

A mérőeszköz-bővítéstől a tesztelési folyamat vizsgálatáig: számítógépes tesztelés nagymintás nemzetközi vizsgálatokban

Napjainkban világszerte ugyanazon kérdésekre más-más hangsúlyokkal keresik a válaszokat kutatók, oktatásügyi szakemberek és pedagógusok egyaránt: a folyamatosan változó világ és környezet hogyan hat az oktatásra, illetve milyen készségeket és képességeket kell az egyénnek elsajátítania a mindennapi boldoguláshoz, sikerhez (Trilling és Fadel, 2009). Ennek oka, hogy a 21. században megjelenő technológiai innovációkhoz való hatékony alkalmazkodás érdekében számos új készségcsoport (például digitális média műveltség) elsajátítása vált szükségessé. Ugyanakkor sok olyan készség, képesség, illetve kompetenciaterület azonosítható (például kritikai gondolkodás, problémamegoldás), amelyek az elmúlt században éppoly relevánsak voltak, mint napjainkban, azonban elsajátításuk módja, valamint hétköznapi alkalmazásuk dinamikusan változik (Partnership for 21st CenturySkills, 2009). Ezeket a javarészt a tudományos-technológiai forradalom hatásainak szükségszerű „melléktermékeként” létrejövő, illetve megújuló készségeket, kompetenciákat a szakirodalom „21. századi készségek” vagy „21. századi kompetenciák” gyűjtőfogalom alatt tömöríti (Partnership for 21st CenturySkills, 2009).

A 21. századi készségek vizsgálatát számos nemzetközi projekt (például: OECD DeSeCo: *21. századi készségek mérése és tanítása*–[Assessment and Teaching of 21st CenturySkills, Partnershipfor 21st CenturySkills]) tűzte ki célul. Egyetértés mutatkozik abban, hogy a jelenkor tanulónak a 21. században való sikeres boldoguláshoz tantárgyi tudás, specifikus készségek, szakértelem és műveltség kiegyensúlyozott egyvelegét kell elsajátítaniuk és használniuk, amelyek alapvetően négy kategóriába sorolhatóak (Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci és Rumble, 2012). Az első (Binkley és mtsai, 2012 alapján) a gondolkodás, ahová a kreativitás, a kritikai gondolkodás, problémamegoldás, döntéshozatal és tanulás tartozik. A következő két kategória a munkában való sikeres létezés szükséges készségeket (kommunikáció, kollaboráció) és eszközöket, illetve azok használatát (IKT-eszközök és infokommunikációs műveltség) tömöríti. Míg az előbbieken felsoroltak az egyén boldogulásához szükségesek, a negyedik csoportba tartozó készségek (például: állampolgárság, élet és karrier, személyes és

társadalmi felelősségvállalás) a társadalomban való eligazodást segítik. E területek hatékony mérésére szolgálnak az IKT-eszközök, azon belül is a számítógépek különböző formái, hiszen e mérési területek gyakran interaktív, illetve dinamikusan változó környezetben mérhetőek a leghatékonyabban és a leginkább objektíven.

A számítógépek nyújtotta előnyök és a tanulók értékelése közötti szinergia megteremtése érdekében a kognitív műveltségi területek, valamint az affektív készségek elektronikus környezetben történő méréséhez a korábbi tartalmi és értékelési keretek átgondolása szükséges. Ennek oka, hogy az új tesztkörnyezet új lehetőségeket kínál, amelyeket az új tartalmi kereteknek is le kell képezniük, valamint új változókat von(hat) be a pedagógiai vizsgálatokba, amelyek hatással lehetnek a mérőeszköz validitására is (*Huff és Sireci, 2001*).

Jelen tanulmányunk az új tesztkörnyezetben rejlő, a pedagógiai mérés-értékelés során hasznosítható lehetőségek áttekintését tűzte ki célul: egyrészt bemutattuk, hogy az IKT-eszközök hogyan integrálhatók a pedagógiai mérés-értékelés gyakorlatába, másrészt megvizsgáljuk, hogy az elektronikus környezetben végzett tesztelés milyen jellemzőkkel bír a tradicionális, nyomtatott médiumon történő teszteléssel szemben. Ehhez nemzetközi szinten végzett vizsgálatok módszertani jellemzőit és példafeladatait használjuk fel.

A számítógépes mérés-értékelésben rejlő lehetőségek

A pedagógiai tesztek adatfelvétele során az IKT-eszközök két módon integrálhatók. Az egyik mód: a már meglévő papír-ceruza tesztek digitalizáljuk, azaz változtatás nélkül vagy minimális változtatással számítógépes környezetbe ültetjük át (*Tóth és Hódi, 2011*), majd e számítógépes teszt segítségével végezzük a diákok felmérését (például: *Bennett, Braswell, Oranje, Sandene, Kaplan és Yan, 2008; Csapó, Molnár és R. Tóth, 2009; Kim és Huynh, 2007*). Ez esetben a digitalizálás célja, hogy a korábban papíralapon mért konstruktum ne változzon meg, csak az adatfelvétel működjön hatékonyabban és eredményesebben. A másik lehetőség, hogy új típusú, kizárólag elektronikus környezetben alkalmazható itemeket készítünk. Például a papíralapú tesztekkel szemben az online olvasási képesség mérésekor – a teljesség igénye nélkül – a következő itemtípusokkal gazdagíthatjuk a számítógépes tesztünket: fogd és vidd ('drag and drop') technikával, itembuborék használatával és hiperlinkekkel (*Scalise és Gifford, 2006*). A felsorolt elemeket tartalmazó és egyéb új típusú itemek alkalmazása lehetővé teszi, hogy az elektronikus mérőeszköz a már meglévő papír-ceruza itemekkel lefedett területet kibővítse, vagy egy papíralapon nem mérhető, teljesen új konstruktum (lásd például: *OECD, 2011; Greiff, Wüstenberg és Funke, 2012*) mérése valósítsa meg (*van Lent, 2009*).

Összegezve tehát az IKT használata lehetőséget teremt a pedagógiai mérés-értékelés területén (*Csapó, Ainley, Bennett, Latour és Law, 2012* alapján):

1. a tradicionális adatfelvétel hatékonyabb és eredményesebb lebonyolítására;
2. a már meglévő mérőeszközök bővítésére és gazdagítására;
3. új konstruktumok (papíralapon részben vagy egyáltalán nem mérhető készségek és képességek) mérésére; illetve
4. arra, hogy információt szerezzünk a pedagógiai mérés folyamatáról, megvizsgáljuk a diákok tesztkörnyezettel folytatott interakcióját.

A papír-ceruza adatfelvétel IKT-eszközökre történő átültetése során leggyakrabban idézett előnyök: a technológia-alapú mérés-értékelés idő-, illetve hosszú távú költség-hatékonyasága (*Thurlow, Lazarus, Albus és Hodgson, 2010*) a nyomtatott médiumon történő adatfelvétellel szemben, valamint az azonnali pontozás és visszajelentés lehetősége

(Noyes és Garland, 2008). Továbbá a számítógépes környezetnek köszönhetően a diákok szívesebben és nagyobb élvezettel oldják meg a számítógépes tesztet, mint a papír-ceruza változatot (OECD, 2010a). Mindazonáltal a számítógépes adatfelvételt önbevallás alapján nem kevésbé veszik komolyan a diákok, mint a nyomtatott változatot (OECD, 2010a). A papír-ceruza és számítógépes tesztfelvétel, illetve teszteredmények összehasonlításáról részletesen lásd például van Lent (2009), magyarul pedig Csapó, Molnár és R. Tóth (2008), valamint R. Tóth (2009) tanulmányait. (E tanulmányban a digitalizált és papíralapú tesztek további összehasonlításától a terjedelmi korlátok miatt eltekintünk.)

A digitalizált tesztfelvételen túlmutatva a mérőeszközök eszköztárának bővítésével, gazdagításával elérhetjük, hogy a papíralapon mért készség, képességterület vagy tudás vizsgálatát kiterjesszük új típusú, papír-alapon nem alkalmazható itemtípusokkal. Ezáltal ugyanazt a konstruktumot más típusú itemekkel is mérhetjük, valamint kiegészíthetjük a vizsgálatot nyomtatott formátumban nem mérhető elemekkel. Ilyen típusú vizsgálatokra mutat példát a 2. fejezet.

Ugyanakkor léteznek olyan 21. századi készségek, képességek, illetve kompetenciák, melyek papíralapon nem mérhetőek. Ilyen a dinamikus problémamegoldás, a digitális szövegértés vagy egyes számítógéphasználathoz kötődő területek (e-mail rendezés, webes keresés, stb.) vizsgálata. Ezért e pedagógiai/pszichológiai felmérések kizárólag technológiailag gazdag környezetben történhetnek. A harmadik fejezet az új konstrukciók mérésére irányuló vizsgálatokat összegzi.

Függetlenül attól, hogy papír-ceruza tesztjeinket digitalizáljuk, a korábban papíralapon mért domaint bővítjük ki, vagy új mérési területhez készítünk itemeket és azokkal mérjük fel a diákjaink képességeit, teljesítményét, a technológia-alapú mérés lehetőséget nyújt arra, hogy a tesztelés folyamatát megvizsgáljuk. A számítógépes tesztelés során a diákok egyéni, a tesztelő szoftverrel folytatott interakcióit rögzítjük, majd az eltárolt adatok alapján kiértékelhetjük azokat. E témáról a 4. fejezet nyújt részletesebb áttekintést.

Mérőeszközök bővítése és gazdagítása

A már korábban, papír-ceruza alapon tesztelt kompetenciákat, készségeket és képességeket a számítógépes tesztelés során új, a papír-ceruza mérőeszközöknél gazdagabb vizuális és auditív ingereket tartalmazó feladatok segítségével vizsgálhatjuk meg. Például a 2006-os OECD-PISA tanulói teljesítménymérő programsorozat természettudományi eszköztudás vizsgálatában (OECD, 2010a) a nyomtatott tesztek mellett opcionális lehetőségként megjelent a technológia-alapú mérés-értékelés. E vizsgálati ciklusban 15 éves tanulók természettudományos eszköztudását mérték papír-ceruza alapon és számítógépes környezetben. A számítógépes adatfelvétel során a diákok fix formátumú, multimédiás elemekkel gazdagított számítógépes itemeket (például lásd 1. ábra) oldottak meg (OECD, 2010a). A mérés három ország részvételével (Dánia, Izland, Korea, részletesebben lásd. Halldórsson, McKelvi, és Björnsson, 2009; Lee, 2009; Sørensen és Andersen, 2009) zajlott. A vizsgálatra 45 itemet fejlesztettek, és minden item két részből állt: egy stimulus tartalmú részből és egy feladat részből (OECD, 2010a). A stimulus terület szöveget, képet, videót vagy flash animációt tartalmazott, míg a feladat megoldását feleletválasztó ('multiple choice' – MCQ – vagy komplex MCQ) itemekkel végezték (OECD, 2010a). Az 1. ábrán látható item stimulus része egy videót, a válaszalternatívák pedig képeket tartalmaznak.


PISA CBAS - Item Viewer - Form_01-SOURCE.cta

File View Locate (en_IE) Help

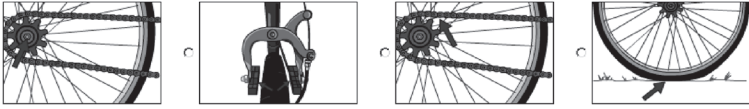
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43

Question 5: Bicycle

Danny is riding along a track and has to stop quickly. He applies the brake strongly and **the bike skids to a halt**. When Danny applies the brake, a calliper pushes rubber pads against the rim of the wheel.



Which part of the bicycle is likely to have the highest temperature just as the bicycle comes to a stop?



Previous Next

Start [Icons] Jango - Free Music - I... Engleke Filer PISA CBAS - Item V... Dokument 1 - Microsof... 10:27

1. ábra. Az OECD-PISA 2006-os természettudományi eszköztudás mérésének egyik példája (OECD, 2005, idézi: Sørensen és Andersen, 13. o.)

Egy másik példa, a Test of English as a Foreign Language Internet-based Test (TOEFL IBT) nyelvvizsga teszt, mely szintén a mérőeszköz gazdagítására ad példát. A tesztet 2005-ben vezették be, és fokozatosan tértek át a internet-alapú adatfelvételtől az internet-alapú vizsgáztatásra (*Educational Testing Service*, a továbbiakban rövidítve: *ETS*, 2008). A TOEFL IBT teszt a korábbi papír-alapú angol nyelvvizsga elektronikus megvalósítása, ugyanakkor nem kizárólag digitalizálása, mert a nyelvvizsga papír-alapú és internetes változata különbözik egymástól. Van olyan szövegértési feladat, melyben, kihasználva a technológia nyújtotta előnyöket, a vizsgázók online környezetben rákattinthatnak egy jelölt szóra vagy kifejezésre, és megnézhetik annak definícióját, illetve magyarázatát (*ETS*, 2008.). Például a 2. ábrán a szövegben megtalálható „shamans” szó jelentését a bal oldalon olvashatjuk. Hallásértési feladatok esetén pedig számítógépes környezetben nemcsak a szöveget hallják a vizsgázók, hanem képek is segítik az elhangzottak kontextualizálását és a beszélők szerepeinek, szándékainak beazonosítását, amely segítségével könnyebben elképzelhetik az adott szituációt (*ETS*, 2008).

The screenshot shows a TOEFL Reading test interface. At the top, it says 'TOEFL Reading' and 'Question 3 of 12'. The time is 00:18:54. The question asks: 'According to paragraph 5, why do some scholars refer to a trance state to help understand the cave paintings?'. There are four multiple-choice options:

- To explain the state of consciousness the artists were in when they painted their pictures
- To demonstrate the mythical significance of the strange geometric shapes
- To indicate that trance states were often associated with activities that took place inside caves
- To give a possible reason for the strange appearance of the men painted on the cave walls

Paragraph 5 is marked with an arrow [→].

A glossary window is open, defining 'shamans' as 'holy people who act as healers and diviners'.

The passage text on the right explains that a third opinion takes psychological motivation further into the realm of tribal ceremonies and mystery, mentioning 'shamans' who believe that an animal's spirit and energy is transferred to them while in a trance. It also notes that perhaps no satisfactory answers to the cave images exist, but their existence reveals that ancient human beings were not without intelligence, skill, and sensitivity.

2. ábra. TOEFL IBT példaitem (ETS, 2008)

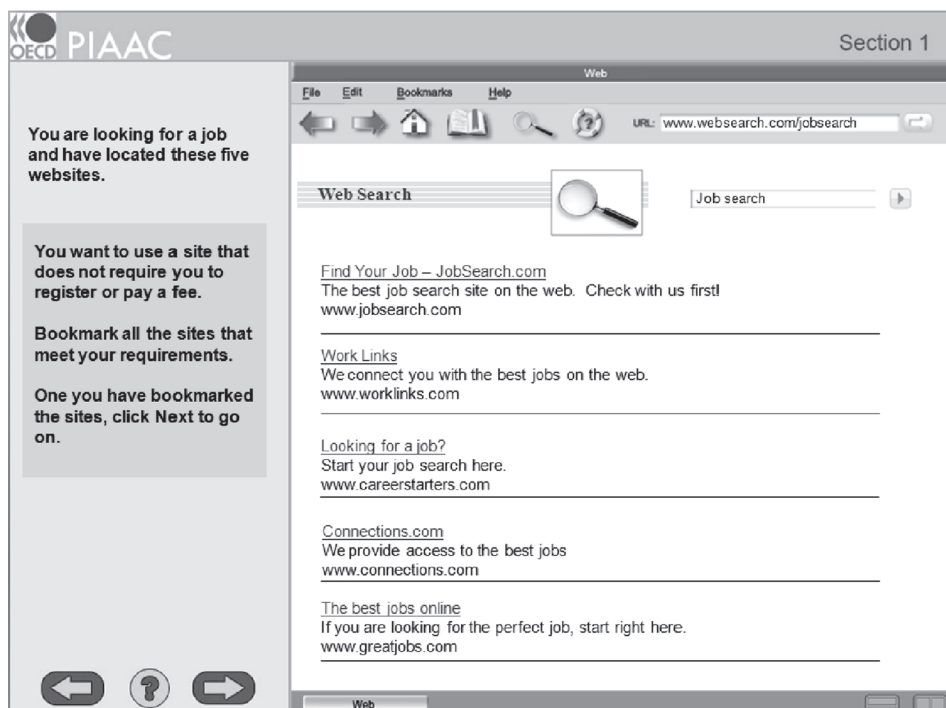
Papír-alapon nem mérhető konstruktumok: lehetőségek a 21. századi képességek mérésére

Léteznek olyan konstruktumok, melyek a 21. században fontos szerepet töltenek be, ugyanakkor természetükből fakadóan kizárólag elektronikus környezetben vizsgálhatók. Példaként említhetjük a dinamikus problémamegoldást (Greiff és mtsai, 2012) és a digitális szövegértést (például: OECD PISA – Electronic Reading Assessment, lásd: OECD, 2011). Az olyan, számítógéphez, mobil eszközök használatához kapcsolódó mérések, mint a számítógép-, illetve internethasználati ismeretek feltérképezése: fájlkeresés, e-mail küldés, e-mail rendszerezés, kulcsszavas keresés (Kammerer és Gerjets, 2012; Lennon, Kirsch, Davier, Wagner és Yamamoto, (2003); Pfaff és Goldhammer, megjelenés alatt; OECD, 2012) szintén kizárólag interaktív környezetben valósíthatóak meg leghatékonyabban.

Noha a 21. századi képességek, a számítógép-, illetve internethasználatához fűződő ismeretek egy része statikus, papír-ceruza alapon is vizsgálható (például Dancsó, 2009), az egyes feladatok megvalósításának folyamatát, azt, hogy a diák nem csak ismeri, de képes az egyes műveletek elvégzésére, csak számítógépes környezetben, szimulációk segítségével vizsgálhatjuk. Ilyen szimulációs vizsgálat zajlott az OECD Programme for the International Assessment of Adult Competencies (felnőtt kompetenciák nemzetközi mérési programja, rövidítve: PIAAC) 2011-es „Problémamegoldás technológiailag gazdag környezetbe” című mérése során, melynek középpontjában az információs-kommunikációs technológiák használatához kapcsolódó problémamegoldás állt (OECD, 2012). A PIAAC mérésben használt problémamegoldás definíciója alapján (lásd: OECD, 2012) a technológia-alapú környezetben végzett problémamegoldás magába foglalja a digitális

technológia, kommunikációs eszközök és hálózatok használatát, az információhoz való hozzáférést és annak értékelését, a másokkal történő kommunikációt és hétköznapi problémák megoldását (például foglalás, vásárlás, munkakeresés).

A PIAAC próbamérésben alkalmazott feladatra ad példát a 3. ábra, mely feladat egy webes keresőfelületet szimulál, ahol a „Job search” (munkakeresés) kifejezésre kapott találatok láthatók a webböngészőben (‘browser’). A feladat megoldásához a tesztet megoldó személyek hiperlinkek segítségével megtekinthetik a keresésben kapott oldalakat, majd könyvjelzővel (‘bookmark’) el kell látniuk azokat a weboldalakat, amelyek teljesítik a feladatban foglaltakat, azaz amelyek nem igényelnek regisztrációt és fizetést.



3. ábra. PIAAC példaitem (OECD, 2012, 53. o.)

Szintén webes keresés eredményének feldolgozási folyamatát vizsgálta Pfaff és Goldhammer (megjelenés alatt) a németországi National Educational Panel Study* (NEPS) keretében. A szerzők által közölt vizsgálat egy webes keresőfelületet valósított meg, ahol a diákok egy kulcsszavas keresés eredményeit láthatták. A feladat felépítése megegyezett a 3. ábrán látható példaitemmel: a diákok a képernyő bal oldalán az instrukciót, jobb oldalán az interaktív stimulus felületet láthatták. E jobb oldali webes keresési eredmények linkjeit követve a diákoknak el kellett dönteniük, melyik link tartalmazza a leginkább hiteles információt egy feltett kérdésre. A feladatok megoldásához tehát a tanulóknak tudniuk kell alkalmazni a hiperlinkeket, navigációs (előre, hátra, home) gombokat, és a weboldalak közül azt kellett megjelölniük, mely adatait tartják a leginkább

* <http://tba.dipf.de/en/projects/neps-national-educational-panel-study>

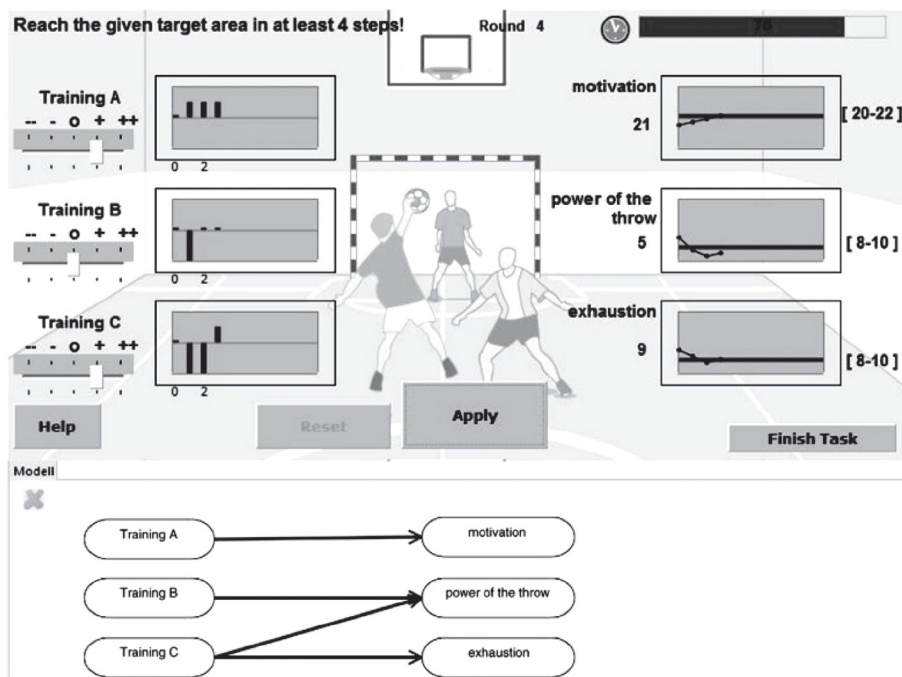
megbízhatónak, és a feltett kérdés tekintetében leginkább relevánsnak (*Pfaff és Goldhammer*, megjelenés alatt).

A digitális olvasás-szövegértés vizsgálatára ad példát az OECD-PISA 2009-ben 19 ország részvételével végzett elektronikus szövegolvasás és – értés-vizsgálata (*OECD*, 2011). A tesztitemek itt is két részből álltak: egy stimulus részből és egy kérdésből. A stimulus részben található digitális szövegek hipertextusok, azaz írott szövegek és képi információk halmazai, amelyek navigációs eszközökkel vannak összekötve (*OECD*, 2011. 40. ol alapján). Így a feladatok egy több oldalból álló webes környezetet jelenítenek meg, melyben a diákoknak hiperlinkek segítségével kellett navigálniuk.

A digitális olvasás fejlettségének vizsgálatakor tehát az olvasás-szövegértés képessége mellett az információhoz való hozzáférést is mérjük, hiszen a digitális szövegek hatékony megértéséhez a tanulónak képesnek kell lennie explicit és beágyazott hiperlinkek alkalmazására, és jártasnak kell lennie a navigációs eszközök használatában (*Mendelovits, Lumley és Searle*, 2009 alapján). Ehhez hasonló vizsgálatot terveztek (*Webs olvasás-szövegértés – Web-based reading* címmel) 9-10 éves tanulók körében a Progress in International Reading Literacy Study (PIRLS) – Nemzetközi Szövegértés-vizsgálat keretében, melynek 2011-re tervezett kipróbálását 2016-ra módosították (*Mullis, Martin, Kennedy, Tron, és Sainsburg*, 2009).

A digitális olvasás-szövegértés feladatok mellett szintén kitüntetett szerepet kapott a problémamegoldó gondolkodás vizsgálata elektronikus környezetben, mivel e terület az OECD PISA 2012-es mérésében is szerepelt (*OECD*, 2010b). E nemzetközi vizsgálatban a statikus problémaszituációk mellett olyan interaktív feladatok is helyet kaptak, melyek csak a technológia-alapú mérés-értékelés keretei között oldhatók meg (*OECD*, 2010b).

Az ilyen interaktív feladatokat tartalmazó problémamegoldó gondolkodás mérésére Greiff és munkatársai (2012) bevezettek egy ún. MicroDYN alkalmazást, mely szorosan kapcsolódik Dörner operacionális intelligencia elméletéhez, és a problémamegoldási folyamatot három facet segítségével adja meg: információ kinyerés ('information retrieval'), modellképzés ('modellbuilding') és előrejelzés ('forecasting'). A rendszer működésének prezentálására bemutatunk (*Wüstenberg, Greiff és Funke*, 2012 alapján) egy példaitemet (4. ábra), melyben a diákoknak három megadott edzéstípus (Training A, Training B és Training C) hatását kell vizsgálniuk. Majd meg kell állapítaniuk, hogyan hat a három edzéstípus a csapat jellemzőire ('motivation' - motiváció, 'power of throw' - dobáserősség, és 'exhaustion' - kimerültség). A csúszkák pozitív (+,++) vagy negatív (-, --) irányba történő elmozdítása növeli/csökkenti az adott edzésforma mennyiségét. A problémamegoldás első fázisában a diákok kipróbálják a rendszer működését, azaz feltérképezik, hogyan függnek össze az endogénváltozók (edzéstípusok) az exogénváltozókkal (a csapatok jellemzőivel). Ezt követi a modellképzés, ahol a diákoknak grafikusán kell ábrázolniuk, hogyan függnek össze a változók, azaz milyen inputra, milyen outputot ad a rendszer. Végül (3. facet: forecasting) előre meghatározott értékeket – csapatjellemzőket – kell a diákoknak elérniük, azaz alkalmazniuk kell az előzőekben megszerzett tudást (*R. Tóth, Molnár, Wüstenberg, Greiff és Csapó*, 2011).



4. ábra. Problémamegoldás példaitem (Wüstenberg és mtsai, 2012. 5. o.)

Még komplexebb számítógépes rendszer meglétét feltételezi majd a kollaboratív problémamegoldás (Collaborative Problem Solving- CPS) vizsgálata, mely az egyénnek az a képessége, amely elősegíti a hatékony részvételt és együttműködést egy olyan problémamegoldási folyamat során, amelynek kettő vagy annál több résztvevője (ágense) van (OECD, 2013 alapján). Ilyen CPS vizsgálat szerepel majd a 2015-ös OECD PISA mérésekben, amely során a tanulók különböző képességekkel rendelkező szimulált résztvevőkkel kollaborálnak, hogy a kollaboratív készségekről (például csoportgondolkodás, a hatékony csoport- és egyéni szintű interakció megteremtéséhez szükséges kommunikációs készségek) valid információkat szerezzenek (OECD, 2013).

A tesztelési folyamat vizsgálata

A számítógépes tesztelés lehetőséget nyújt arra, hogy ne csak a végső tanulói választatot, vagy elkészített produktumot (például fogalmazást) értékeljük, hanem információkat gyűjtsünk a tesztmegoldás folyamatáról is. Ezen információk értékelésbe történő bevonása lehetőséget teremt pontosabb tudományos következtetések levonására, illetve azok finomítására. Ehhez azonban pontosan rögzítenünk kell a diákok tesztmegoldó tevékenységét.

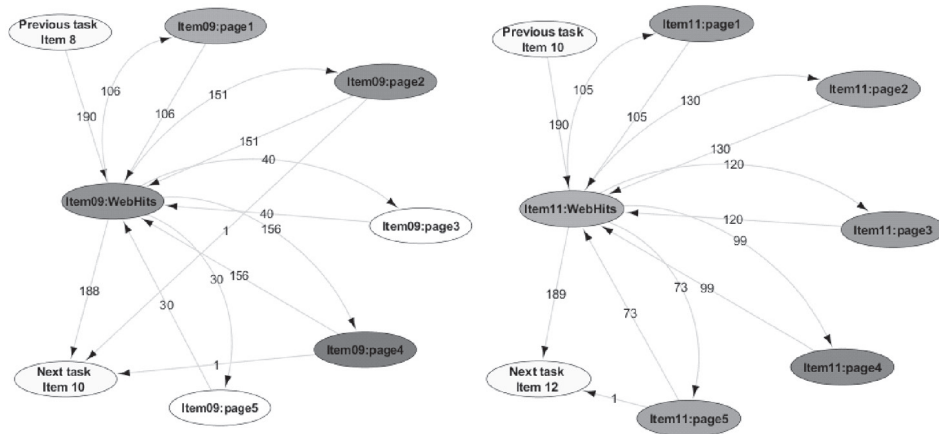
Az elektronikus tesztelés során a diákok tesztelőrendszerrel folytatott interakcióját az adatfelvevő szoftver rögzíti (Tóth, Rölke és Goldhammer, 2012a). Ezen információk alapján tájékozódhatunk arról, hogy egy diák mennyi időt töltött egy feladat megoldásával, hányszor javított a válaszában, melyik kérdést hagyta ki, melyik kérdésre tért vissza, nem lineáris tesztek esetén milyen sorrendben oldotta meg a tesztfeladatokat, hogyan

tudta kezelni az egeret, a billentyűzetet, stb. Azonban a diákok tesztmegoldási folyamatának adatai gyakran nem kerülnek bele az elemzésekbe. Ennek egyik lehetséges magyarázata, hogy az adott vizsgálatban csak a végeredmény számít, azaz, hogy helyes vagy nem helyes a tanuló válasza. Egy másik lehetséges ok, hogy a nagy mennyiségű tanulói tevékenységeket rögzítő fájlok (ún. logfájlok) feldolgozásához programozói ismeretekre van szükség (Tóth, Rölke, Naumann és Goldhammer, 2012b), ez pedig nehezíti a tesztmegoldási folyamatból származó információk felhasználását.

A logfájlok feldolgozásának egyik legelterjedtebb módszere, hogy a loggolt adatokból előre definiált változók értékeit kinyerjük (Tóth és mtsai, 2012a, 2012b), és ezekkel a változókkal jellemezzük az egyes tanulók tevékenységeit. Ilyen változók lehetnek például: a tesztmegoldással töltött idő, illetve webes keresőfelületeket megvalósító feladatok esetén, hogy hány darab weboldalt nézett meg a diák, vagy hány darab, a feladatmegoldás szempontjából releváns oldalt nézett meg a tanuló a teszt megoldása során (OECD, 2011). Ilyen és ezekhez hasonló változók segítségével képet kaphatunk a tesztitem és a diák interakciójáról, a tesztitem működéséről, mellyel segíthetjük a tesztkészítők munkáját.

Egy tesztitemmel folytatott tanulói interakciók vizuális megjelenítésére ad példát az 5. ábra (Tóth és mtsai, 2012a alapján). A vizualizáció alapjául szolgált egy komplex tesztitem, mely a 3. ábrához hasonló, öt weblinket tartalmazó webes keresőfelületet valósított meg. A feladat megoldásához végzett tevékenységeket a tesztelőszoftver rögzítette és logfájlokban tárolta. A logfájlokból kinyerték a tanulók összes tevékenységét, majd a kinyert akciókat aggregálták, és az alábbi irányított gráf segítségével ábrázolták. A gráf csúcsai (oválisok) egy-egy weboldalt reprezentálnak. Például Item09:Webhits a tesztitem nyitóoldala, mely a webes keresés eredményét jeleníti meg. Az Item09:page1 pedig az első hiperlink segítségével elérhető weboldalt reprezentálja. A gráf élei azt mutatják meg, hogy egy weboldalról hány tanuló navigált el egy másik weboldalra. (Például a kilencedik item nyitóoldaláról 106 diák navigált az első elérhető weboldalra.) Az 5. ábráról leolvasható, hogy a kilencedik itemen a diákok a harmadik és az ötödik weboldalt nagyon ritkán nézték meg, míg a 11. item esetén kiegyensúlyozottabb eloszlást mutat a weboldalak felkeresésének száma (Tóth és mtsai, 2012a alapján).

A gráf az interakciók gyakorisági eloszlása mellett az egyes weboldalakon eltöltött időt is reprezentálja. Ezt az információt a gráf csúcsainak színe adja meg. Minél sötétebb egy csúcs, annál több időt töltöttek ott a diákok átlagosan. E két gráf azonban eltérést mutat, hiszen a bal oldali gráfon a harmadik és ötödik oldalon a diákok nagyon rövid időt töltöttek, míg a jobb oldali gráf kiegyensúlyozottabb időeloszlást igazol. Ezen eredmények alapján megállapítást nyert, hogy a 9. item esetén a harmadik és ötödik webes találat nem tölti be a disztraktor funkcióját, mert a diákok ezeket az oldalakat ritkán nézik meg, és akik megnézik, nagyon rövid idő alatt eldöntik róla, hogy nem releváns a feladatmegoldás szempontjából (Tóth és mtsai, 2012a alapján).



5. ábra. Aggregált interakciók (Tóth és mtsai, 2012. 2068. o.)

Az itemműködés vizsgálatán túl egyéb pedagógiai területen is kihasználhatjuk a logfájlokban kódolt információkat. Számos vizsgálat vette például górcső alá a diákok navigációs profilját (például: *Barab, Bowdish és Lawless, 1997; Lawless és Kulikowich, 1998; Niederhauser, Reynolds, Salmen és Skolmoski, 2000*), melyek legalább három navigációs profilt azonosítottak (*Lawless, Brown, Mills és Mayall, 2003* alapján): (1) információ-/tudáskereső ('knowledge seekers'), (2) főbb jellemzők felfedezői ('feature explorers') és (3) apatikus/közönyös felhasználók ('apathetic hypertext users'). A tudáskereső (akiket könyvimádókként is szoktak aposztrofálni) kategóriája azon olvasókat jelöli, akik a hipertextus tartalmához kötődő információkra vadásznak. Ők olyan képernyők felé navigálnak, amelyek tartalma szükséges a jobb megértéshez. A felfedező (őket forrás-élvezőknek is hívják) aránytalanul sok időt töltenek multimédiás elemekkel. Látszólag ezek az egyének több időt töltenek azzal, hogy megértsék, hogyan működik a hipertextus, mint azzal, hogy releváns információkat gyűjtsenek az írott szövegből. A közönyös felhasználókat nem érdekli a hipertextusban rejlő lehetőségek feltárása; sem információ kinyerésére, sem pedig a jellemzők feltérképezésére nem szánnak időt (*Lawless, Brown, Mills és Mayall, 2003*).

A felhasználói profil (angolul 'user profile') elkészítésén kívül a navigációs jellemzők a tanulói teljesítmények magyarázatára is felhasználhatók. Az OECD PISA digitális olvasás-szövegértés vizsgálatában például három navigációs tényező (a megtekintett weboldalak száma, a releváns – a feladat megoldásához szükséges információt tartalmazó – oldalak megtekintésének száma, valamint a tanuló által megtekintett releváns oldalak száma) és a tanulói teljesítmények összefüggését vizsgálták (*OECD, 2011*). További példaként szolgál Tóth és munkatársai (2012b) vizsgálata, melyben a navigációs tevékenységek klaszteranalízise során kapcsolatot tártak fel a navigációs viselkedés és az itemmegoldás sikeressége között.

A navigációs viselkedés ('navigation behaviour') egyik tényezőjének, a hiperlinkek kiválasztásának kutatásával is több tanulmány foglalkozik (lásd: *Salmeron és García, 2011; Kammerer és Gerjets, 2012*). Salmeron és García (2011) a hiperlinkek kiválasztási sorrendjének felmérésére egy wiki-dokumentumot készített, és megvizsgálták, milyen tényező alapján választották ki a diákok a linkeket és milyen sorrendben. Kammerer és Gerjets (2012) német orvostanhallgatók webes keresőfelületekkel (például: Google)

kapcsolatos linkkiválasztási stratégiáját vizsgálta. A szerzőpáros arra kereste a választ, mennyire bíznak a hallgatók a Google rangsorolásában.

További, nem webes keresőfelületet megvalósító, interaktív itemek alkalmazása során szintén elvégezhetjük a tesztmegoldási folyamat és a teszteredmények kapcsolatának vizsgálatát, azonban a feladat típusának megfelelő változók definiálására van szükség. Például problémamegoldás terén, amelyre a 4. ábra mutat példát, Tóth, Wüstenberg, Rölke és Greiff (2012c) az alábbi változókat definiálta: a diák által használt beállítások száma, ismételt beállítások száma, tesztidő stb. E tesztmegoldási folyamatot leképező változók segítségével a tanulói teljesítmények több, mint 80 százaléka volt magyarázható (Tóth és mtsai, 2012c).

Összegzés

Vitathatatlan tény, hogy a napjainkban zajló technológiai fejlődés életünk szerves részét képezi, formálja azt. Ez a hatás megmutatkozik abban, hogy a mindennapokban való boldogulásunkhoz szükséges készségek repertoárja folyamatosan bővül vagy a hagyományos készségek átalakulnak a technológiai fejlődés hatására. Ez a változás az oktatás mikrokörnyezetét sem hagyja érintetlenül. Az IKT-eszközök széleskörű használatának köszönhetően a tanítási-tanulási folyamat, ezen belül a mérés-értékelés is jelentős átalakuláson megy keresztül.

Jelen tanulmány elsődleges céljaként betekintést nyújtottunk a számítógépes mérés-értékelésben rejlő lehetőségekbe, valamint megvizsgáltuk, hogy különböző nemzetközi értékelési programok hogyan integrálják a számítógépet a pedagógiai mérés-értékelés folyamatába. Nemzetközi példák segítségével bemutattuk, hogy a már meglévő mérőeszközök bővítése és gazdagítása hogyan és milyen mértékben teszi lehetővé a korábban papíralapon mért konstruktumok újszerű vizsgálatát. Az OECD-vizsgálatok módszertani jellemzőire támaszkodva prezentáltuk, milyen új konstruktumok mérését tűzték ki célul a nemzetközi nagymintás keresztmetszeti vizsgálatok, valamint azt, hogy milyen módszerekkel és feladattípusokkal mérték azokat. Górcső alá vettük, hogyan szerezhettünk információt a számítógép-alapú pedagógiai mérés folyamatáról, valamint a diákok tesztkörnyezettel folytatott interakci-

A tanulmányban foglaltak alapján megállapítható, hogy az IKT-eszközök és azok közül is például a számítógép egy olyan sokoldalú eszköz a pedagógiai mérés-értékelés folyamatában, amely bármely tesztelési folyamatba hatékonyan integrálható. Haszna több szinten is megmutatkozik: növeli a mérhető készségek, konstruktumok skáláját s jó alapot szolgáltat arra, hogy interaktív környezetben vizsgálhassuk meg számos 21. századi kompetencia fejlettségi szintjét. További nem elhanyagolható előnyei, hogy egyrészt mélyebb betekintést enged a tesztelési folyamatba és teljesebb képet ad a tanuló tesztkitöltési folyamatáról, másrészt hozzájárul az IKT jártasság és a digitális műveltség fejlődéséhez, implicit fejlesztéséhez.

ójáról, hiszen a tesztelési folyamat vizsgálata számos lehetőséget kínál a tanulói teljesítmények vizsgálata terén.

Összegezve elmondható, hogy a számítógép, mint tesztmédiium a mérőeszköz digitalizálástól kezdve, a mérőeszköz bővítésen/gazdagításon át a tesztelési folyamat vizsgálataig többféle módon alkalmazható a pedagógiai mérés-értékelési folyamatban. Azonban a hatékony alkalmazás feltétele, hogy a számítógépes környezet nyújtotta lehetőségeket a mért konstruktum jellemzőihez, valamint a mérési célhoz igazítsuk.

Irodalomjegyzék

- Barab, S. A., Bowdish, B. E., és Lawless, K. A. (1997): Hypermedia navigation: Profiles of hypermedia users. *Educational Technology Research and Development*, **45**. 3. sz. 23–42.
- 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. 2013. 10r 2.-i megtekintés, Journal of Technology, Learning, and Assessment [on-line], <http://www.jtla.org>.
- Binkley, M., Erstad, O., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M. és Rumble, M. (2012): Defining twenty-first century skills. In: Griffin, P., McGaw, B. és Care, E. (szerk.): *Assessment and Teaching of 21st century skills*. Springer, New York. 17-66.
- Csapó, B., Ainley, J., Bennett, R., Latour, T. és Law, N. (2012): 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, Bő, Molnár, Gy. és R. Tóth, Ka (2009): Comparing paper-and-pencil and online assessment of reasoning skills. A pilot study for introducing electronic testing in large-scale assessment in Hungary. In: Scheuermann, F. és Björnsson, J. (szerk.): *The transition to computer-based assessment. New approaches to skills assessment and implications for large-scale testing*. Office for Official Publications of the European Communities, Luxemburg. 113-118.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papíralapú teszteléstől a számítógépes adaptív tesztelésig. A pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, **18**. 3-4. sz. 3-16.
- Dancsó Tünde (2009): *A tanulók informatikai készségeinek fejlettsége az általános és a középiskola végén*. SZTE BTK Neveléstudományi Tanszék, Szeged.
- ETS (2008): *TOEFL IBT Tips. How to prepare for TOEFL IBT*. [Brosúra]. Educational Testing Service, USA. 2013. 10. 15-i megtekintés, http://www.ets.org/Media/Tests/TOEFL/pdf/TOEFL_Tips.pdf
- Greiff, S., Wüstenberg, S. és Funke, J. (2012): Dynamic Problem Solving: A new assessment perspective. *Applied Psychological Measurement*, **36**, 3. sz. 189-213.
- Halldórsson, A., McKelvie, P. és Björnsson, J. (2009): Are Icelandic boys really better on computerized tests than conventional ones: Interaction between gender test modality and test performance. In: Scheuermann, F. és Björnsson, J. (szerk.): *The transition to computer-based assessment. New approaches to skills assessment and implications for large-scale testing*. Office for Official Publications of the European Communities, Luxemburg. 178–193.
- Huff, K. L. és Sireci, S. G. (2001): Validity Issues in Computer-Based Testing. *Educational Measurement: Issues and Practice*, **20**. 3. sz. 16–25.
- Noyes, J. M. és Garland, K. J. (2008): Computer- vs. paper-based tasks: Are they equivalent? *Ergonomics*, **51**. 9. sz. 1352-1375.
- Kammerer, Y. és Gerjets, P. (2012): *In Google we trust? Die Bewertung der Vertrauenswürdigkeit von Suchergebnissen bei der Websuche zu einem medizinischen Thema*. Előadás: 48. Kongress der Deutschen Gesellschaft für Psychologie (DGPs), Bielefeld, 2012. szeptember 22-27.
- Kim, D. H. és Huynh, H. (2007): Comparability of Computer and Paper-and-Pencil Versions of Algebra and Biology Assessments. *Journal of Technology, Learning, and Assessment*, **6**. 4. sz. 2013. 10. 2.-i megtekintés, Journal of Technology, Learning, and Assessment [on-line], <http://www.jtla.org>.
- Lawless, K. A. Brown, S. W., Mills, R. és Mayall, H. J. (2003): Knowledge, interest, recall and navigation: A look at hypertext processing. *Journal of Literacy Research*, **35**. 3. sz. 911-934.
- Lawless, K. A. és Kulikowich, J. M. (1998): Domain knowledge, interest, and hypertext navigation: A study of individual differences. *Journal of Educational Multimedia and Hypermedia*, **7**. 1. sz. 51–70.
- Lee, M-K. (2009): CBAS in Korea: Experiences, Results and Challenges. In: Scheuermann, F. és Björnsson, J. (szerk.): *The transition to computer-*

- based assessment. *New approaches to skills assessment and implications for large-scale testing*. Office for Official Publications of the European Communities, Luxemburg, 187–193.
- Lennon, M., Kirsch, I., Von Davier, M., Wagner, M. és Yamamoto, K. (2003): Feasibility study for the PISA ICT literacy assessment. Report to network A. ACER, ETS, NIER. 2013. 10. 15-i megtekintés, <http://www.pisa.oecd.org/dataoecd/35/13/33699866.pdf>
- Mendelovits, J., Lumley T. és Searle, D. (2009): *Assessing Reading Literacy in the Digital Age*, Előadás: PISA Research Conference 2009, University of Kiel, Germany, 2009. szeptember 14-16.
- Mullis, I. V. S., Martin, M. O., Kennedy, A. M., Trong, K. L. és Sainsbury, M. (2009): *PIRLS 2011 Assessment Framework*. TIMSS & PIRLS International Study Center Lynch School of Education, Boston College, Amsterdam. 2013. 10. 15-i megtekintés, http://timss.bc.edu/pirls2011/downloads/PIRLS2011_Framework.pdf
- Niederhauser, D. S., Reynolds, R. E., Salmen, D. J., és Skolmoski, P. (2000): The influence of cognitive load on learning from hypertext. *Journal of Educational Computing Research*, 23. 3. sz. 237–255.
- OECD (2010a): *PISA Computer-Based Assessment of Student Skills in Science*. OECD Publishing, Paris.
- OECD (2010b): *PISA 2012 Field Trial Problem Solving Framework. Draft Subject to Possible Revision After the Field Trial*. OECD Publishing, Paris. 2013. 10. 15-i megtekintés, <http://www.oecd.org/pisa/pisaproducts/46962005.pdf>
- OECD (2011): *PISA 2009 Results: Students on Line: Digital Technologies and Performance (Volume VI)*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264128859-en>
- OECD (2012): *Literacy, Numeracy and Problem Solving in Technology-Rich Environments: Framework for the OECD Survey of Adult Skills*. OECD Publishing, Paris. <http://dx.doi.org/10.1787/9789264128859-en>
- OECD (2013): *PISA 2015 Draft Collaborative Problem Solving Framework*. OECD Publishing, Paris. 2013. 10. 20-i megtekintés, <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Collaborative%20Problem%20Solving%20Framework%20.pdf>
- Partnership for 21st Century Skills (2009): *P21 Framework Definitions*. Partnership for 21st Century Skills, 2013. 10. 17-i megtekintés, http://www.p21.org/storage/documents/P21_Framework_Definitions.pdf
- Pfaff, Y. és Goldhammer, F. (megjelenés alatt): *Evaluating Online Information: The influence of cognitive components*
- R. Tóth Krisztina (2009): Papír-ceruza és számítógépes tesztek eredményeinek összehasonlító vizsgálata. In: Vajda Zoltán (szerk.): *Bölcsész-műhely 2009*. JATEPress, Szeged.
- R. Tóth K., Molnár Gy., Wüstenberg, S., Greiff, S. és Csapó, B. (2011): *Measuring adults' dynamic problem solving competency*. Előadás: 14th European Conference for the Research on Learning and Instruction. Exeter, United Kingdom, 2011. augusztus 30- –szeptember 3.
- Sameron L. és Garcia, V. (2011): Reading skills and children's navigation strategies in hypertext. *Computers in Human Behaviour*, 27. 3. sz. 1143–1151.
- 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.
- 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, Luxemburg. 194–207.
- Thurlow, M., Lazarus, S. S., Albus, D, és Hodgson, J. (2010): Computer-based testing: Practices and considerations (Synthesis Report 78). University of Minnesota, National Center on Educational Outcomes, Minneapolis, MN. 2013. 10. 17-i megtekintés, <http://www.cehd.umn.edu/NCEO/onlinepubs/Synthesis78/Synthesis78.pdf>
- Tóth Krisztina és Hódi Ágnes (2011): Számítógépes és papír-ceruza teszteredmények összehasonlító vizsgálata az olvasás-szövegértés területén. *Magyar Pedagógia*. 111. 4. sz. 313-332.
- Tóth, K., Rölke, H, és Goldhammer, F. (2012a). Investigating Test-taking Behaviour in Simulation-based Assessment – Visual Data exploration. In: *EDULEARN12 Proceedings CD*.
- Tóth, K., Rölke, H., Naumann, J. és Goldhammer, F. (2012b): 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. (2012c): *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.
- Trilling, B. és Fadel, C. (2009): *21st Century Skills: Learning for Life in Our Times*. Jossey-Bass, San Francisco.

van Lent, G. (2009): Risks and Benefits of CBT versus PBT in High-Stakes Testing Introducing key concerns and decision making aspects for educational authorities. 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. 83-92.

Wüstenberg, S., Greiff, S. és Funke, J. (2012): Complex problem solving— More than reasoning? *Intelligence*, **40**. 1–14.

Vétkek és választások

Szerkesztette: Réz Anna

Gondolat Kiadó, 2013
ISBN 978 963 693 342 5
3400 Ft

Dicsérünk és hibáztatunk, bosszankodunk és felháborodunk, szemrehányást teszünk, megbocsátunk, büntetünk és jutalmazunk. Állami intézményeinkben és a saját otthonunkban, szavakkal, tettekkel vagy érzelmekkel, de erkölcsi szempontból egyfolytában reagálunk a másik tetteire. Amikor így teszünk, morálisan felelősnek tartjuk egymást.

Ez a könyv azzal foglalkozik, hogy milyen körülmények fennállása esetén helyénvaló ezt tennünk. A *Vétkek és választások* az elmúlt ötven év morális felelősségről szóló szakirodalmának legfontosabb szövegeit gyűjti össze, melyek többek között az alábbi kérdésekre keresik a választ: Megítélhetők vagyunk-e olyan dolgokért, amelyeket nem befolyásolhatunk? Felelősek vagyunk-e az érzelmeinkért és vágyainkért?

Milyen szerepet játszanak az érzelmek a felelősnek tartás és felelősségvállalás gyakorlatában? Túl filozófiai jelentőségükön a cikkek az angolszász tradíció méltó képviselői: elméleti tisztaságukkal, szellemes példáikkal és elegáns érveikkel kitűnően alkalmasak arra, hogy a felsőoktatásban (vagy akár a középiskolában) tanulók számára világossá tegyék a kérdéses problémák jelentőségét, és kedvet csináljanak a filozófia és etika tanulmányozásához.

