukkal és jelszavukkal bejelentkeznek a tesztrendszerbe, megoldják a feladatokat. A feladatok megoldásához bármely kellően gyors számítógép és bármely általános célú böngésző felhasználható. A tesztelés befejeztével, azonnal megjelenik a képernyőn az elért teljesítményük százalékban kiszámolva. A tanulók azonosítóit és jelszavát előzetesen a személymodulba (testees modul) regisztrálja a tesztelést levezető informatikus szakember. Ez az attribútum-lista (azonosító és jelszó páros) tetszés szerint bővíthető, így további adatokat (pl. a tanuló neve, iskola neve, címe, diák neme, osztálya) tárolhatunk a diákokról.

A „regisztrált” diákokat a csoportmodulban (groups modul) csoportokba rendezhetjük. Ez az adott tesztet megoldó konkrét minta kijelölésének alapja. Célja, hogy megkönnyítse a teszt-diák összerendelést. Az azonos bemeneti feltételekkel rendelkezőket (pl.: azonos tesztet megoldókat) egy részmintába rendezve, elegendő a tesztet az adott csoporthoz hozzárendelni ahhoz, hogy az a csoport összes tagja számára hozzáférhető legyen.

A teszt közvetítésének modulja (delivery server modul) felelős a tesztelés zavartalan lebonyolításáért. A tesztadminisztrátor meghatározza, milyen időszakban és összesen hány alkalommal hozzáférhető az adott teszt, valamint beállítja egy diákcsoporthoz rendelt tesztek sorrendjét. Ez a modul szabályozza továbbá, hogy az eredmények hol kerülnek mentésre. A válaszok kiértékelését és az elért teljesítmény visszajelentését is ez a komponens végzi.

A tesztelési folyamat során keletkezett adatokat a rendszer automatikusan generált .log fájlok formájában menti a megadott szerverre, a megfelelő személy-, csoport-, item- és tesztspecifikus adatokkal együtt, valamint az eredmény modulban (results modul) grafikus formában is megjeleníti azokat. A tesztelést követően a tesztadatok táblázatos formába rendezhetők és .csv fájlként exportálhatók.

Szerver oldalról tekintve a TAO akár egy egyszerű laptopról, vagy akár egész szerverfarmról is futtatható – a minta nagysága és a tesztelés egyéb jellemzői által igényelt

teljesítménytől függően. A rendszert többnyelvű alkalmazásra tervezték, a felhasználói felület célnyelvi adaptációja egyetlen fájl lefordításával megvalósítható. A nyílt forráskód előnyeit kihasználva, új itemtípusok, tesztalgoritmusok integrálhatók a rendszerbe, melyek implementálásához programozási ismeretekre van szükség, továbbá a forráskód ismerete lehetővé teszi más rendszerek integrációját, összekötését a TAO-val.

A platformot a projektet elindító munkacsoport folyamatosan fejleszti. 2011-ben már az alaposan kibővített 2.0 változat bejelentésére került sor. További tervek között szerepel a platform mobil eszközökre adaptálható verziójának elkészítése (részletesebben lásd ARCANI, JARS, DEVOOGHT, BOGAERTS ÉS JADOUŁ 2010). A TAO-t felhasználó és adaptáló munkacsoportok ugyancsak folyamatos fejlesztő munkát végeznek, az eredmények jelentős része beépül a fejlesztő központban készülő újabb verziókba.

A bemutatott, folyamatosan fejlődő rendszer prezentált funkciói, tulajdonságai alapot teremtenek e rendszer oktatásban történő hatékony alkalmazására.

A számítógép-alapú tesztelés teljes körű, kis- és nagymintás, alacsony vagy magas tétellel bíró, szummatív, formatív és diagnosztikus tesztelés megvalósításához.

## A TAO NEMZETKÖZI ALKALMAZÁSAI

A számítógép-alapú mérés-értékelés bevezetésében Európában elsőként Luxemburg előkelő helyet foglal el, mivel a TAO platform segítségével ott került elsőként bevezetésre egy országos, online mérés-értékelési rendszer. A jelenleg futó rendszer egyszerre képes tesztelni a teljes diákpopulációt (Luxemburgban egy teljes évfolyam mintegy 5000 tanulót jelent). A rendszer a tesztelés eredményéről automatikus visszajelzést készít a tanárok részére (PLICHART, JADOUŁ, VANDENABEELE ÉS LATOUR, 2004; PLICHART, LATOUR, BUSANA ÉS MARTIN 2008).

A TAO rendszer legkiterjedtebb alkalmazására az OECD mérési programjai keretében került sor (OECD, 2008, 2009; Sørensen és Andersen 2009). Az OECD nem online méréseket alkalmazott, adatfelvételei során csak a TAO egyes moduljainak felhasználására került sor, ugyanakkor ezen alkalmazással több új funkcióval is kiegészült a rendszer, továbbá számos nyelvre adaptálták.

Érdekes új kihívást jelentett 2009-ben a számítógépes olvasás-felmérés (Electronic Reading Assessment – ERA; HALDANE 2009; OECD 2008), amely a számítógépes munka során felmerülő olvasási feladatok megoldását mérte fel életszerű helyzetekben. A számítógép-alapú mérés-értékeléshez való egyre pozitívabb viszonyt mutatja, hogy míg 2006-ban a természettudományi tudás számítógép-alapú adatfelvételében csak három ország (Dánia, Izland és Korea) vett részt, addig 2009-ben az elektronikus olvasásvizsgálat adatfelvétele már több mint 20 ország részvételével történt (MOLNÁR 2010a, 2010b).

A nyílt forráskódú szoftverek előnyeit kiaknázva mind az ERA, mind a PIAAC vizsgálatban a TAO funkcióit kiegészítették a Softcon IT Service által fejlesztett, CBA Item Builder szoftverrel (BARKOW ÉS DOROCHESKY 2009; DOROCHESKY 2010). E rendszer alkalmas komplexebb elemeket tartalmazó, interaktív és innovatív itemformátumok létrehozására is.



A rendszer-integráció további tapasztalatait példázza a TAO és a GeoGebra<sup>5</sup> – interaktív, oktatási célra kifejlesztett matematikai szoftver, mely geometriai formák, függvények stb. ábrázolására, az alakzatok dinamikus változtatására alkalmas – összekötésével készített pilot hatásvizsgálat (KREIS, DORDING, PORRO ÉS JADOUL 2010).

Szintén a TAO rendszer átvételén és fejlesztésén dolgozik a DIPF (Deutsches Institute für Internationale Pedagogische Forschung) egyik munkacsoportja, egy jelentős TBA (Technology-Based Assessment) projekt keretein belül. A munka végső célja az, hogy e rendszer képezze a német nemzeti értékelési rendszer alapját, ami jelenleg a hazai kompetenciavizsgálatokhoz hasonlóan papíralapú vizsgálat.

A DIPF által vezetett TBA projekt másik lényeges eleme a dinamikus problémamegoldás mérését lehetővé tevő interaktív felület és mérési modell kidolgozása (WIRTH ÉS KLIEME 2003; Wirth és Funke 2005). A MicroDYN teljes mértékben kihasználja a technológia adta lehetőségeket. A tesztelt személyeknek egy dinamikusan változó környezetben kell a feladatokat megoldaniuk (BLECH ÉS FUNKE 2005; GREIFF ÉS FUNKE 2008). Ez a rendszer képezi 2012-ben az OECD-PISA problémamegoldás modul számítógép-alapú adatfelvételének alapját.

## A TAO MAGYAR ADAPTÁCIÓJA ÉS A DIAGNOSZTIKUS TESZTELÉS

Az említett előnyök kihasználására az MTA-SZTE Képességkutató Csoport néhány évvel ezelőtt elindította az online teszteléssel kapcsolatos kutatásokat, a TAO rendszer implementálását. Az első nagymintás online adatfelvételre 2008 májusában került sor. A vizsgálat úgy zajlott, hogy minden egyes diák mind papír, mind számítógépes formátumban megoldotta az alkalmazott tesztet, ami nemzetközi viszonylatban is egyedi kutatásnak számított. A kutatási elrendezés lehetőséget biztosított a diákszinten történő összehasonlításra is, ami által pontosabb képet kaphattunk az alkalmazott közvetítő eszköz befolyásoló hatásáról. Az eredmények alapján megállapítható, hogy bár a teljesítmények között szoros kapcsolat volt megfigyelhető, mégis voltak közvetítőeszköz-függő itemek és diákok, azaz olyan itemek és diákok, amelyek és akik viselkedése jelentősen különbözött papír-ceruza és számítógépes környezetben. Ennek az időszaknak az elemzése elsősorban a papíralapú és az online tesztelés összehasonlítására, a különbségek feltárására (médiá hatás) és a számítógépes tesztelés validitásának elemzésére irányultak (CSAPÓ, MOLNÁR ÉS R. TÓTH 2009; MOLNÁR 2010a, 2010b).

Az első kísérletek nyomán bebizonyosodott, hogy Magyarországon az iskolákban már rendelkezésre állnak azok az alapvető feltételek, amelyek között a fejlesztő munkát el lehet indítani. Ezt követően a Szegedi Tudományegyetem Oktatásméleti Kutatócsoportja egy TÁMOP program keretében elindította egy online diagnosztikus értékelő rendszer kidolgozásának első fázisát.

A TÁMOP pályázat „A diagnosztikus tanulói teljesítménymérések elektronikus alapra helyezését szolgáló eszközök adaptációja, kipróbálása” című részprojektje keretein belül további online felmérésekre került sor (lásd MOLNÁR, R. TÓTH ÉS TÓTH 2010). A részprojekt célja a TAO

5 [www.geogebra.org](http://www.geogebra.org)

rendszer adaptálása, továbbfejlesztése és oktatási kontextusban történő használatának kidolgozása. A 2009 tavaszán elvégzett online felméréseknek az volt a célja, hogy még pontosabb képet kapjunk a papír-ceruza és számítógép-alapú tesztelés egymáshoz való viszonyáról. Ennek fényében más-más kutatási elrendezésben három fő műveltségi területen (matematika, olvasás és problémamegoldás) történt meg a párhuzamos papír- és számítógép-alapú adatfelvétel. A mintát ez esetben 2. és 6. évfolyamos diákok alkották. (A kutatás és eredményeinek részletesebb ismertetését lásd MOLNÁR 2010c, 2010d; R. TÓTH 2009; HÓDI és R. TÓTH 2009.) E kutatás eredményei alátámasztották korábbi tapasztalatainkat, miszerint lényeges kérdés a közvetítő eszköz validitást befolyásoló szerepének vizsgálata, azaz annak a kérdésnek a megválaszolása, vajon ugyanazt mérjük-e papír, mint számítógép alapon. E vizsgálat keretein belül területfüggetlenül jól azonosíthatóak voltak azok a diákok, akiknek a viselkedése nagyon, kevésbé vagy egyáltalán nem függött az alkalmazott közvetítő eszköz típusától (R. TÓTH 2010).

Ezt követően 2010 tavaszán még nagyobb mintán került sor a felmérésekre. Egyrésztől egy tágabb életkori spektrumban (1–8. évfolyam) vizsgáltuk a diákok egy lényeges általános gondolkodási képességének, az induktív gondolkodás fejlettségi szintjét. A vizsgálat kisiskolások körében lezajlott részének egyediségét az adja, hogy a közvetítő eszköz teljesítménybefolyásoló hatásának vizsgálatára fókuszáló hazai és nemzetközi kutatások szinte kivétel nélkül idősebb diákokkal és felnőttekkel foglalkoztak, továbbá nem vizsgálták kisiskolás diákok számítógép-alapú tesztkörnyezetben való viselkedését (MOLNÁR, R. TÓTH és CSAPÓ 2010). 2010 tavaszán ismételten sor került a három fő műveltségi terület (matematika, természettudomány és olvasás) bevonásával a párhuzamos papír- és számítógép-alapú adatfelvételre. Ebben az esetben az alkalmazott tesztek különbözősége miatt horgony itemeket alkalmaztunk az adatfelvétel során, amelyek (1) lehetővé teszik az eredmények közös képességskálára konvertálását, (2) e kutatási elrendezés alkalmazásával lényegesen több feladat paraméterezésére kerülhetett sor. Szintén 2010 tavaszán a TAO platform kérdőívmodulja (TAO CAPI) alkalmazásával megtörtént egy tanároknak szóló nem lineáris kérdőív felvétele. A kutatás a tanulói teljesítménymérések tanítási folyamatra vonatkozó hatását vizsgálta általános iskolában dolgozó pedagógusok nézetei alapján (TÓTH 2009; TÓTH és CSAPÓ 2010).

A számítógép-alapú tesztelés széles körű bevezetésének egy fontos állomása az iskolák infrastruktúrájának felmérése, ehhez 2010 őszén–2011 januárjában a mérésekben részt vevő iskolák rendszergazdjával vagy az informatikai eszközökért felelős kollégájával kitöltöttünk egy IKT ellátottságra vonatkozó kérdőívet, ami előzőleg a Kutatócsoport keretein belül került kidolgozásra (TÓTH, MOLNÁR és CSAPÓ 2011). Ezzel párhuzamosan sor került egy másik adatfelvételre is, a korábbi online tesztelések felügyelő tanárait kértük meg arra, hogy töltsék ki többek között IKT szokásaikra kérdező kérdőívünket. A kérdőív eredményei által árnyaltabb képet kapunk a tanárok online teszteléssel kapcsolatos attitűdjéről, a tesztelés lebonyolításának sikerességéről és a diákok hozzáállásáról. 2011 tavaszán túlléptünk az eredetileg papír alapra készült, statikus mérőeszközök digitális alkalmazásán, és adaptáltunk egy német fejlesztésű

dinamikus problémamegoldó gondolkodást vizsgáló tesztet. Így túljutottunk a hagyományos itemtípusok alkalmazásán, sőt, eddig nem vizsgált és vizsgálható képesség- és készségterület mérését valósítottuk meg. A hazánkban elsőként alkalmazott dinamikus teszt felvétele több ezer 6–11. évfolyamos diák körében történt. Szintén 2011 tavaszán került sor egy önkormányzat összes közoktatási intézményében 2–11. évfolyamos tanulók gondolkodási képességei fejlettségének online vizsgálatára (MOLNÁR és CSAPÓ 2011), az önkormányzat iskoláinak infrastrukturális felszereltségének feltérképezésére (TÓTH és MOLNÁR 2011), illetve a pedagógusok értékeléssel kapcsolatos véleményének monitorozására (TÓTH és KINYÓ 2011).

Az MTA-SZTE Képességkutató Csoportja longitudinális vizsgálatában a 2011-ben induló újabb minta esetében domináns közvetítő eszközként a számítógépet alkalmazza. Ezzel empirikus vizsgálataink során fokozatosan eltávolodunk a papíralapú méréstől és korlátaitól, miközben áttérünk a számítógép-alapú tesztelésre, értékelésre. Az adatfelvétel eredményeinek elemzése mellett a tanároknak és igazgatóknak szóló visszajelentés közvetítő eszközét is változtatjuk. Bár a diákok 2011 tavaszán már azonnali visszacsatolást kapnak az aktuális tesztelés végén teljesítményükről, a tanárok és iskolaigazgatók számára az elmúlt két évben kidolgozásra és pilot tesztelésre került online visszajelentő rendszert alkalmazzuk.

Mindezen kutatások megalapozzák a három fő műveltségi területre és az 1–6. évfolyamra fókuszáló online diagnosztikus mérési rendszer hazai kiépítését. A diagnosztikus mérési rendszer célja, hogy rendszeres tanulói szintű visszajelzésekkel segítse a diákok egyéni fejlődését. E cél megvalósításához azonban minden tanulót évente több alkalommal is fel kell mérni, ami lehetővé teszi az egyéni fejlődési görbék pontos felrajzolását. A diagnosztikus értékelés során azonban nem elegendő egy összegző értékelést adni, hanem fel kell tárni a tanulók hiányosságait, meg kell határozni a lemaradás mértékét, a hiányosság pótlásának módját, és vissza kell jelenteni ezeket az információkat a diákok és a pedagógusok számára. Mindennek megvalósítását alapozza meg a kb. 2000 itemet tartalmazó parametrizált feladatbank, a gyors és hatékony adatfelvételt és visszacsatolást lehetővé tevő online platform (TAO) és a visszajelentő rendszer.

\*

Összefoglalásul megállapítható, hogy egy hatékony diagnosztikus értékelési rendszer kiépítésének alapját olyan számítógépes platformnak kell alkotnia, amely lehetővé teszi a felhasználók kollaborációját, parametrizált elektronikus itembank kiépítését, és egyben az online tesztelés lebonyolítását, az eredmények visszacsatolását. Továbbá követelmény a rendszer bővíthetősége, miután új itemtípusok, újabb tesztalgoritmusok kidolgozásával hatékonyabb tesztelés érhető el.

## Köszönetnyilvánítás

A tanulmány megírására a Szegedi Tudományegyetem Oktatásméleti Kutatócsoportja által végzett TAMOP 3.1.9/08/01 program keretében került sor. A TAO adaptálásának első fázisát a MTA-SZTE Képességkutató Csoport támogatta. A tanulmány írása idején Molnár Gyöngyvér Bolyai János Kutatási Ösztöndíjban részesült.

## IRODALOM

- ARCANI, D., JARS, I., DEVOOGHT, K., BOGAERTS, J., ÉS JADOUL, R. (2010): The MATAO platform or the mobile assessment using TAO. *ICL 2010 Conference*. Hasselt, Belgium, 2010. szeptember 15–17.
- BARKOW, I. ÉS DOROCHEVSKY, M. (2009): CBA Item Builder - a new Generation of Test Item Authoring Tools in Computer Based Assessment. *1st Szeged Workshop on Educational Evaluation*. Szeged, Hungary, 2009. április 27–28.
- 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/esprid/dokumente/doc-2005/>
- BLECH05\_01.PDF. (Letöltés ideje: 2010. november 5).
- 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 BENŐ, MOLNÁR GYÖNGYVÉR ÉS R. TÓTH KRISZTINA (2008): A papíralapú tesztek a számítógépes adaptív teszteséig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 3–4. SZ. 3–16.
- CSAPÓ, B., AINLEY, J., BENNETT, R., LATOUR, T. ÉS LAW, N. (2011): Technological issues of computer-based assessment of 21st century skills. In McGaw, B. és Griffin, P. (szerk.): *Assessment and teaching of 21st century skills*. Springer, New York. 143–230.
- CSAPÓ, B., MOLNÁR, GY. ÉS R. TÓTH, K. (2009): Comparing paper-and-pencil and online assessment of reasoning skills: A pilot study for introducing TAO in large-scale assessment in Hungary. In Scheuermann, F. és Björnsson, J. (szerk.): *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing*. Office for Official Publications of the European Communities, Luxemburg. 120–125.
- DOROCHEVSKY, M. (2010): Building rich problem solving items with the CBA Item Builder. *2nd Szeged Workshop on Educational Evaluation*. Szeged, Hungary, 2010. április 14–15.
- GREIFF, S. ÉS FUNKE, J. (2008): Measuring complex problem solving: The MicroDYN approach. *Kézirat*. Heidelberg. <http://www.psychologie.uni-heidelberg.de/ae/allg/>
- /FORSCHUN/DFG-KOMP/GREIFF&FUNKE.2008.MICRODYN.PDF. Letöltés ideje: 2010. november 15.
- HALDANE, S. (2009): Delivery platforms for national and international computer based surveys. 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. 63–67.
- HÓDI ÁGNES ÉS R. TÓTH KRISZTINA (2009): Olvasási képesség mérése számítógépes környezetben. *IX. Országos Neveléstudományi Konferencia*. Veszprém. 2009. november 19–21. 96.

- KOZMA, R. (2009): Assessing and teaching 21st century skills. *Call to action*. 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. 13–23.
- KREIS, Y., DORDING, C., PORRO, V., ÉS JADOUL, R. (2010): GeoGebra and TAO. *Interoperable Interactive Geometry Conference. Hluboká nad Vltavou, Czech Republic, 2010. július 2–3.*
- 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>. Letöltés ideje: 2010. november 10.
- 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.
- MARTIN, R., LATOUR, T., BURTON, R., BUSANA, G., KELLER, U., REICHERT, M., PLICHART, P., JADOUL, R. ÉS SWIETLIK, J. (2007A): TAO: modular architecture; TAO: architecture and use cases of a collaborative, internet-based platform for computer-assisted testing. *12th Biennial Conference of European Association for Research on Learning and Instruction*. Budapest, Hungary, 2007. augusztus 28–szeptember 1.
- MARTIN, R., LATOUR, T., BURTON, R., BUSANA, G., KELLER, U., REICHERT, M., PLICHART, P., JADOUL, R. ÉS SWIETLIK, J. (2007B): TAO: use cases and item templates; TAO: architecture and use cases of a collaborative, internet-based platform for computer-assisted testing. *12th Biennial Conference of European Association for Research on Learning and Instruction*. Budapest, Hungary, 2007. augusztus 28–szeptember 1.
- MCGAW, B. ÉS GRIFFIN, P. (2011): *Assessment and teaching of 21st century skills*. Springer, New York.
- MOLNÁR GYÖNGYVÉR (2010A): Technológia-alapú mérés-értékelés hazai és nemzetközi implementációi. *Iskolakultúra*, 7–8. sz. 22–34.
- MOLNÁR GYÖNGYVÉR (2010B): Technology-based assessment: Challenges and Promises. *Keynote presentation. Sixth European Distance and e-Learning Network Research Workshop*. Budapest, Hungary. 2010. október 24–27. <http://www.eden-online.org/eden.php?menuId=524>. Letöltés ideje: 2010. október 30.
- MOLNÁR GYÖNGYVÉR (2010C): Papír- és számítógép-alapú tesztelés összehasonlító vizsgálata problémamegoldó környezetben. In Perjés István és Kozma Tamás (szerk.): *Új Kutatások a Neveléstudományokban*. Aula Kiadó, Corvinus Egyetem, Budapest. 135–144.
- MOLNÁR GYÖNGYVÉR (2010D): A problémamegoldó gondolkodás vizsgálata számítógépes környezetben. *VIII. Pedagógiai Értékelési Konferencia*. Szeged, 2010. április 16–17. 55.
- MOLNÁR GYÖNGYVÉR, R. TÓTH KRISZTINA ÉS CSAPÓ BENŐ (2010): Papír- és számítógép-alapú tesztelés összehasonlító vizsgálata kisiskolás diákok körében. *X. Országos Neveléstudományi Konferencia*. Budapest, 2010. november 4–6. 167.
- MOLNÁR GYÖNGYVÉR, R. TÓTH KRISZTINA ÉS TÓTH EDIT (2010): Developing online diagnostic assessment – Experiences of a large scale national case study in public education in Hungary. *Sixth European Distance and e-Learning Network Research Workshop*. Budapest, Hungary, 2010. október 24–27. 186–187.

- MOLNÁR GYÖNGYVÉR ÉS CSAPÓ BENŐ (2011): A képességek fejlődésének vizsgálata online tesztkörnyezetben 2–11. évfolyamon. XI. Országos Neveléstudományi Konferencia, Budapest, 2011. november 4–6. 253.
- OECD (2008): Issues arising from the PISA 2009 field trial of the assessment of reading of electronic texts. Document of the 26th meeting of the PISA Governing Board. OECD Directorate for Education, Paris.
- OECD (2009): PISA CBAS analysis and results – Science performance on paper and pencil and electronic tests. OECD, Paris.
- PLICHART, P., JADOU, R., VANDENABEELE, L. ÉS LATOUR, T. (2004): TAO, a Collective distributed computer-based assessment framework built on semantic web standards. *International Conference on Advances in Intelligent Systems – Theory and Application AISTA2004*. Luxembourg, Luxembourg, 2004. november 15–18.
- PLICHART, P., LATOUR, T., BUSANA, G. ÉS MARTIN, R. (2008): Computer based school system monitoring with feedback to teachers. In Luca, J. és Weippl, E. (szerk.): *Proceedings of World Conference on Educational Multimedia, Hypermedia and Telecommunications 2008*. VA: AACE, Chesapeake. 5065–5070.
- R. TÓTH KRISZTINA (2009): A számítógépes és papír-ceruza tesztek eredményeinek összehasonlító vizsgálata háttérváltozók alapján. IX. Országos Neveléstudományi Konferencia. Veszprém, 2009. november 19–21. 97.
- R. TÓTH KRISZTINA (2010): Médiahatás vizsgálata teljesítményen alapuló mintafelosztás segítségével. X. Országos Neveléstudományi Konferencia. Budapest, 2010. november 4–6. 168.
- RAS, E., PLICHART, P. ÉS LATOUR, T. (2010): Addressing the needs of today's e-learning environments via an open and versatile architecture for assessment. *ICL 2010 Conference*. Hasselt, Belgium, 2010. szeptember 15–17. [https://www.tao.lu/html/images/stories/download/Publications/pub\\_tao\\_architecture\\_caf2010\\_final.pdf](https://www.tao.lu/html/images/stories/download/Publications/pub_tao_architecture_caf2010_final.pdf). Letöltés ideje: 2010. december 5.
- 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.
- TÓTH EDIT (2009): A tanárok értékeléssel kapcsolatos attitűdjei, nézetei. IX. Országos Neveléstudományi Konferencia. Veszprém, 2009. november 19–21.
- TÓTH EDIT ÉS CSAPÓ BENŐ (2010): A tanulói teljesítménymérések hatása a tanítási folyamatra általános iskolában dolgozó pedagógusok nézetei alapján. X. Országos Neveléstudományi Konferencia. Budapest, 2010. november 4–6. 169.
- TÓTH EDIT ÉS KINYÓ LÁSZLÓ (2011): Pedagógusok nézetei a mérés-értékelés néhány aspektusáról és ezek kapcsolata a tanulók teljesítményével. XI. Országos Neveléstudományi Konferencia, Budapest, 2011. november 4–6. 254.
- TÓTH EDIT ÉS MOLNÁR GYÖNGYVÉR (2011): Az iskola infrastrukturális fejlettségi szintjének teljesítménybefolyásoló hatása. „XI. Országos Neveléstudományi Konferencia”, Budapest, 2011. november 4–6. 255.
- TÓTH EDIT, MOLNÁR GYÖNGYVÉR ÉS CSAPÓ BENŐ (2011): Az iskolák IKT felszereltsége – helyzetkép országos reprezentatív minta alapján. *Iskolakultúra*, 10–11. sz. 124–137.

WIRTH, J. ÉS KLIEME, E. (2003): Computer-based assessment of problem solving competence. *Assessment in Education: Principles, Policy & Practice*, 10. 3. sz. 329–345.

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.