

**TECHNOLÓGIAALAPÚ TESZTELÉS AZ OKTATÁSBAN:
A PROBLÉMAMEGOLDÓ KÉPESSÉG FEJLŐDÉSÉNEK
ÉRTÉKELÉSE**

MOLNÁR GYÖNGYVÉR

AKADÉMIAI DOKTORI ÉRTEKEZÉS TÉZISEI

2016

BEVEZETÉS

Az elmúlt néhány évtized legjelentősebb oktatási vonatkozású fejlesztései az oktatási rendszerek különböző szintjein megvalósuló visszacsatoló mechanizmusok kiépítésére irányultak, ezért óriási fejlődésen ment át a pedagógiai értékelés elmélete és gyakorlata. A gyors fejlődés következtében az ezredforduló idején leginkább elfogadott és elterjedt papíralapú tesztekre alapuló mérések egyre több korlátba ütköztek, sőt a papíralapú tesztekre alapozott fejlesztés lehetőségei mára teljesen kimerültek. A továbblépéshez, a 21. században jelentkező új mérés-értékelési igények kielégítéséhez alapvető, minőségi változtatásra volt szükség (Scheuermann és Pereira, 2008). A változtatás irányát jelentős mértékben meghatározta a technológia, a számítógépek fejlődése és oktatásban való terjedése, mindennapossá válása. A számítógép-alapú értékelés ma már mind a hagyományos szemtől szemben történő, mind a papíralapú tesztelésnél hatékonyabb mérések megvalósítását teszi lehetővé, ezért belátható időn belül nagy valószínűség szerint megvalósul minden fontosabb mérés és vizsga technológiai alapokra helyezése.

A pszichológia kognitív forradalma, a 21. század gyorsan változó, állandó tanulás igénylő tudástársadalmában új tudáskonceptiók (pl. *OECD*, 2013a, 2013b, 2013c) kialakulása és a mérés-értékelés fejlődése – beleértve a hagyományos eszközökkel is mért területek életszerűbb mérési és fejlesztési lehetőségét, új konstrukciók mérhetőségét, motiváltabb, személy képességszintjéhez illeszkedő tesztek kiközvetíthetőségét – lehetővé tette annak meghatározását, hogy a fejlődés egyes szakaszaiban mit és hogyan érdemes mérni, valamint fejleszteni. Rávilágított és számszerűsítette az oktatás problémáit, valamint követhetővé tette az egyes oktatási beavatkozások hatását.

A hagyományos tesztelési technikákról a technológiaalapú tesztelésre történő átállás azonban nemcsak lehetőségeket, de számos kérdést, problémát és kihívást is állít a kutatók elé. A disszertáció középpontjában a számítógép-alapú tesztelésre való átállás szükségességének, előnyeinek és hátrányainak, valamint lehetőségeinek áttekintése áll, miközben a 21. században kulcsfontosságúnak tartott problémamegoldó képesség fejlettségi szintjét mérő kutatások szintetizálása segítségével konkrét példán keresztül ismertetem az adatfelvétel és adatelemzés fejlődési tendenciáit és a számítógép-alapú tesztelés tág életkori intervallumban történő megvalósíthatóságát.

A disszertáció öt fejezetből áll. Az első fejezetben áttekintem az információs és kommunikációs technológiák tanulásra és tanításra gyakorolt hatását, kitérek a változtatás oktatási szükségességére, valamint felvázolom az információs és kommunikációs technológiák (IKT) oktatási intergrációjának útjait. A második fejezet a pedagógiai mérés-értékelés technikáinak fejlődési tendenciáit ismerteti, míg a harmadik fejezet fő kutatási kérdése, hogy a számítógép-alapú tesztelés alkalmazható-e kisiskolás korban. Az egymásra épülő negyedik és ötödik fejezet a problémamegoldó képesség mérési lehetőségeinek változásán keresztül veszi górcső alá a számítógép-alapú tesztelésben rejlő mai lehetőségeket, miközben tág életkori intervallumban (9-19 éves korban) ismerteti a diákok dinamikus problémamegoldó képességének fejlődését, valamint áttekinti a változás fő tendenciáit és lehetséges okait. A technológiaalapú tesztelés lehetőségeinek kihasználása és az új elemzési eljárások ötvözésével olyan kutatási kérdésekre válaszolhattunk, amelyekre hagyományos technikákkal történő adatfelvételek és klasszikus tesztelméleti elemzések segítségével nem kerülhetett volna sor.

1. AZ INFORMÁCIÓS-KOMMUNIKÁCIÓS TECHNOLÓGIÁK HATÁSA A TANULÁSRA ÉS OKTATÁSRA

A 20. század ipari és 21. század tudás társadalma, gazdasága, kommunikációs szokásai, munkaformái, az értékesnek, versenyképesnek számító tudás jelentős mértékben eltér egymástól (*Trilling és Fadel, 2009*). Az internet kiépülése és tartalommal való feltöltése alapvetően megváltoztatta a tudáshoz való viszonyunkat, a tudáshoz való hozzáférés lehetőségeit, költségeit (*OECD, 2010a, 2012*), sőt fokozatosan háttérbe szorította a többi információforrás szerepét. Az alkalmazható tudás, az új ismeretek létrehozását lehetővé tevő tudás, valamint az információs és kommunikációs technológiák (IKT) az ezredforduló meghatározó kifejezésévé váltak.

Az oktatási rendszerek minősége, az oktatás során alkalmazott eszközök, taneszközök, módszerek ismét a figyelem középpontjába kerültek. Számos korábban értékesnek tartott ismeret, készség és képesség helyét új készség- és képességterületek vették át (*Mayrath, Clarke-Midura és Robinson, 2012*). A 21. század gazdaságában, társadalmában, munkaerőpiacán a tények memorizálásának, egyszerű eljárások implementálásának már kevesebb szerep jut – ezek feladatát átveszik a különböző technológiai eszközök –, a hangsúly a flexibilitáson, a jó komplex problémamegoldó képességen, a hatékony kommunikációs képességen és információkezelésen, a csoportmunkára való alkalmasságon, a kreatív és produktív technológiahasználaton, illetve az ezekkel összefüggő új tudás előállításának képességén van (*Quellmalz, Timms, Buckley, Davenport, Loveland és Silberglitt, 2012; Binkley, Erstad, Herman, Raizen, Ripley, Miller-Ricci és Rumble, 2012*). Mindezen, a 21. században kulcsfontosságúnak számító készségeket és képességeket azonban ritkán tanítják a ma iskolájában (*Graesser, 2012*), holott nehéz elképzelni a jelen és jövő tanulási környezetét IKT eszközök és a vonatkozó készségek és képességek hatékony használata nélkül (*RTTT, 2009; Obama, 2016*).

A technológia adta új lehetőségeknek megvan az a képessége, hogy alapjaiban megváltoztassa, transzformálja az oktatást. Vannak országok, ahol a társadalom IDI indexe (*ICT Development Index – IKT fejlődési index*) jelentős (Korea, Dánia, Izland, Anglia, Svédország, Finnország, Norvégia, Hong Kong; l. *ITU, 2015*), vagy jelentős IDI indexnövekedésen esett át (pl.: Thaiföld, Kazahsztán; l. *ITU, 2015*), ahol az IKT az oktatás-tanulás szerves részévé is vált (l. pl.: *KCC, 2014*). Ezen országok, nemzetközi oktatási eredményeik (*OECD, 2013a, 2013b*) alapján, mind az OECD PISA mérések élvonalában helyezkednek el (pl.: Finnország, Korea), vagy a jelentős fejlődést elért országok között vannak (l. Malajzia, Kazahsztán).

A technológia iskolai integrációja számos új lehetőséget kínál, ugyanakkor társadalmi csoportok és egyes személyek között akár tovább növelheti a fennálló eltéréseket. Míg egy jól átgondolt fejlesztés csökkentheti az iskolák között lévő különbségeket, addig, ha ugyanez a folyamat spontán zajlik, a jobb érdekérvényesítő iskolák még nagyobb előnyre tesznek szert. Ezzel növelve az iskolák közötti különbség nagyságát és ezáltal tovább szélesítve a tanárok, diákok, iskolák között fennálló „digitális szakadékot”. A technológia használata ugyanis több lehetőséget kínál, mint bármely korábbi taneszköz. Nemcsak a többcsatornás ismeretközlést és ezáltal a tudás új reprezentációs formáinak kialakítását teszi lehetővé, hanem motiváló erővel is bír a diákok irányában.

A technológia és információs tudástársadalom adta lehetőségek kihasználásához több feltétel együttes teljesülése is szükséges, azok egymagukban nem oldják meg a problémákat. Az IKT iskolai alkalmazása, oktatásba történő bevonása nemcsak az újabb eszközök tantermi megjelenésével és azzal párhuzamosan a tanárok IKT-s módszertani repertoárjának növelésével lehetséges, hanem az iskolai mérés-értékelés folyamatába történő integrálással is. Mindez csak akkor járul hozzá hatékonyan az oktatás fejlesztéséhez, ha nem a megjelenő technikákhoz keressük a felhasználás lehetőségeit, azaz nem a technológia a cél, hanem az oktatásban jelentkező problémák hatékony megoldására alkalmazzuk azokat (*Csapó, Molnár és R. Tóth, 2009*).

2. A PAPIRALAPÚ TESZTEKTŐL A SZÁMÍTÓGÉPES ADAPTÍV TESZTELÉSIG: A PEDAGÓGIAI MÉRÉS- ÉRTÉKELÉS TECHNIKÁJÁNAK FEJLŐDÉSI TENDENCIÁI

Az elmúlt másfél évtized egyik legdinamikusabban fejlődő területe a pedagógiai mérés-értékelés. Az ezredforduló idején leginkább elfogadott és elterjedt papíralapú tesztekre alapuló mérések a fejlesztések ellenére egyre több korlátba ütköztek, a papíralapú tesztekre alapozott fejlesztés lehetőségei mára teljesen kimerültek. Az áttérés iskolai kontextusban – mint tapasztaljuk a PISA mérések ez irányú változtatásai kapcsán – azonban csak fokozatosan lehetséges, minden lépésben gondosan ellenőrizve, és kiszűrve a nemkívánatos mellékhatásokat (*Csapó, Molnár és R. Tóth, 2008*). Az értekezés e fejezetében (1) a kötött formátumú, hagyományos, papíralapú tesztek főbb jellemzőinek fényében ismertetjük a kötetlen formátumú és a valószínűségi tesztelmélet adta lehetőségeket, (2) áttekintjük a technológiaalapú tesztelés különböző szintjeit, előnyeit, hátrányait és kihívásait, valamint (3) a pedagógiai mérés-értékeléssel foglalkozó főbb (reprezentatív mintán alapuló) kutatások hazai megjelenésének és irányvonalainak fényében bemutatjuk a technológiaalapú mérés-értékelés hazai és nemzetközi tendenciáit.

2.1. A kötött formátumú papíralapú tesztek, valamint a kötetlen formátumú és a valószínűségi tesztelmélet adta lehetőségek

Az úgynevezett hagyományos, közismert, papíralapú tesztek nagyon fontos szerepet játszottak és játszanak ma is a tanítási-tanulási folyamatok irányításában, az oktatás eredményességének felmérésében. Ezek a tesztek többnyire rögzített formátumúak, ami azt jelenti, hogy a tesztek feladatait mindig azonos formai elrendezésben kapják meg a tesztelt személyek. Az oktatási kontextusban alkalmazott mérések azonban többnyire nem egyetlen kötött formátumú tesztet igényelnek, mert például olyan nagy tudásterületet vizsgálnak, vagy olyan széles képességfejlődési spektrumot kellene átfogniuk, amelyek technikai okokból sem férnek bele egyetlen tesztbe. A probléma megoldására számos technika született. Ezek közé tartozik a teljes lefedés elve, amikor egy nagyobb tudásterület teljes felméréséhez a lehetséges összes feladat elkészül. Egy másik megoldás a feladatbankok alkalmazása, amikor lényegében a teljes lefedés előzőekben bemutatott elveit alkalmazva, tesztváltozatokba sorolva kerül sor a

feladatok bemérésére. Egy további probléma – különösen a képességtesztek esetében –, hogy a tanulók között nagyobbak a különbségek, mint amekkorát egy kötött formátumú teszttel le lehet képezni.

A klasszikus tesztelmélet által kínált eljárásokat alkalmazva ki lehet számítani a teszt sokféle jellemzőjét, azonban a paraméterek többsége szigorúan véve csak a teszt bemérésére alkalmazott minta (tanulócsoport) esetében lesz érvényes. A diákok képességszintjének meghatározása jelentős mértékben függ a kutatás során alkalmazott teszt(ek) nehézségi szintjétől.

Az a probléma, hogy a mintában minden egyes diákról megmondható legyen, hogy ő az adott, közös feladatbankban lévő feladatot milyen valószínűség mellett oldaná meg, még akkor is, ha a konkrét feladat megoldására nem kerül sor, túlmutat a klasszikus tesztelmélet határain, a valószínűségi tesztelmélet oldotta meg. Ezzel megnyílt az út a változatos összetételű, kötetlen formátumú tesztek alkalmazása előtt. A valószínűségi tesztelmélet a mérés során elkövetett hibát és az itemek tulajdonságait más módon, nem determinisztikusan, hanem valószínűségi alapon kezeli.

A Rasch modell a valószínűségi tesztelméleti modellek között csak egy modell, mégis speciális tulajdonságainak köszönhetően kiemelt szerepet kap a pedagógiai kutatások során (l. *Bond és Fox*, 2001, 2015; *Molnár*, 2005, 2006, 2008). A speciális objektivitás biztosítja, hogy a mintában bármely két személy összehasonlítása független attól, hogy az adott konstruktumot mérő itemek közül melyiken tesszük azt, illetve bármely e tulajdonsággal bíró két item összehasonlítása független attól, hogy milyen képességszintű személy oldotta meg azokat. Ez a féle objektivitás, függetlenség az IRT modellek közül csak a Rasch modell tulajdonsága, ami biztosítja a teszt és mintafüggetlen elemzéseket.

2.2. A technológiaalapú tesztelés típusai

A technológia mérés-értékelésben betöltött szerepe sokrétű lehet, ebből adódóan nincs egységes definíció arra vonatkozólag, hogy mit jelent a technológiaalapú tesztelés, megvalósításának számos lehetősége ismert. Magában foglalja az összes olyan mérési-értékelési rendszer alkalmazását, ahol az adatgyűjtésre valamilyen információs-kommunikációs technológiai eszközt használunk.

A technológiaalapú mérésen belül – elterjedtségéből fakadóan is – a legtöbb lehetőséget a számítógép-alapú értékelés rejti magában, ahol az adatfelvétel közvetítő eszköze a számítógép, a mérőeszköz (teszt, kérdőív) kérdései, feladatai, itemei a számítógép monitorján jelennek meg. A hálózatalapú mérés egy gyakori alkalmazása, amikor az adott hálózaton belül egyszerre több gépen zajlik a tesztelés, a tesztelést egy külön számítógépről irányítják, ahol az adatok összegyűjtése, elemzése történik (*Csapó, Molnár és R. Tóth*, 2008). Az internet alapú adatfelvétel irányítását egy központi szerver végzi, azaz a helyi számítógépekre nem kell felinstallálni a tesztelést végző programot. Mivel a szoftver külső gépen fut a gépekkel szemben felállított követelmények alacsonyabbak (*Csapó, Ainley, Bennett, Latour és Law*, 2012), ugyanakkor megfelelő számú és minőségű internetkapcsolattal (és internetes böngészővel) rendelkező számítógépre van szükség. Az internet alapú tesztelés legkifinomultabb, legtöbb információt szolgáltató formája az adaptív tesztelés. Az adaptív tesztelési technika személyre szabott, a diákok képességszintjéhez igazodó, kötetlen formátumú tesztek kiközvetítését valósítja meg.

Adaptív tesztelés során a fix teszteknel tapasztalt mérési pontosság eléréséhez kevesebb feladatra, illetve rövidebb időre van szükség. Ugyanolyan mennyiségű feladat és idő alkalmazása mellett viszont biztosított a nagyobb mérési pontosság (Molnár, 2013b; Frey, 2007). Mindennek megvalósításában a Rasch modell a feladatok, résztesztek paraméterezésében játszik szerepet, ugyanis az adaptív tesztelés esetén egy korábban paraméterezett, indexelt feladatbank áll a tesztelés háttérében.

2.3. A technológiaalapú tesztelésre való átállás előnyei, hátrányai és kihívásai

A technológia fejlődése és terjedése, általánossá válása, hozzáférhetősége rendkívüli lehetőségeket kínál és kínál a pedagógiai mérés-értékelés gyakorlatának fejlesztésére. Alkalmazásának segítségével pontosabb, változatosabb, lényegesen komplexebb képességeket vizsgáló tesztelési eljárásokat, életszerűbb, alkalmazásorientáltabb, természetesebb környezeteket, feladatokat alakíthatunk ki a korábbi papíralapú mérésekhez képest (Bridgeman, 2010; Martin, 2010; Christakoudis, Androulakis és Zagouras, 2011; Csapó, Ainley, Bennett, Latour és Law, 2012; Beller, 2013; Breiter, Groß és Stauke, 2013). A disszertáció e részében részletesen kifejítjük, hogy a technológiaalapú mérésre való átállás milyen előnyökkel, illetve hátrányokkal bírhat. Hatékonysága, a hatékonyság növekedése bizonyos feltételek mellett a mérés-értékelés minden egyes szintjén kimutatható:

- a) a tesztelés gazdaságossága (pl.: Farcot és Latour, 2008; Peak, 2005);
- b) a tesztszerkesztés változatosága (Csapó és mtsai, 2012), a kiközvetítés és adatáramlás gyorsasága (Csapó, Lőrincz és Molnár, 2012),
- c) az azonnali, objektív, standardizált visszacsatolás biztosításának lehetősége (Becker, 2004; hosszabb, írott szöveges válaszok értékelése kapcsán pl.: Dikli, 2006; Valenti, Neri és Cucchiarelli, 2003),
- d) változik a diákok tesztelés iránt mutatott motivációja (Meijer, 2010; Sim és Horton, 2005),
- e) innovatív feladatszerkesztési lehetőségek, multimédiás, dinamikus, interaktív itemek, második és harmadik generációs tesztek alkalmazása (Strain-Seymour, Way és Dolan, 2009; Pachler, Daly, Mor és Mellar, 2010), amelyek korábban papíralapon nem voltak kivitelezhetőek (Molnár, 2010; Molnár, Greiff, Wüstenberg és Fischer, 2016).
- f) Elérhetővé válik az adaptív tesztalgoritmus; amelynek segítségével pontosabbá válik a tudás- és képességszintbecslés (pl.: Magyar és Molnár, 2013; Frey, 2007; Molnár, 2015a);
- g) bővül a tesztelésbe bevonhatóak köre (pl. meghallgathatóak a feladatok; Csapó, Molnár és R. Tóth, 2008; Molnár, 2015b) és
- h) lehetővé válik a kontextuális adatok hatékony (pl. például mérhetjük az egyes itemek megoldásához szükséges időt, feltérképezhetjük hányszor javított a diák a megoldásán, hány és hányszor kattintott a tesztelés folyamán) rögzítése és elemzése is (Csapó, Lőrincz és Molnár, 2012; Bodmann és Robinson, 2004). Ennek következtében a papíralapú tesztelésnél megszokott egyedüli indikátor, a teszteredmény helyett gazdag és jól strukturált, a diák tesztelés alatt mutatott viselkedésének pontosabb követését lehetővé tevő adatbázis áll rendelkezésünkre (Molnár és Lőrincz, 2012).
- i) Javulhatnak a tesztek jószágmutatói (Jurecka és Hartig, 2007; Ridgway és McCusker, 2003; Csapó, Molnár és Nagy, 2014, 2015).

2.4. Technológiaalapú mérés-értékelés kialakulása, fejlődése, nemzetközi és hazai tendenciái

A technológiaalapú, azon belül is kiemelt szerepet játszó számítógép-alapú teszteléssel kapcsolatos kutatások már három évtizedes múlttal rendelkeznek. A három évtized alatt jelentős mértékben változott a technológia mérés-értékelésben betöltött szerepe és változtak a számítógép-alapú tesztelés lehetőségei. A kezdetekben – 90-es évek elején – még magán a technológián és az egyes eszközök adta mérés-értékelési lehetőségeken volt a hangsúly, kiemelt szerepet kapva a drága, kis számban, esetlegesen csak kognitív laborokban elérhető eszközöket. Ebből adódóan a technológiaalapú tesztelés kezdeti fázisában még nem a papíralapú tesztelés egy lehetséges alternatívájaként, felváltójaként jelent meg a kutatásokban (*Baker és Mayer, 1999*), hanem egy új lehetőségeket kínáló, de alapvetően drága tesztelési módként.

Egy évtizeddel később, az ezredforduló körüli években, amikor rendszeressé váltak a nemzeti monitor-vizsgálatok és nemzetközi felmérések, a technológiai eszközök, különös tekintettel a számítógép elterjedésével ismét a mérés-értékelés fókuszába került a technológiaalapú tesztelés. Ezen a ponton már kiemelt szerepet képviselt a széles körben rendelkezésre álló számítógépek segítségével történő adatfelvétel, mint a korábbi, papíralapú tesztelés felváltásának, a továbblépésnek egy lehetséges alternatívája (*Csapó, Ainley, Bennett, Latour és Law, 2012; OECD, 2010a; Bennett és mtsai, 1999*). A kapcsolódó kutatások alapvetően nem a technológia adta plusz lehetőségekre, hanem a médiahatás vizsgálatára koncentráltak, ami ma már jól dokumentált kutatási területnek számít.

A számítógépes tesztelés szélesebb körű alkalmazása a világ különböző régióiban az oktatás helyi sajátosságaitól függően más-más formában indult el (bővebben l. erről *Molnár, 2010; Csapó, Ainley, Bennett, Latour és Law, 2012*). A leglassúbb fejlődés Európában, míg a leggyorsabb az USA-ban volt tapasztalható. Előbbi oka részben a tesztelési hagyományok hiánya, részben pedig a technológia lassúbb terjedése.

Európában az első országos és teljes körű számítógépes felmérésre Luxemburgban került sor, ami az országban tanuló diákok száma miatt bizonyos szempontból leegyszerűsítette a technológiai problémákat. A Luxemburgban kidolgozott TAO (a *Test Assisté par Ordinateur* rövidítése) nevű nyílt forráskódú tesztelő platformot (l. <http://www.taotesting.com/>) több ország is átvette, közöttük kezdetben a Szegedi Tudományegyetem Oktatásméleti Kutatócsoportja is jelentős mértékben hozzájárult a fejlesztésekhez.

Ázsia több országában jelentős szerepet kap, a tanterv részét képezi a technológia oktatási alkalmazása, ezért az értékelésben is megjelentek és rohamosan terjednek a különböző technológiai eszközök (*UNESCO, 2014*). Szingapúrban 1997-ben, Hong Kong-ban 2001-ben zajlott le vonatkozó átfogó tantervi reform, aminek hatására az infokommunikációs készségek fejlesztése meghatározó szerepet kapott az oktatásban. Japán e-tanulás stratégiája az 1990-es évek vége óta meghatározó és jelentős mértékben befolyásolja az oktatás fejlesztését is.

Az Amerikai Egyesült Államokban nem csak az oktatásban, hanem az élet más területein (humán erőforrás menedzsment, katonaság) is jelentős hagyományai vannak a tesztek alkalmazásának. A jelentős mérési hagyományok következtében természetesnek tekinthető, hogy Amerikában kezdődtek el legkorábban a számítógép-alapú tesztelésben rejlő lehetőségeket kihasználó kutatási-fejlesztési programok.

A nemzetközi tendenciákat figyelemmel kísérve, követve és segítve a vonatkozó fejlesztéseket hazánkban is jelentős előrelépés történt a számítógép-alapú tesztelések területén, azonban e fejlesztések nem a nagy téttel bíró tesztek számítógépesítésére, hanem a diák és

iskola szintjén is kis tétellel bíró diagnosztikus értékelés irányába történtek. A Szegedi Tudományegyetem Oktatásméleti Kutatócsoportja egy uniós finanszírozású projekt keretein belül Diagnosztikus mérések fejlesztése címmel a TAO platform implementálását követően (Csapó, Molnár és R. Tóth, 2009; Latour és Martin, 2007) elindította saját online értékelési rendszerének kidolgozását, az eDia fejlesztéseket (Molnár és Csapó, 2013; Molnár, 2015a, 2015b; Molnár, Papp, Makay és Ancsin, 2015). A program az online diagnosztika koncepciójának kidolgozásától a mérések tartalmi kereteinek fejlesztésén, egy elektronikus platform létrehozásán és feladatbakokkal való feltöltésén keresztül a gyakorlati kipróbálásig a teljes kutatási-fejlesztési és implementálási folyamatot átfogja (Steklács, Molnár és Csapó, 2015; Korom, Molnár és Csapó, 2015; Csikos, Molnár és Csapó, 2015). Ennek következtében a diagnosztikus értékelés technológiai megvalósítása terén jelenleg Európában Magyarországon folyik a legjelentősebb fejlesztési program. Az eDia (elektronikus diagnosztikus mérés-értékelési rendszer) kidolgozásának célja egy 1–6. évfolyamos diákok és pedagógusok által is használható, a tanítási, fejlesztési folyamat tervezését és megvalósítását segítő rendszer kialakítása.

2.4.1. A disszertációban felhasznált kutatásaink módszertanának áttekintése

A disszertációban 14 kutatás eredményére támaszkodom. A kutatásba bevont vizsgálatok módszertana, mintajellemzője a vonatkozó fejezetekben részletesen kifejtésre kerül. A számítógép-alapú képességvizsgálatokat megelőző papíralapú kutatásaink mintáinak főbb jellemzőit a 2.1. táblázat foglalja össze, ezért a 2002 és 2011 közötti időszakban kivitelezett öt empirikus adatfelvétel közel 24.000 diákjára vonatkozó adatokat nem integráltuk a 2.2. táblázatba, ami a technológiaalapú adatfelvételek főbb jellemzőit ismerteti.

A kutatások felépítésének, a minta kiválasztásának, az adatfelvételi pont megválasztása során minden esetben figyelem előtt tartottuk, hogy a kutatások eredményei egyrészt alkalmasak legyenek nemzetközi szinten is új eredmények létrehozására, másrészt az azonos konstruktum mérését célzó vizsgálataink eredményei összekálázhatóak legyenek, azaz biztosított legyen a megfelelő mértékű horgonyzás a különböző évfolyamokon és adatfelvételi pontokon keresztül is. Az egyes kutatások originalitására a vonatkozó fejezetben térek ki részletesen.

A vizsgált konstruktumok között kiemelkedő szerepet tölt be a problémamegoldó képesség. E képesség mérésnek változásán keresztül szemléltetem a technológia mérés-értékelés terén gyakorolt hatását, azokat a lehetőségeket, amelyek megvalósítása számítógép-alapú mérés-értékelés nélkül nem lenne lehetséges. A problémamegoldó képesség vizsgálata kezdetben hagyományos adatfelvételi technikákkal (első generációs tesztek), majd 21. századi adatfelvételi technikák alkalmazásával (második és harmadik generációs tesztek) valósult meg. Mára a problémamegoldó képesség olyan dimenziójának mérése (dinamikus problémamegoldó képesség) is megvalósítható, melyre 15 évvel ezelőtt, hagyományos technikák alkalmazásával még nem volt lehetőség. A technológiaalapú mérést feltételező logfájlelemzések és modern elemzési eljárások segítségével olyan kutatási kérdésekre is megtalálhatjuk a választ, amelyekre szintén nem volt lehetőség egy évtizeddel ezelőtt.

A 2.2. táblázatban felsorolt kutatások adatfelvételei alapvetően online adatfelvételek voltak, a kérdőívek, illetve tesztek kiközvetítését 2008-ig a TAO, majd a 2008 utáni kutatásokban az eDia platform segítségével történtek. A számítógép-alapú adatfelvételek az

iskolák számítógéptermeiben, az iskolák infrastruktúráját használva valósult meg. A tesztek, kérdőívek megoldására, kitöltésére egy tanóra, azaz 45 perc állt a diákok rendelkezésére.

2.1. táblázat. Az adatelemzésekbe bevont papíralapú kutatások mintáinak főbb tulajdonságai

Adatfelvétel éve	Évfolyam	N	Főbb tulajdonság	Anyá átlagos iskolai végzettsége (adatok hány százalékában ismert)
2002	3.	543	nagyvárosi diákok	3,51 (72,0)
	4.	558		3,28 (82,0)
	5.	515		3,15 (87,6)
	6.	556		3,28 (91,5)
	7.	498		3,18 (87,6)
	8.	491		3,03 (91,0)
	9.	614		3,22 (93,2)
	10.	590		3,22 (94,1)
	11.	525		3,26 (92,6)
2004	3.	1015	hátrányos helyzetű diákok	n. a.
	4.	1029		n. a.
	5.	998		3,82 (79,9)
	6.	1047		3,56 (88,3)
	7.	1053		3,44 (92,0)
	8.	912		3,60 (90,2)
2006	7. (követéses)	937	hátrányos helyzetű diákok	3,82 (70,6)
	7.	3451	országos reprezentatív	3,83 (91,0)
	11.	2376	országos reprezentatív	3,94 (77,8)
2010	7.	3572	országos reprezentatív	3,70 (73,6)
2011	3.	309	illesztett országos reprezentatív	3,95 (61,8)
	4.	373		4,32 (47,2)
	5.	470		3,97 (58,7)
	6.	432		4,02 (61,1)
	7.	520		4,06 (78,5)
	8.	320		4,06 (82,5)
	11.	218		3,92 (100,0)
2002–2011 együtt	3-11.	23922	-	3,65 (74,8)

Az elemzések során alkalmazott eljárások között szerepelnek klasszikus és valószínűségi tesztelméleti elemzések, valamint strukturális egyenletekkel történő elemzések. Az elemzések során az SPSS21.0, SigmaPlot12.0, MPlus6.0 és ConQuest programcsomagokat használtuk.

2.2. táblázat. A disszertációban bemutatásra kerülő technológialapú empirikus kutatások listája és főbb jellemzői

Kutatás megnevezése	Adat-felvétel időpontja	Vizsgált jelenség	Elemzésbe bevont	Vizsgált életkori sáv
Papír- és számítógép-alapú tesztelés összehasonlító vizsgálata kisiskolás diákok körében	2010	induktív gondolkodás - médiahatás	PP: 952 CB: 377	1. évfolyam
A dinamikus problémamegoldó képesség fejlődése 5-11. évfolyamon	2011	dinamikus problémamegoldó képesség	855	5-11. évfolyam
Az iskolakezdés kulcsfontosságú előkészítéseinek online mérése kisiskolás korban	2012	beszédhanghallás, relációszőkincs, elemi számlálás, tapasztalati összefüggés-megértés, tapasztalati következtetés	FF: 364/5423/ 1895/424/ 1094 CB: 364/1740/ 435/402/ 416	1. évfolyam
A dinamikus problémamegoldó képesség fejlődése 3-9. évfolyamon	2013	dinamikus problémamegoldó képesség	1291	3-9. évfolyam
Egér és billentyűzethasználati képességek fejlettségi szintje kisiskolás diákok körében	2013	egér és billentyűzethasználat	1195	1-4. évfolyam
Egér és billentyűzethasználati képességek fejlettségi szintje első évfolyamos diákok körében	2014	egér és billentyűzethasználat	4952	1. évfolyam
A dinamikus problémamegoldó képesség fejlettségi szintje 7. évfolyamon (longitudinális kutatások keretein belül kivitelezett adatfelvétel)	2014	dinamikus problémamegoldó képesség	1191	7. évfolyam
A dinamikus problémamegoldó képesség fejlődése 3-11. évfolyamon	2014	dinamikus problémamegoldó képesség	4371	3-11. évfolyam
A dinamikus problémamegoldó képesség fejlettségi szintje első évfolyamos egyetemisták körében	2015	dinamikus problémamegoldó képesség	1259	egyetem, 1. évfolyam (BA, osztatlan, nappali)

3. A TECHNOLÓGIAALAPÚ TESZTELÉS ALKALMAZHATÓSÁGA KISISKOLÁS KORBAN

Miután a jelentős, nemzetközi szummatív mérés-értékelési programok idősebb diákok tudás- és képességszintmérését célozzák, a technológiaalapú tesztelésre történő átállást segítő kutatások mintája is rendszerint idősebb korosztályokra fókuszálnak. Kisiskolás tanulók körében lényegesen kevesebb kutatás valósult meg (*Carson, Gillon és Boustead, 2011; Csapó, Molnár és Nagy, 2014, 2015*). Ugyanakkor, amíg idősebb diákok esetében feltételezhető a megfelelő szintű számítógép-használati jártasság, addig kisiskolás korban ez a kritérium nem feltétlenül adott.

Az óvodai, illetve az iskolába lépés időszakában megvalósuló technológiaalapú mérések esetében ezért elengedhetetlen annak vizsgálata, hogy a diákok képesek-e a különböző adatbeviteli eszközök adekvát használatára, felkészültek-e az online tesztekkel történő empirikus kutatásokra. A nem megfelelő szintű eszközhasználat komoly validitási problémákat eredményezhet, valamint a tesztelés frusztrálóvá válhat a gyerekek számára, és ellenézés alakulhat ki bennük a technológiaalapú mérésekkel szemben.

3.1. Az egér- és billentyűzethasználati képességek online mérése kisiskolás diákok körében

A kutatás fő célja annak feltérképezése, hogy (1) egy egységes, vagy az egér- és a billentyűzethasználatot tekintve kétdimenziós, vagy műveletenkénti bontásban háromdimenziós konstruktumnak tekinthető-e az egér- és a billentyűzet-használati képességek; (2) alkalmas-e az iskolába lépő tanulók egér- és billentyűzet-használati képessége arra, hogy egyéb tudás- és képességszintjüket online mérjük; (3) vannak-e olyan egér- és/vagy billentyűzethasználatot igénylő műveletek, amelyek problémát jelentek számukra, ezért tesztbeli alkalmazásuk kerülendő; (4) a feladatokon, az egyes műveleteken nyújtott teljesítményüket időkorlát alkalmazása esetén mennyiben befolyásolja egér- és billentyűzetkezelési képességeik fejlettségi szintje; (5) hogyan változik és miként jellemezhető 1–4. évfolyamos diákok egér- és billentyűzethasználati képességeinek fejlettségi szintje; (6) milyen háttérváltozók befolyásolják az egér- és billentyűzethasználatra vonatkozó képességeik fejlettségi szintjét.

Mind a pilot vizsgálatra kidolgozott, mind a pilot mérés eredményei alapján a nagymintás adatfelvételhez átdolgozott 44 ítemes teszt feladatai alapvetően két csoportba sorolhatók: (1) egérkezelést, (2) billentyűzetkezelést tartalmazó feladatok. A feladatok fokozatosan nehezedtek a teszten belül. Az egyes műveletek elvégzésének gyorsaságát nemcsak az idő logolásával, hanem időkorlátos feladatok alkalmazásával is mértük. Az egérkezelési feladatokban a kurzor/mutató megfelelő helyre történő navigálása mellett a kattintási, kijelölési műveletek pontosságát és a vonszolás technikájának elsajátítási szintjét is vizsgáltuk.

A nagymintás adatfelvétel 44 íteemből álló tesztjének reliabilitásmutatója (Cronbach- α) 0,89 volt, azaz a teszt alkalmasnak bizonyult a diákok egér és billentyűzethasználati képességei fejlettségi szintjének meghatározására. A vizsgált konstruktum dimenzionalitásának vizsgálatára irányuló elemzések alapján a háromdimenziós mérési modell, ahol külön dimenzióba került a billentyűhasználat és az egérhasználati készségeken belül a kattintás és

vonszolás művelete, illeszkedett leginkább az adatokhoz ($\chi^2=201,313$, $df=2$, $p<0,001$), azaz e képességet nem tekinthetjük egy egydimenziós konstruktumnak. A kattintás, a vonssolás és a gépelés műveletei más-más készségek működtetését igénylik, ezért egyik működéséből nem feltétlenül következtethetünk a másik hasonló szintű működésére, azaz műveleti szinten mindhárom mérése-értékelése, fejlesztése lényeges feladat.

Az eredmények értelmében a legfiatalabb tanulók részére kiközvetített tesztek kidolgozása során javasolt a következők figyelembe vétele: itemtípusok tekintetében elsősorban a kattintásra alapozó feladatok, esetlegesen a néhány gyakori karakter begépelését igénylő feladattípusok, illetve a nagy objektumok nagy területre történő mozgatását kívánó feladatok használata ajánlott. A kis elemek kis területre történő mozgatására épülő vonssolás típusú műveleteket, illetve a bonyolultabb gépelést igénylő feladatok vagy időkorlátos feladatok tesztbe történő integrációja esetén a teszten nyújtott teljesítményt erősebben befolyásolhatja a diákok egér- és billentyűzet-használati képességei, ezért javasolt azok kerülése az első évfolyamos diákoknak írt feladatokban. A diákok olvasási képességének fejlettségi szintjében lévő különbségek fülhallgató és a feladatok meghallgathatóságának biztosításával kiküszöbölhető.

A műveletek feltételezett könnyű taníthatóságát támasztja alá az az eredmény, hogy összevetve az azonos műveletek elvégzését kívánó, tesztben korábban és később szereplő itemek nehézségi indexét, egyértelmű tendencia bontakozik ki: művelettől függetlenül a tesztben korábban szereplő, azonos művelet elvégzését igénylő item nehezebbnek bizonyult, mind a később szereplő item.

A háttérváltozókkal történt összefüggés-vizsgálatok eredménye alapján megállapítható, hogy sem a diákok neme, sem általánosságban a szülők által biztosított gazdasági háttér nem befolyásolta technológiai műveltségük e szeletének fejlettségi szintjét. Feltételezésünkkel ellentétben a diákok géphasználati gyakorisága nem bizonyult meghatározó tényezőnek, csak az, hogy valamilyen gyakorisággal használnak-e számítógépet, internetet. Ennek oka lehet, hogy egyrészt a technológiai műveltség online tesztek megoldásához szükséges képességszintjét jóval kevesebb számítógép-használattal is el lehet érni, annak nem feltétele az akár napi több óra internetezés vagy számítógépezés, másrészt az érintett fiatal korban a napi több órát internetezők a táblagépek elterjedtsége miatt feltételezhetően érintőképernyős számítógépekkel játszanak és nem egér-, vagy billentyűzethasználatot igénylő eszközökkel.

3.2. Az iskolakezdés kulcsfontosságú előkészítéseinek online mérése kisiskolás korban

A felkészült iskolakezdés alapvetően meghatározza a későbbi évek tanulási eredményeit, a sikeres iskolai tanulmányok kulcsfontosságú feltétele (Nagy, 2000, 2007). E feltétel meglétének mérése, diagnosztizálása kutatási-fejlesztési feladat. Olyan megbízható, érvényes és könnyen alkalmazható tesztek (Snow, 2006), mérőeszközök kidolgozását igényli, amelyekkel pontosan meg lehet állapítani, hol tart egy gyermek a sikeres iskolakezdéshez szükséges készségek fejlődése tekintetében (Snow és Van Hemel, 2008).

A Szegedi Műhelyben négy évtizedes hagyománya van az iskolakészültség tesztek fejlesztésének, amelyek eddig kizárólag szemtől szembeni, azaz egyéni adatfelvételben valósultak meg. A továbblépés természetes lépése a technológia mérés-értékelésben nyújtott előnyeinek kiaknázása.

A fejezet középpontjában álló iskolaérettségi tesztet, a DIFER-t úgy tervezték, hogy a használata ne igényeljen különleges szakértelmet. Ez egyrészt a teszt sorozat hátránya is, ugyanis könnyen sérülhet a tesztek felvételének objektivitása. Előfordulhat, hogy a tesztet felvevő pedagógus kicsit másképpen olvassa fel az instrukciókat az egyes tanulóknak, és a válaszok pontozása is kismértékben változó lehet. A teszt sorozat alkalmazhatóságának másik korlátja, hogy csak személyesen és egyénileg végezhető el felvétele, így egy tanuló teljes körű felmérése akár több órát is igénybe vehet. A feladatok online közvetítésével, mindkét probléma kiküszöbölhetővé válhat, előbbi esetén például előre rögzített hanginstrukciókkal (a szövegeket képzett szakemberek olvassák fel) és a válaszok automatikus pontozásával, utóbbi pedig a számítógép-alapú adatfelvétel adta csoportos tesztfelvételi lehetőséggel.

A disszertáció vonatkozó fejezetében ismertetett elemzésekbe a DIFER beszédhanghallás, relációsózikincs, tapasztalati összefüggés-megértés, tapasztalati következtetés, elemi számolás tesztjei mellett bevontunk egy induktív gondolkodás tesztet is, ami eredeti formájában nem szemtől szembeni, hanem papíralapú adatfelvétellel valósult meg.

A hat teszt különböző pszichológiai jellemzőket mér, és az online változataik különböző technológiai megoldásokat igényelnek. A mérőeszközöknek ez a változatos összetétele lehetőséget biztosít annak bemutatására, hogy a technológiaalapú tesztelés nem csupán kivitelezhető ilyen fiatal korban, hanem sok tekintetben még alkalmasabb is, mint a személyes felmérés vagy papíralapú tesztelés.

A kutatások célja annak feltérképezése volt, hogy alkalmazhatóak-e ezen speciális, a területek sokfélesége miatt változatos, ezért számos kihívást is jelentő tesztek a hagyományos oktatási-nevelési gyakorlatba integrálva, az érintett, iskolába lépő célcsoportban csoportos adatfelvétel céljából. Milyen megbízhatósággal tehető ez, növelhető-e a tesztek reliabilitása a tesztíngerek feletti kontrol és a kiértékelés objektivitásának növelésével? Változik-e a tesztek validitása és a diákok eredményei azért, mert ők számítógép előtt és nem szemtől szemben, egyéni adatfelvétel során, vagy papíralapú környezetben oldják meg a tesztek? Ugyanúgy viselkednek-e az alacsonyabb és a magasabb képességszintű diákok a megváltozott környezetben? Mennyiben befolyásolja a tesztek tartalma és a teszteken belül alkalmazott itemtípusok a hagyományos és a számítógépes környezetben nyújtott teljesítményeket?

Az öt teszt digitális változatát más-más diákok oldották meg. A számítógépes adatfelvételben tesztenként kb. 400 diák vett részt. A beszédhanghallás teszt esetében a személyes és a számítógépes adatfelvétel mintája megegyezett, a többi esetben nem. A relációsózikincs, az elemi számolás és a tapasztalati következtetés tesztek online eredményeit összevetettük egy egyéni adatfelvételen alapuló országos reprezentatív mintán felvett felmérés eredményeivel. A relációsózikincs teszt mind a négy változatát használtuk, ezért mindegyiket a teljes minta egynegyedén vettük fel. A tesztek számítógépesítése során különös figyelmet fordítottunk arra, hogy ne legyen szükség speciális számítógépes készségekre a teszt kitöltéséhez.

A tesztek belső konzisztenciája (Cronbach- α) megfelelőnek bizonyult mind az egyéni, mind a számítógép-alapú adatfelvétel során. Előbbi esetben a reliabilitásmutatók értékei 0,71 és 0,89 között mozogtak, utóbbi esetben pedig 0,74 és 0,94 között. Általában az online adatfelvétel mutatói magasabbnak bizonyultak, mint a szemtől szemben egyénileg történt adatfelvétel eredményei alapján számolt mutatók értékei. Az kisiskolás diákok részére kidolgozott 37 itemes induktív gondolkodás teszt belső konzisztenciája mind papír, mind számítógépes környezetben megfelelőnek bizonyult (Cronbach- α_{PP} =0,86; Cronbach- α_{CB} =0,86). A reliabilitásmutató értéke nem változott a közvetítő eszköz megváltozása hatására.

Ennek oka lehet, hogy jelentős viselkedésbeni módosítás nélkül történt a papíralapú teszt digitalizációja.

A mérési invariancia tesztelése után megállapíthattuk, hogy a beszédhanghallás és az induktív gondolkodás teszt esetén összevethetőek a teszteken nyújtott teljesítmények, miután áll a mérési invariancia. A relációszókincs, elemi számolás és tapasztalati következtetés tesztekénél mind a metrikus, mind a Lagrange-multiplikátor teszt alapján felállított részleges metrikus invariancia modell szignifikánsan gyengébben illeszkedett, mint a konfigurális modell, azaz nem beszélhetünk mérési invarianciáról. A két tesztelési módból adódó különbségeknek számos oka lehet. Annak ellenére, hogy szigorúan véve a mérési invarianciaelemzések eredménye alapján csak a beszédhanghallás és induktív gondolkodás teszten nyújtott teljesítmények fejezhetőek ki egy skálán, azaz hasonlíthatóak össze, a tesztek tartalmának, a feladatok típusának, és a vizsgált személyek jellemzőinek hatásai elemzése céljából a többi teszten nyújtott teljesítményeket is górcső alá vettük.

Elemeztük a tesztelés megváltozott módjának hatását a diákok teljesítményre. Miután a DIFER kritériumorientált tesztekben áll, ezért az első osztályos tanulók többsége már elérte az iskolakezdéshez szükséges fejlettségi szintet. Az eredmények alapján nyolc esetből hét esetben a számítógép-alapú teszteken gyengébb eredményt értek el a tanulók, mint a hagyományos, személyes adatfelvételen alapuló teszteken.

A beszédhanghallás teszt igen magas eredményeiben jelentős mértékű esés volt megfigyelhető (91,4 százalékról 82,6 százalékos teljesítményre). A relációszókincs teszt négy változatában mindkét irányú változásnak tanúi lehettünk: jelentős teljesítménycsökkenés két esetben, kismértékű növekedés a másik kettőben. Az elemi számolás esetében nem észleltünk számottevő változást. A különbségnek két magyarázata is lehet. Egyrészt a tesztet felvevő pedagógusok gyakran magasabb pontszámokat adnak a gyerekeknek, mint a számítógép automatikusan (és objektíven) kiszámolt pontszámait, vagy a tanulóknak nehézséget jelentett a számítógép kezelése, ami rontott teljesítményükön. Miután sem az induktív gondolkodás teszt kapcsán, sem az elemi számolás esetén sem találtunk utóbbira utaló jelet, ezért a számítógép kezelési nehézség nem lehet kielégítő magyarázat a teljesítménykülönbségekre. Valószínűbb, hogy a tesztek felvevő pedagógusok elnézőbbek voltak a gyerekek válaszaival szemben.

E kutatási eredmény megerősíti az előző részben ismertetett eredményeket, miszerint kisiskolás diákok körében is sikeresen alkalmazhatóak a számítógép-alapú tesztek, sőt a diákok egér- és számítógéphasználati képessége fejlettségi szintjére vonatkozó eredményeket figyelembe véve – igény esetén – kidolgozhatóak olyan tesztek, melyek viselkedése azonos papír- és számítógép-alapú környezetben.

Az elemzések korlátja, hogy az eredeti DIFER tesztek az óvodából az iskolába átlépő gyerekek mérésére tervezték, ezért a már iskolás gyerekekénél gyakran plafoneffektusba ütköztünk az elemzések során, ami nem kedvezett a tesztek működésének vizsgálatához. Ez kiküszöbölhető egy jövőbeni kutatással, melynek célpopulációját az óvodások alkotják, annak ellenére, hogy a számítógép-használat ebben a korban még további megvalósíthatósági tanulmányokat igényel (Csapó, Molnár és Nagy, 2015).

4. A SZÁMÍTÓGÉP-ALAPÚ TESZTELÉS ALKALMAZÁSA 9-19 ÉVES KORBAN: A PROBLÉMAMEGOLDÓ KÉPESSÉG MÉRÉSI LEHETŐSÉGEINEK VÁLTOZÁSA

A problémamegoldó képesség az emberiség egyik legfontosabb, túlélését és fejlődését leginkább meghatározó képessége. Megvalósítja, hogy tetteink előtt gondolkodjunk, ami jelentős mértékben megnöveli az ember és környezete közötti interakció hatékonyságát, és biztosítja, hogy lehetőségeink szerint a legjobb döntést hozzuk meg az adott szituációban. Problémamegoldó képességünk lehetővé teszi, hogy olyan célokat, megoldási lehetőségeket is megvalósítsunk, melyeket különben nem érünk el.

A disszertáció e fejezetének célja a mérés-értékelés módszertanában történő fejlődési tendenciák szemléltetése a problémamegoldó képességgel kapcsolatos empirikus kutatások és a megválaszolható kutatási kérdések körének bővülésén, változásán keresztül; külön hangsúlyt fektetve (1) a pedagógiai kontextusban történő mérési lehetőségekre; (2) a jelen kutatások előzményeként az elmúlt 10 év vonatkozó papíralapú tesztekre alapozó kutatásaink szintetizálására; (3) a számítógép-alapú tesztelés nyújtotta lehetőségek és igények ismertetésére. A fejezet második felében (4) több, egymásra épülő, különböző korosztályban megvalósuló harmadik generációs számítógép-alapú, dinamikus problémamegoldó képesség fejlődését összességében tág életkori intervallumban vizsgáló kutatás-sorozat céljairól, módszereiről és az alkalmazott mérőeszközök, problémák működéséről számolunk be.

A kutatások felépítése a valószínűségi tesztelmélet adta eszközrendszer felhasználásával lehetővé teszi, hogy az elmúlt tíz év kutatási eredményeit egyesítve közös képességskálán jellemezzük a 3–11. évfolyamos diákok problémamegoldó képességének fejlődését, a különböző típusú problémák megoldási sikerességének alakulását, valamint nem, iskolatípus és szocioökonómiai faktor mentén összehasonlítsuk a képzett résztminták teljesítményének változását. A fejezetben bemutatott elemzésekbe hat nagymintás empirikus vizsgálat adatbázisát, azaz közel 24000 diák adatát vontuk be (4.1. táblázat).

A közel tíz évvel ezelőtt induló kutatássorozat keretein belül a problémamegoldó képességgel kapcsolatban egy új modell és párhuzamosan egy új szemléletű tesztsomag kidolgozására került sor. A modell integrálja az amerikai és az európai irányzatok szemléletét, valamint a PISA vizsgálatok műveltségkoncepcióját. A modell értelmében kidolgozott tesztek a matematika és a természettudományok területén életszerű, szemantikailag gazdag környezetben vizsgálják a diákok alkalmazható tudását, valamint tág életkori intervallumban és iskolai kontextusban valószínűsítik meg a diákok képességvizsgálatát.

A valószínűségi tesztelmélet adta eszközöket kihasználva az egyes keresztmetszeti vizsgálatok tesztjeit úgy dolgoztuk ki, hogy azok – mind az adott mérés keretein belül szereplő különböző nehézségű tesztek, mind a különböző adatfelvételek során alkalmazott tesztek – horgonyitemek segítségével összekapcsolhatóak legyenek. A mérések során összesen 100 jól működő item kidolgozása és paraméterezése valósult meg, az elemzésekből kizártuk a kevésbé jól működő itemeket.

A problémamegoldó képesség tesztjein életszerű helyzetekbe (pl. iskolai vagy családi kirándulás, pizzarendelés, vásárlás) ágyaztuk az alapvetően matematikai eszközökkel megoldandó problémákat, melyek három csoportba sorolhatók: (1) a megoldásához szükséges információt tartalmazza a feladatlap, ám nem a megszokott iskolai, hanem életszerű formában;

(2) a megoldásához nem minden információ adott, de a hiányzó információk a tananyag részét képezik; (3) a megoldásához nem található meg minden háttérinformáció, azonban azokkal a diákok, ha nem is az iskolai tanórán, de a hétköznapi életben találkozhattak.

Az eredmények szerint 2002 és 2011 között minden évfolyamon átlagos képességszint-csökkenés figyelhető meg, ami nem magyarázható kizárólagosan az egyes adatfelvételek mintájának speciális tulajdonságaival. A 2011-es adatfelvétel 3. évfolyamos diákokra vonatkozó eredményei e tendencia megváltozását jelzik. Összességében az oktatás nem csökkenti, ám nem is növeli tovább a hátránnyal indulók helyzetét ezen a képességterületen.

A problémamegoldó képesség fejlődése is – hasonlóan más gondolkodási képesség fejlődéséhez – jól jellemezhető egy négyparaméteres logisztikus görbével. A görbe inflexiós pontja, mérési ponttól függetlenül, 7. évfolyamra esik, azaz ebben az életkorban a leggyorsabb a fejlődés és ezért hatékony a fejleszthetőség, majd ezen életkor után az addig gyorsuló fejlődés lassuló növekedésbe vált. A fejlődés mértéke – hasonlóan más általános gondolkodási képesség fejlődéséhez lassú, 3–11. évfolyamok teljes időszakát tekintve évenként átlagosan a szórás ötöde. 4. évfolyamon minimális, 7-8. évfolyamon intenzív fejlődés figyelhető meg, míg extrapolálva a fejlődési folyamatokat a 3. évfolyam előtt és a 11. évfolyam után is, bár lassuló ütemben, de folytatódik e képesség fejlődése. Több éves fejlődésnek megfelelő mértékű teljesítménykülönbség van a különböző iskolatípusban tanuló diákok között, ami egy évfolyamon belül a diákok szintjén még markánsabban realizálódik.

A három különböző problématípus megoldási sikeressége alapján megállapított képességszintek közötti különbség mértéke egy évfolyamon belül is jelentős, években mérhető. A valós helyzetekben történő problémamegoldás sikerességét leginkább tükröző rosszul definiált, hiányos és szemantikailag gazdag problémahelyzetek évfolyamtól függetlenül jelentős kihívást jelentettek a diákoknak. Képességfejlődésük alacsony mértéke arra utal, hogy e típusú problémahelyzetek megoldására egyáltalán nem készíti őket fel az iskola és spontán fejlődése igen lassú.

A szülő iskolai végzettségének korábban jelentős mértékű teljesítménybefolyásoló szerepe változott az elmúlt tíz év alatt. Egyrészt általános tendencia, hogy az oktatás expanziójával a szülők átlagos iskolai végzettsége is növekedett, ez azonban – a várakozásokkal ellentétben – a problémamegoldó gondolkodás képességterületén nem vonta maga után a diákok átlagos képességszintjének növekedését. Másrészt új tendencia, hogy a korábban az egyik legjelentősebb előrejelző és képességszint-befolyásoló tényezőnek számító anya iskolai végzettsége teljes mértékben megszűnőben van, 2011-ben egyik évfolyamon sem volt kimutatható közepes vagy erős szignifikáns kapcsolat a diákok problémamegoldó képességének fejlettségi szintje és az anyák iskolai végzettsége között.

Az utóbbi években tapasztalt jelentős mértékű fejlődés a pedagógiai mérés-értékelés terén, valamint a technológia rapid változása, rohamos elterjedése és iskolai megjelenése új lehetőségeket teremtett a problémamegoldó képesség vizsgálata terén, háttérbe szorítva a korábbi papíralapú, statikus problémahelyzetekre alapozó kutatásokat, és előtérbe állítva a 21. századra jellemző dinamikusan változó, interaktív problémahelyzetek segítségével megvalósuló méréseket. A technológia adatfelvételbe történő integrációja olyan új eszközöket (pl.: multimédia alkalmazása, dinamikusan változó problémakörnyezet, társas problémamegoldás) kínált és kínál a problémamegoldó képesség fejlődését és fejlesztési lehetőségeit vizsgáló kutatók számára, melyek forradalmasítják a korábbi problémamegoldásra vonatkozó modelleket és tesztek.

4.1. A problémamegoldó képesség számítógép-alapú mérése: célok és módszerek

A 21. század polgára már nem boldogulhat az életét minden területen körbevevő technológiai eszközök használata nélkül, melyek használatához az adott géppel történő interakciók sorozatára van szükség: különböző gombok bizonyos sorrendben történő lenyomására, tekerésére, esetleg az eszközt vezérlő program érintőképernyőn keresztül történő irányítására. A mindezt lehetővé tevő új szoftveres és hardveres technológiák állandó tanulásra és ezzel párhuzamosan problémamegoldásra készítetnek minket. De hogyan lehet ezeket a 21. században kulcsfontosságúnak számító képességeket mérni?

Az oktatási kontextusban alkalmazásra kerülő tesztek eredményeinek pontossága és validitása egyrészt azok elméleti alátámasztottságán, másrészt a tesztek jó pszichometriai jellemzőin múlik (*Benson, Hulac és Kranzler, 2010*). Ha a tanulók interaktív problémamegoldó képességének fejlettségi szintje áll egy kutatás fókuszában, első lépésként szükséges egy minden szempont szerint megfelelő mérőeszköz kidolgozása. A nemzetközi együttműködés keretein belül fejlesztett mérőeszközsorozat minimálisan komplex rendszerekre épít, mely a problémamegoldás két fázisát a problémamegoldó képesség működését jellemző próba és hiba (*trial and error*) stratégia alkalmazásával történő ismeretek elsajátításának és ismeretek alkalmazásának sikerességét méri. A disszertációban bemutatásra kerülő elemzések elsődleges célja

- egy tág életkori intervallumban alkalmazható mérőeszközsorozat összeállítása;
- a mérőeszközök pszichometriai jóságmutatóinak jellemzése mellett a mérőeszközfejlesztés mögött lévő elméleti kétdimenziós mérési modell empirikus verifikációja;
- a különböző csoportok összehasonlítását lehetővé tevő mérési invariancia tesztelése, annak ellenőrzése és számszerűsítése, hogy a teszt ugyanazon felépítésű konstruktumot méri-e (*Byrne és Stewart, 2006*) a különböző évfolyamokon, azaz releváns-e a teszten nyújtott teljesítmények egymással történő összehasonlítása ezen csoportok között;
- a különböző évfolyamos és iskolatípusban tanuló diákok teljesítményének összehasonlítása, fejlődési trendvonalak tulajdonságainak azonosítása és összevetése korábbi problémamegoldó képesség fejlődésére vonatkozó eredményekkel, illetve más, gondolkodási képességeket vizsgáló (pl.: induktív gondolkodás) kutatási eredményekkel;
- az iskoláztatás éveitől a problémamegoldó képesség fejlesztésére érzékeny időszak meghatározására;
- a problémamegoldó képesség, az intelligencia, az induktív gondolkodás és a diákok tanulmányi sikeressége (iskolai jegyek) és gazdasági-társadalmi, szociális háttérváltozói (pl. szülők iskolai végzettsége) összefüggéseinek, strukturális relációinak és azok időbeli állandóságának vagy esetleges változásának feltérképezése, konstruktum validitás tesztelése;
- annak számszerűsítése, hogy lehetséges-e harmadik generációs tesztekkel mérhető képességek fejlettségi szintjének meghatározása, előrejelzése hagyományosnak számító első generációs teszteredmények alapján;
- a logfájlelemzések alapján milyen problémamegoldó stratégiák, milyen típusú profillal jellemezhető problémamegoldók azonosíthatóak;
- a metaadatok szerint hogyan változik a diákok által alkalmazott problémamegoldó stratégia időben a dinamikus problémamegoldó teszt megoldása közben.

A kutatások mintájának főbb tulajdonságait a 2.1. táblázat foglalja össze. A kutatásokban alkalmazott problémák a MicroDYN megközelítésen alapultak, felépítésükben azonosak voltak a PISA 2012 kreatív problémamegoldás modul kutatásban alkalmazott interaktív problémákkal (OECD, 2014). A számítógép-alapú problémák a diákok által kedvelt, ismerős (pl.: mindennapi élet, videojátékok) kontextusban kerültek megfogalmazásra, ugyanakkor szerkezetük miatt számukra újak voltak, a megoldás során előzetes ismereteiket nem tudták alkalmazni. A teszt kiközvetítése az első kutatás keretein belül a TAO platformon, a többi az eDia platform segítségével történt. Az adatfelvételekre az iskolák számítógépes termeiben került sor.

A probléma megoldásának első fázisában a diákoknak fel kellett fedeznie a rendszert, azaz a bemeneti változó értékeit szabadon változtatva és megfigyelve a kimeneti változók értékváltozását, fel kellett ismerni a probléma háttérében lévő összefüggésrendszert. A változók egymással való kapcsolatát nyilak segítségével a probléma alatt található, a bementi és kimeneti változókat megjelenítő modellen meg is kellett jeleníteni. A válaszok kiértékelése során akkor kapott 1 pontot a diák, ha a rendszerben lévő összefüggések mindegyikét pontosan felrajzolta, azaz tökéletes modellt állított el. Ellenkező esetben 0 pontot kapott a probléma e fázisának megoldására.

A problémák második fázisában működtetni kellett a rendszert, azaz megismerve a valódi összefüggéseket (a program megjeleníti a problémamegoldó számára a helyes összefüggésrendszert), a bemeneti változók értékeit állítva elérni a kimeneti változók előre meghatározott célértékeit (részletesen l. Molnár, 2013a). Mindezt a problémák megoldójának maximum 4 lépésben és kutatástól függően 90 vagy 180 másodperc alatt kell elérni. A válaszok értékelése során kizárólag abban az esetben kapott az adott diák 1 pontot a probléma ezen részének megoldására, ha megadott időn és lépésszámon belül sikeresen elérte az összes kitűzött célértéket, ellenkező esetben 0 ponttal értékeltük teljesítményét.

Eredményeink szerint életkortól függetlenül megbízhatóan használhatóak a nemzetközi kooperációban kidolgozott tesztek. A 15 különböző szerkezetű és kontextusú (problémánként 2 itemet tartalmazó), interaktív, dinamikus problémából összeállított harmadik generációs dinamikus problémamegoldó képesség tesztek (DPK) belső konzisztenciája minden egyes évfolyamon megfelelőnek bizonyult (Cronbach- α =0,75-0,92). Az empirikus adaton nyugvó strukturális elemzések eredményei alátámasztották a kétdimenziós elméleti modell felépítését. Az invarianciaelemzések eredményei megalapozták a különböző korú és nemű diákok teszten nyújtott teljesítményének összehasonlíthatóságát, ugyanis nem volt szignifikáns különbség sem az egyre szigorúbb modellek közötti illeszkedésben, sem azok konfigurális modellilleszkedéshez való viszonyításában sem. Ennek következtében a teszten mutatott teljesítmények évfolyamtól és nemtől függetlenül akár nyerspont érték szerint is összehasonlíthatóak egymással, ugyanis ugyanazon látens konstruktum mérésére került sor.

A disszertáció ötödik fejezetében mind fejlődési, mind strukturális perspektívából áttekintjük és összehasonlítjuk a tudás elsajátításában, alkalmazásában és transzferálásában kiemelt szerepet játszó gondolkodási képességek első és harmadik generációs tesztekkel történő mérési eredményeit. Számszerűsítjük az első generációs teszteken nyújtott problémamegoldó képességre vonatkozó teljesítmények előrejelző erejét a harmadik generációs problémamegoldó képességet mérő teszteredmények függvényében, illetve szintetizáljuk az interaktív helyzetekben és tág életkori intervallumban mért problémamegoldó képesség fejlődésére, a fejlődést befolyásoló háttérváltozókra és a logfajlelemzésekre alapozó, a diákok tesztelés során mutatott problémamegoldó stratégiáit és azok változását feltérképező kutatási kérdésekre válaszoló elemzéseket és eredményeket.

5. A DINAMIKUS PROBLÉMAMEGOLDÓ KÉPESSÉG FEJLŐDÉSE ÉS A FEJLŐDÉST BEFOLYÁSOLÓ TÉNYEZŐK

Az 5. fejezetben bemutatott kutatási eredmények kivétel nélkül a 4. fejezetben ismertetett dinamikus problémamegoldó képesség mérésére fókuszáló vizsgálatok adatbázisain alapulnak. A 4. fejezet kapcsolatos elemzései a problémák és a tesztek működését járták körül, az 5. fejezetben pedig a diákok problémamegoldó képességének fejlődése, a fejlődést befolyásoló és előrejelző tényezők feltérképezése, valamint a diákok problémamegoldó stratégiáinak feltérképezése áll a középpontban. Utóbbit a logfájllemzések alapján, a dinamikus problémamegoldó teszten mutatott viselkedés segítségével tesszük.

5.1. A dinamikus problémamegoldó képesség fejlődése

A fejlődési folyamatok ismertetése során a 8. évfolyamos diákok eredményei alapján egy 500 átlagú és 100 szórású skálára transzformáltuk a fejlődést leíró görbét. A dinamikus problémák megoldottsága alapján megállapítható, hogy minden egyes évfolyamon nő vagy stagnál a diákok interaktív környezetben mutatott problémamegoldó képességének fejlettségi szintje. A fejlődés mértéke – hasonlóan a többi gondolkodási képességhez (pl. induktív gondolkodás), az explicit iskolai fejlesztés hiánya miatt – nagyon lassú, évfolyamonként változó, átlagosan évenként a szórás ötöde. A legjelentősebb, az éves átlagos képességfejlődés kétszeresének megfelelő gyorsaságú fejlődés 5. és 7. (80 pont) évfolyam között van, míg 8. évfolyamon stagnálás tapasztalható.

Az empirikus adatokra illesztett négyparaméteres logisztikus görbe jól illeszkedik az adatokhoz ($R^2=0,98$), ezért a görbe paraméterei alkalmasak a képesség fejlődésének jellemezésére. Az illesztett logisztikus görbe inflexiós pontja, azaz a fejlesztés szempontjából szenzitív intervallum, a 6. évfolyamra tehető. Az egy évfolyamon belül tapasztalható fejlettségbeli különbségek nagysága meghaladja a 9-19 éves korosztályban történő fejlődés mértékét. Az iskolaátmeneti pontokon (általános – középiskola, középiskola – egyetem), ha implicit is, de egyértelmű a problémamegoldó képesség, mint tudás elsajátítás és alkalmazás mentén történő szelekció.

5.1.1. A problémamegoldó képességszint kapcsolata az intelligencia és az induktív gondolkodás fejlettségi szintjével

Az eredmények értelmében a diákok statikus problémák megoldása során nyújtott teljesítménye jelentős mértékben összefüggött intelligenciaszintjükkel ($r=0,48$, $p<0,01$). Az összefüggés szorossága erősebbnek bizonyult ($z=2,11$, $p<0,05$), mint az intelligencia és a dinamikus problémákon nyújtott teljesítményszint ($r=0,39$, $p<0,01$) kapcsolata, sőt erősebb volt, mint a dinamikus és a statikus problémamegoldó teszten nyújtott ($r=0,35$, $p<0,01$) teljesítmények kapcsolata ($z=3,84$, $p<0,01$). A parciális korreláció értékei a lineáris (Pearson-féle) korrelációkénál minden esetben szignifikánsan alacsonyabbak voltak, azaz mindhárom kapcsolat erősségében meghatározó szerepet játszott a harmadik konstruktum fejlettségi szintje. A diákok általános intelligenciájának az alsóbb évfolyamokon kisebb, majd egyre erősödő

szerepe mutatkozott mind a területspecifikus és statikus, mind a területfüggetlen és dinamikus problémák megoldása kapcsán.

Az induktív gondolkodás és a különböző problémahelyzetekben mutatott problémamegoldó képesség fejlettségi szintjének kapcsolata közepes volt. A kapcsolat erőssége független volt az alkalmazott probléma dinamizmusától ($r_{SPK}=0,43$ és $r_{DPK}=0,44$, mindkettő $p<0,01$), ugyanakkor az induktív gondolkodás fejlettségi szintje és a problémamegoldó képesség fejlettségi szintje közötti kapcsolat erősebbnek bizonyult ($z=1,80$, $p<0,05$), mint a statikus és a dinamikus problémahelyzetekben nyújtott problémamegoldó sikeresség közötti kapcsolat ($r=0,35$, $p<0,01$). A parciális korreláció értékei a lineáris korrelációkénál minden esetben szignifikánsan alacsonyabbak voltak, azaz az intelligencia és a problémamegoldó képesség összefüggéseinek elemzése során is tapasztalt jelenséggel talákoztunk, mindhárom kapcsolat erősségében meghatározó szerepet játszott a harmadik konstruktum fejlettségi szintje ($r_{IG_SPK}=0,26$; $r_{IG_DPK}=0,33$; $r_{SPK_DPK}=0,26$, mindegyik $p<0,01$).

Az eredmények értelmében az induktív gondolkodás, a statikus és a dinamikus problémahelyzetekben alkalmazott problémamegoldó képesség egymással összefüggő, de nem azonos konstruktumok. Mindhárom vizsgált képesség olyan kognitív képességeket is működtet, amelyek nélkülözhetetlenek a szabályfelismerés, szabálygenerálás és szabályalkalmazás területén. Bármely vizsgált konstruktum fejlesztése másodlagos hatásként pozitív transzferhatást gyakorol a többi vizsgált konstruktum fejlődésére.

Az induktív gondolkodás fejlettségi szintje, mint a problémamegoldás folyamatában az információfeldolgozás során működtetett alapvető gondolkodási képesség, erősebb hatással volt a különböző környezetben történő problémamegoldó sikerességre, mint a statikus problémahelyzetekben nyújtott problémamegoldó sikeresség a dinamikusra vagy fordítva. Ez alátámasztotta azon korábbi kutatási eredményeket (pl.: *Klauer*, 1996; *Hamers, De Koning és Sijtsma*, 2000), miszerint az induktív gondolkodás fejlettségi szintje alapvetően befolyásolja a tudás (idegennyelvtudást is; *I. Csapó és Nikolov*, 2009) elsajátításának és alkalmazásának sikerességét, ezért alapvető szerepet játszik mind a területspecifikus, mind a területfüggetlen, mind a statikus, mind a dinamikus problémák megoldása során.

5.1.2. Az Országos kompetenciamérés, az iskolai sikeresség és a demográfiai háttérváltozók előrejelző ereje a diákok problémamegoldó képességszintjére

A 2014-es longitudinális vizsgálat és a 2013-as OKM mérés adatbázisait egyesítve felépítettük az OKM által mért területek, az iskolai sikeresség és a demográfiai háttérváltozók előrejelző hatását magába foglaló modellt a diákok DPK teszten nyújtott teljesítményére. A végső SEM modell illeszkedése ($CFI=0,99$, $TLI=0,99$, $RMSEA=0,04$) jó volt. Az útelemzések alátámasztották, hogy a diákok induktív gondolkodásának fejlettségi szintje jelentős prediktív erővel bír problémamegoldó gondolkodásuk fejlettségi szintjére, valamint felhívták a figyelmet az iskolai sikeresség és a tudás alkalmazhatósága, a problémamegoldó gondolkodás fejlettségi szintje közötti szakadéokra. Az induktív gondolkodás mellett a gondolkodtató feladatokat és nem a tanultak szó szerinti visszaadását kérő Országos kompetenciamérés matematika tesztjén elért eredmények előrejelző ereje is előtérbe került az elemzések során, amit a diákok neme és angol nyelvtudása követett (utóbbi jelentős mértékben összefügg az induktív gondolkodás fejlettségi szintjével). Az iskolai jegyek és a diszciplináris tudásra építő tesztek ugyanakkor egyáltalán nem jelezték előre a diákok problémamegoldó képességfejlettségi szintjét. Az a kutatási eredmény, hogy a diákok tanulmányi átlagából, iskolai sikerességéből egyáltalán nem

következtethetünk problémamegoldó képességük – beleértve a tudásuk alkalmazhatóságának képességét – fejlettségi szintjére, elgondolkodtató és komoly problémákat vet fel iskolarendszerünk hatékonyságát és működését illetően.

A hét év távlatában elvégzett elemzések kiemelték a DIFER teszt és az első évfolyamos diákok számolási készségének előrejelző hatását a hetedik évfolyamos problémamegoldó képességteszten nyújtott teljesítményükre. E képességterületek első évfolyamos teljesítményei alapján 20-25%-ban magyarázható a diákok hetedik évfolyamos problémamegoldás képességteszten nyújtott teljesítménye, ami alátámasztja az erős szelekció következtében tapasztalt jelenségeket (*Csapó, Molnár és Kinyó, 2009*). Ugyanakkor a hatodik évfolyamos teszteredményeket integrálva a modellbe, az első évfolyamon mért területek előrejelző ereje eltűnt és egyértelműen az OKM matematika tesztjén és az induktív gondolkodás fejlettségét mérő teszten elért eredmények előrejelző ereje került előtérbe, amit a diákok neme és angol nyelvtudása követett. Utóbbi erős összefüggést mutatott a diákok induktív gondolkodásának fejlettségi szintjével ($r=0,60$, $p<0,01$). A harmadik generációs teszttel történt adatfelvétel jelentős eredménye, hogy a mérés során sikerült teljes mértékben kiküszöbölni a diákok olvasási képességének fejlettségi szintjének teljesítménybefolyásoló hatását, ami a hagyományos tesztkörnyezetben végzett képességszintvizsgálatokban jelentős hatással bírt, jelentős mértékben befolyásolta a diákok más, komplex képességterületen nyújtott teljesítményét.

5.2. A dinamikus problémamegoldó környezetben alkalmazott felfedező stratégiák hatékonysága és azok változása: logfájlelemzések

A disszertáció korábbi részeiben tapasztalhattuk, hogy a technológiaalapú mérés-értékelés új lehetőségeket és kihívásokat is hozott a pedagógiai empirikus kutatásokba. Alapvetően három fő tényező támogatja és motiválja a technológiaalapú tesztelésre való átállást: a hagyományos, papíralapon is mért területek kapcsán tapasztalt megnövekedett mérési pontosság (pl. *Csapó, Molnár és Nagy, 2014*); olyan területek mérési lehetősége, amelyre hagyományos eszközökkel nem kerülhetett sor (pl.: komplex, dinamikus problémamegoldás, l. *Greiff, Wüstenberg és Funke, 2012; Greiff és mtsai, 2013*); végül a közvetlenül megfigyelt adatokon kívül a metaadatok rögzítése, majd a logfájlek elemzésének lehetősége (l. pl.: *Tóth, Rölke, Greiff és Wüstenberg, 2014*).

A pedagógiai adathányászás (*educational data mining*, EDM), a logfájlek elemzése, hozzájárulhat ahhoz, hogy mélyebben és alaposabban megértsük és megmagyarázzuk a vizsgált jelenségek működését. Az EDM segítségével olyan kutatási kérdések megválaszolása is lehetséges, melyekre korábban, a hagyományos technikákkal nem volt lehetőség. A disszertáció e részében a DPK kutatások logfájleire alapozva elemezzük azokat a hatékony és kevésbé hatékony problémamegoldó stratégiákat, melyeket a diákok a DPK teszt megoldása közben alkalmaztak.

A DPK teszt problémái olyan minimálisan komplex rendszerek, ahol a diákoknak azáltal, hogy manipulálhatják a bemeneti változók értékeit, amelyek, oksági kapcsolatban állnak a kimeneti változókkal, fel kell fedezniük a bemeneti és kimeneti változók közötti összefüggéseket (*Wüstenberg és mtsai, 2014*). Ezek a kapcsolatok a legtöbb esetben nem nyilvánvalóak, ugyanakkor a rendszer szisztematikus kontrollálásával, megfelelő felfedező és problémamegoldó stratégia alkalmazásával detektálhatóak, feltérképezhetőek. A diákok által

alkalmazott stratégiák, lépéskombinációk sokfélék, miután minden egyes bemeneti változónak öt értéke lehetséges és nem korlátoztuk a lépések számát (csak az időt, ami alatt a lépéskombinációkat alkalmazhatták). Ezért első lépésként szükség volt a logolt adatok átalakítására, kezelhetővé, elemezhetővé, értelmezhetővé tételére, rendszerbe foglalására.

Az egyes lépések kódolása mellett az összes, elméletileg helyes stratégiára alapozva felépítettünk egy olyan matematikai modellt, aminek segítségével elemezhetővé, áttekinthetővé váltak az alkalmazott lépéskombinációk, stratégiák. A szakirodalomból ismert úgynevezett változók kontrollja (*control of variables*, CVS; Greiff, Wüstenberg, Csapó, Demetriou, Hautamäki, Graesser és Martin, 2014) vagy egyszerre egy dolog változtatása (*very vary-one-thing-at-a-time*, VOTAT; Tschirgi, 1980; Funke, 2014) stratégián belül e modell segítségével számos további stratégia definiálását valósítottuk meg. Időelemzések segítségével teszteltük az alkalmazott stratégiák tudatosságát, elemeztük a diákok által alkalmazott stratégiák teszten belüli változását, látens profil elemzések segítségével definiáltuk a különböző típusú problémamegoldókat, valamint összevetettük a különböző felfedező stratégiák kapcsolatát a tudás elsajátítás képességét mérő itemeken nyújtott teljesítményekkel, valamint az általános problémamegoldó képességgel. Végül elemeztük, hogy a probléma megoldásával töltött idő és kattintások száma hogyan függ össze a teljesítményekkel.

Az elemzéseket a két 2014-es és a 2015-ös egyetemisták körében történt adatfelvétel eredményeivel (n=1259) és logadataival végeztük. 2014-ben egyrészt 3–12. évfolyamos tanulók részvételével történt kutatás, melynek 6. (n=677), 7. (n=607) és 8. (n=942) évfolyamos részmintáját vontuk be az elemzésekbe, másrészt a longitudinális adatfelvétel 7. évfolyamos (n=2237) adataira alapoztuk az e részben ismertetett eredményeket.

A kutatás és mintaválasztás oka egyrészt a korábbi elemzések alapján a fejlődés szempontjából szenzitív és gyorsan változó 6-8. évfolyamra eső korszak választása, a hasonló korosztályban történt 2013-as adatfelvételhez képest nagyobb mintaelemszám, továbbá lehetőség az évfolyamok közötti összehasonlításra volt. Miután a 6-8. évfolyamos részmintán – a kutatás felépítése következtében – ugyanazon DPK teszt került kiközvetítésre, ezért az adott kutatáson belül nemcsak a problémák, hanem a problémák teszten belüli pozíciója, elhelyezkedése is azonos volt.

A problémamegoldók látens osztályaink meghatározása érdekében látens profil (Collins és Lanza, 2010) elemzést végeztünk. A látens profil elemzés a problémamegoldók olyan típusait keresi, akik az alkalmazott problémamegoldó stratégiák tekintetében hasonló mintázatot mutatnak. A látens profil elemzés kategorikus látens változóit a logfájlok alapján generált 0,1,2 kategóriákat tartalmazó változók adták. A problémák kapcsán minden egyes diákhhoz hozzárendeltünk még egy kategoriális változót. A változó azt jellemezte, hogy az adott diák az adott probléma kapcsán mennyire alkalmazta a teljes mértékben elszigetelt változókezelésre építő VOTAT-stratégiát. Ha egyáltalán nem tette, akkor az új változó értéke 0. Ha részlegesen tette, azaz például nem mind a három bemeneti változó kapcsán, hanem csak kettő vagy egy változóra alkalmazva, akkor 1-el kódoltuk az új változó értékét. Ha a problémában szereplő összes bemeneti változóra alkalmazta, akkor pedig 2-es kódolást kapott az új változó.

A tudás elsajátítás hatékonyságának fázisát mérő részteszt megbízhatósági mutatója jelentős mértékben megnövekedett, ha azt a diákok teszten nyújtott, logfilék segítségével kinyert viselkedése alapján számoltuk. A generált, logfiléken alapuló, az alkalmazott stratégia helyességét jellemző dichotóm változók segítségével számolt megbízhatósági mutató értéke (10 item): $\alpha=0,91$, azaz a diákok felfedező stratégiáinak direkt kódolása megbízhatóbban jellemezte

a diákok tudás elsajátítása terén mutatott képességszintjét, mint a tanultak modellépítés formájában történő leképezése. Évfolyamonkénti bontásban $\alpha_{\text{evf6}}=0,91$, $\alpha_{\text{evf7}}=0,92$, $\alpha_{\text{evf8}}=0,91$. A 30 itemet együtt kezelve, mint a problémamegoldás három fázisának résztesztjeit, a megbízhatósági mutató értéke a teljes mintán $\alpha=0,88$ ($\alpha_{\text{evf6}}=0,88$, $\alpha_{\text{evf7}}=0,89$, $\alpha_{\text{evf8}}=0,90$).

A logfájlelemzések alapján megállapítottuk, hogy a diákok által alkalmazott stratégia hatékonysága nem minden esetben egyezett meg teljesítményszintjükkel. Annak ellenére, hogy a minimálisan komplex rendszerek kiismerése, feltérképezése számos problémamegoldó stratégiával megvalósítható, mégis a szakirodalomban közel egyedülként tárgyalt VOTAT-stratégiákhoz sorolható teljes mértékben elszigetelt változókezelésen alapuló stratégiák alkalmazása bizonyult a leghatékonyabbnak. A többi, elméletileg jó stratégia alkalmazása esetén minden esetben magasabb volt a helytelen értelmezés, a helytelen válaszok aránya, mint a jó megoldásoké. A tudatos stratégiahasználók nagyobb arányban értelmezték helyesen és képezték le jól a rendszerből kinyert információkat, mégsem jelentett egyértelműen helyes megoldást a tudatos stratégiaalkalmazás. Minél bonyolultabb volt a probléma mögött húzódó rendszer, annál inkább elkülönült egymástól a tudatos és kevésbé tudatos problémamegoldók teljesítménye.

A látens profilelemzések alapján az alkalmazott problémafeltérképező stratégia függvényében hat különböző tulajdonságokkal jellemezhető csoportot különíthetünk el egymástól: a legegyszerűbb rendszereket sem kezelő problémamegoldókat, ide sorolhatóak a 6-8. évfolyamos diákok harmada. A legegyszerűbb rendszereket alacsony fokon átlátó és kisebb fokú tanulást mutató problémamegoldókat (16%), az egyszerű problémákat kisebb sikerekkel átlátókat, de a bonyolultabb rendszerekkel nem boldoguló problémamegoldókat (21%). Azon diákok, akik e három csoport valamelyikébe sorolhatóak, alapvetően alacsony szintű explorációs képességekkel rendelkeztek (a diákok 70%-a) és tanulási képességük is kis határfokú volt. A diákok 8-10%-a magas teljesítményt mutatott az egyszerű felépítésű problémák megoldásában, ugyanakkor a bonyolultabb rendszereknél már egyértelműen az alulteljesítő problémamegoldók csoportjához tartoztak. Tőlük egyértelműen elkülöníthetőek, bár átlagos teljesítményük tekintetében hasonlóak, a nagyon hatékony tanulási képességekkel rendelkező diákok (6-8%), akik gyorsan megtanulták a rendszer használatát, aminek következtében a teszt végére már a legbonyolultabb problémákat is a szakértő problémamegoldók szintjén oldották meg. A 6-8. évfolyamosok mindössze 17%-át sorolhattuk a szakértő problémamegoldók közé, akik mind az egyszerű, mind a bonyolult problémákat hatékonyan térképezték fel. A tág életkori intervallumot átfogó, ezért több iskolaváltás hatását is vizsgáló elemzések megerősítették azt a feltételezésünket, hogy bár e fontos képesség explicit fejlesztése nem valósul meg a közoktatás folyamán, mégis a szelekció során nagyon lényeges szerepet játszik, sőt a szelekció mértéke erősen magyarázható a diákok problémamegoldó és explorációs képességének fejlettségi szintjével. Például az egyetemista korosztályban, az érettségit követő szelekció hatására a szakértő problémamegoldók csoportjába sorolható diákok aránya meghaladja a 60%-ot, miközben arányuk a teljes mintában – a fejlődési görbék alapján extrapolálva az eredményeket – 20% körüli. A látens profilelemzések alátámasztották azt az elméletet, miszerint a képességek fejlődése mind mennyiségi, mind minőségi változáson alapul (Csapó, 2003). A technológiaalapú tesztelés és az új elemzési eljárások segítségével mindkettő számszerűsíthető és objektív eszközökkel jellemezhető.

6. ÖSSZEFOGLALÁS ÉS JÖVŐBELI FELADATOK

A disszertáció öt fejezetében összefoglaltuk a technológia oktatási integrációjának lehetőségeit és fontosságát; a pedagógiai mérés-értékelés technikáinak fejlődési tendenciáit, külön hangsúlyt fektetve a technológiaalapú tesztelésre való átállás előnyeire, hátrányaira, kihívásaira, elterjesztésének feltételrendszerére és a vonatkozó nemzetközi és hazai tendenciákra. Külön fejezetben tárgyaltuk a számítógép-alapú tesztelés bevezetésének nemzetközi szinten legkevésbé kutatott és leginkább kérdéses életkori intervallumban, kisiskolás korban történő alkalmazhatóságára vonatkozó kutatási eredményeket. Végül a problémamegoldó képesség mérési és értékelési lehetőségeinek változásán keresztül szemléltettük a számítógép-alapú tesztelésben rejlő mai lehetőségeket, miközben tág életkori intervallumban ismertettük a diákok egy 21. században kulcsfontosságú gondolkodási képességének, a problémamegoldó képességének fejlődését, a mennyiségi és minőségi változás fő tendenciáit és lehetséges okait. Az elemzések során nemcsak a klasszikus, hanem a valószínűségi tesztelmélet és a neveléstudományi kutatásokban jelenleg leginnovatívabb, új lehetőségeket adó strukturális egyenletek eszközrendszerét is alkalmaztuk. Az elemzett adatbázisok csak egy része vonatkozott a diákok konkrét válaszainak elemzésére, a dolgozatban nagy hangsúlyt fektettünk, a nemzetközi szinten is most megjelenő és ezért kiemelt fontossággal bíró logfájl-elemzésekre és a legmodernebb elemzési módszerek logadatok értelmezésében való felhasználására. A számítógép-alapú adatfelvétel és a modern elemzési eljárások ötvözése lehetőséget teremtett a vizsgált jelenségek alaposabb megértéséhez.

Összességében megállapítható, hogy a technológiaalapú tesztelés és az új módszertani repertoár integrálása olyan új lehetőségeket teremtett a mérés-értékelés terén, amire hagyományos, papíralapú és szemtől szembeni technikák, illetve a klasszikus tesztelmélet alkalmazásával nem volt mód. Ezen eszközök alkalmazásával a korábbinál pontosabbá, objektívabbá, a diákok számára motiválóbbá tehető a tesztelés folyamata. A logfájl-elemzések segítségével pontosabban rekonstruálható, mit tett, hogyan gondolkodott a diák a tesztelés során, ami új lehetőségeket teremtett a vizsgált jelenségek alaposabb megértése kapcsán. Az értékelés azonnali visszacsatolásának lehetőségét kihasználva a korábbi szummatív dominanciájú megközelítés mellett jelentős hangsúlyt kaphat a diagnosztika, az egyénre szabott, hatékony, tanulást segítő tesztelés. Ezzel megvalósulna a mérés-értékelés átdefiniálása és tanulást segítő funkciójának kihasználása. Alkalmazása a 21. században – egy fejlett oktatási rendszerrel rendelkező ország esetén – elkerülhetetlen.

További kutatási irány a középiskolás diákok problémamegoldó képességszintje fejlődésének pontosabb mennyiségi és minőségi jellemzése; a tanulási módszerek és stílusok, valamint a problémák mögött húzódó rendszer megismerése során használt explorációs stratégiák kapcsolatának meghatározása; a dinamikus problémamegoldó képesség komponensképességeinek teljesebb feltérképezése, melyeket nemzetközi kontextusban is elhelyezünk. Az eredmények alkalmazhatóak lesznek a magyar diákok problémamegoldó képességének fejlesztése kapcsán. Egy további kutatási irány, a kollaboratív problémamegoldó képesség mérésére alkalmas valid, jó tesztelmélet mutatókkal bíró, azonnali visszacsatolást biztosító teszt kidolgozása, amely lehetővé teszi a kollaboratív problémamegoldó képesség jellemzőinek, illetve fejlődésének feltérképezését. E téren már elkészültek az eDia online mérés-értékelési rendszer azon funkciói, ami lehetővé teszi az ember-ember (*human-to-human*) kooperációt, sőt az első pilot vizsgálatok is lezajlottak már. A kutatás nemzetközi szinten is

egyedülálló, ugyanis a kooperatív problémamegoldó képesség vizsgálatát általában ember-számítógép (*human-to-agent*) együttműködésben valósítják meg. További kutatások, valamint új módszertan, elemzési repertoár kidolgozása szükséges a logfájlelemzésekben lévő lehetőségek magasabb szintű kihasználásához.

7. IRODALOMJEGYZÉK

- Baker, E. L. és Mayer, R. E. (1999): Computer-based assessment of problem solving. *Computer sin Human Behavior*, **15**. 269–282.
- Beller, M. (2013): Technologies in large-scale assessments: New directions, challenges, and opportunities. In: von Davier, M., Gonzalez, E., Kirsch, I. és Yamamoto, K. (szerk.): *The role of international large-scale assessments: Perspectives from technology, economy, and educational research*. Springer, Netherlands. 25–45.
- Bennett, R. E., Goodman, M., Hessinger, J., Kahn, H., Liggett, J., Marshall, F. és Zack, J. (1999): Using multimedia in large-scale computer-based testing programs. *Computers in Human Behavior*, **15**. 3–4. sz. 283–294.
- Benson, N., Hulac, D. M. és Kranzler, J. H. (2010): Independent examination of the Wechsler Adult Intelligence Scale – Fourth Edition: What does the WAIS-IV measure? *Psychological Assessment*, **22**. 1. sz. 121–130.
- Binkley, M., Erstad, E., Herman, J., Raizen, S., Ripley, M., Miller-Ricci, M. és Rumble, M. (2012): Defining 21st century skills. In: Griffin, P., McGaw, B. és Care, E. (szerk.): *Assessment and Teaching of 21st Century Skills*. Springer, Dordrecht. 17–66.
- 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.
- Bond, T. és Fox, C. M. (2001): *Applying the Rasch model. Fundamental measurement in the human sciences*. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.
- Bond, T. és Fox, C. M. (2015): *Applying the Rasch model. Fundamental measurement in the human sciences*. 3rd edition. Lawrence Erlbaum Associates, Publishers, Hillsdale, New Jersey.
- Breiter, A., Groß, L. M. és Stauke, E. (2013): Computer-based large-scale assessments in Germany. In: Passey, D., Breiter, A. és Visscher, A. (szerk.): *Next generation of information technology in educational management*. Springer Berlin, Heidelberg. 41–54.
- Bridgeman, B. (2010): Experiences from large-scale computer-based testing in the USA. In: Scheuermann, F. és Bjornsson, J. (szerk.): *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing*. European Communities, Brussels. 39–44.
- Byrne, B. M. és Stewart, S. M. (2006): The MACS approach to testing for multigroup invariance of a second-order structure: A walk through the process. *Structural Equation Modeling*, **13**. 2. sz. 287–321.
- Carson, K., Gillon, G. és Boustead, T. (2011): Computer-administrated versus paper-based assessment of school-entry phonological awareness ability. *Asia Pacific Journal of Speech, Language and Hearing*, **4**. 2. sz. 85–101.

- Christakoudis, C., Androulakis, G. S. és Zagouras, C. (2011): Prepare items for large scale computer based assessment: Case study for teachers' certification on basic computer skills. *Procedia-Social and Behavioral Sciences*, **29**. 1189–1198.
- Collins, L. M. és Lanza, S. T. (2010): *Latent class and latent transition analysis: With applications in the social, behavioral, and health sciences*. Wiley: New York.
- Csapó, B. és Nikolov, M. (2009): The cognitive contribution to the development of proficiency in a foreign language. *Learning and Individual Differences*, **19**. 2. sz. 209–218.
- Csapó Benő (2003): *A képességek fejlődése és iskolai fejlesztése*. Akadémiai Kiadó, Budapest.
- Csapó, B., Ainley, J., Bennett, R. E., Latour, T. és Law, N. (2012): Technological issues for computer-based assessment. In: Griffin, P., McGaw, B. és Care, E. (szerk.): *Assessment and teaching of 21st Century skills*. Springer, New York. 143–230.
- Csapó Benő, Lőrincz András és Molnár Gyöngyvér (2012): Innovative assessment technologies in educational games designed for young students. In: Ifenthaler, D., Eseryel, D. és Ge, X.: *Assessment in game-based learning: Foundations, innovations, and perspectives*. Springer, New York. 235–254.
- Csapó Benő, Molnár Gyöngyvér és Kinyó László (2009): A magyar oktatási rendszer szelektivitása a nemzetközi összehasonlító vizsgálatok eredményeinek tükrében. *Iskolakultúra*, **19**. 3–4. sz. 3–13.
- Csapó Benő, Molnár Gyöngyvér és Nagy József (2015): A DIFER tesztek online változatával végzett mérések tapasztalatai. In: Csapó Benő és Zsolnai Anikó (szerk.): *Online diagnosztikus mérések az iskola kezdő szakaszában*. Oktatókutató és Fejlesztő Intézet, Budapest. 163–182.
- Csapó, B., Molnár, Gy. és Nagy, J. (2014): Computer-based assessment of school-readiness and reasoning skills. *Journal of educational psychology*, **106**. 2. sz. 639–650.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papíralapú teszektől a számítógépes adaptív tesztelésig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, **18**. 3–4. sz. 3–16.
- Csapó Benő, Molnár Gyöngyvér és R. 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.
- Csíkó Csaba, Molnár Gyöngyvér és Csapó Benő (2015): A matematika online diagnosztikus mérések tartalmi kereteinek elméleti alapjai. In: Csapó Benő, Csíkó Csaba és Molnár Gyöngyvér (szerk.): *A matematikai tudás online diagnosztikus értékelésének tartalmi keretei*. Oktatókutató és Fejlesztő Intézet, Budapest. 13–26.
- Dikli, S. (2006): An overview of automated scoring of essays. *The Journal of Technology, Learning, and Assessment*, **5**. 1. sz. <http://ejournals.bc.edu/ojs/index.php/jtla/article/view/1640/1489>.
- 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.

- Frey, A. (2007): Adaptives Testen. In: Moosbrugger, H. és Kelava, A. (szerk.): *Testtheorie und Testkonstruktion*. Springer, Berlin, Heidelberg. 261–278.
- Funke, J. (2014). Analysis of minimal complex systems and complex problem solving require different forms of causal cognition. *Frontiers in psychology*, **5**. 739.
- Graesser, A. C. (2012): Foreword. In: Mayrath, M. C., Clarke-Midura, J., Robinson, D. H. és Schraw, G. (szerk.): *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Information Age Publishing, Charlotte, NC. vii-ix.
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J. és Csapó, B. (2013). Complex problem solving in educational contexts – Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, **105**. 2. sz. 364–379.
- Greiff, S., Wüstenberg, S. és Funke, J. (2012): Dynamic problem solving: A new assessment perspective. *Applied Psychological Measurement*, **36**. 3. sz. 189–213.
- Greiff, S., Wüstenberg, S., Csapó, B., Demetriou, A., Hautamäki, H., Graesser, A. C. és Martin, R. (2014): Domain-general problem solving skills and education in the 21st century. *Educational Research Review*, **13**. 74–83.
- Hamers, J. H. M., De Koning, E. és Sijtsma, K. (2000): Inductive reasoning in the third grade: Intervention promises and constraints. *Contemporary Educational Psychology*, **23**. 132–148.
- ITU (2015): *Measuring the Information Society Report 2015*. International Telecommunication Union, Geneva, Switzerland. <http://www.itu.int/en/ITU-D/Statistics/Documents/publications/misr2015/MISR2015-w5.pdf>
- Jurecka, A. és Hartig, J. (2007). Computer- und Netzbasiertes Assessment. In: Hartig, J. és Klieme, E. (szerk.): *Möglichkeiten und Voraussetzungen technologiebasierter Kompetenzdiagnostik*. Bundesministerium für Bildung und Forschung, Bonn, Berlin. 37–48
- KCC (2014): Annual report 2013. Korea Communications Commission. Korea. <http://eng.kcc.go.kr/download.do?fileSeq=42390>
- Klauer, K. J. (1996): Begünstigt induktives Denken das Lösen komplexer Probleme? Experimentellen Studien zu Leutners sahel-Problem. *Zeitschrift für Experimentelle Psychologie*, **43**. 1. sz. 361–366.
- Korom Erzsébet, Molnár Gyöngyvér és Csapó Benő (2015): A természettudományi online diagnosztikus mérések tartalmi kereteinek elméleti háttére. In: Csapó Benő, Korom Erzsébet és Molnár Gyöngyvér (szerk.): *A természettudományi tudás online diagnosztikus értékelésének tartalmi keretei*. Oktatókutató és Fejlesztő Intézet, Budapest. 13–30.
- 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.
- Magyar Andrea és Molnár Gyöngyvér (2013): Adaptív és rögzített formátumú tesztek alkalmazásának összehasonlító hatékonyságvizsgálata. *Magyar Pedagógia*, 3. sz. 181–193.
- Martin, R. (2010): 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.* European Communities, Brussels. 172–177.
- Mayrath, M. C., Clarke-Midura, J. és Robinson, D. H. (2012): Introduction to technology-based assessments for 21st Century skills. In: Mayrath, M. C., Clarke-Midura, J., Robinson, D. H. és Schraw, G. (szerk.): *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research.* Information Age Publishing, Charlotte, NC. 1–12.
- Meijer, R. (2010): Transition to computer-based assessment: Motivations and considerations. In: Scheuermann, F. és Björnsson, J. (szerk.): *The transition to computer-based assessment: New approaches to skills assessment and implications for large-scale testing.* European Communities, Brussels. 104–107.
- Molnár Gyöngyvér és Csapó Benő (2013): Az eDia online diagnosztikus mérési rendszer. XI. Pedagógiai Értékelési Konferencia. Szeged, 2012. április 11-13. 82.
- Molnár, G. és Lőrincz, A. (2012): Innovative assessment technologies: Comparing ‘face-to-face’ and game-based development of thinking skills in classroom settings In: Chen, D. (szerk.): *International proceedings of economics development and research. Management and education innovation. Vol. 37.* IACSIT Press, Singapore. 150–154.
- Molnár Gyöngyvér (2005): Az objektív mérés megvalósításának lehetősége: a Rasch-modell. *Iskolakultúra*, **15**. 3. sz. 71–80.
- Molnár Gyöngyvér (2006): A Rasch-modell alkalmazása a társadalomtudományi kutatásokban. *Iskolakultúra*, **16**. 12. sz. 99–113.
- Molnár Gyöngyvér (2008): A Rasch modell kiterjesztése nem dichotóm adatok elemzésére: a rangskálás és parciális kredit modell. *Iskolakultúra*, **18**. 1. sz. 66–77.
- Molnár Gyöngyvér (2010): Technológiaalapú mérés-értékelés hazai és nemzetközi implementációi. *Iskolakultúra*, **20**. 7–8. sz. 22–34.
- Molnár Gyöngyvér (2013a): Terüleetspecifikus komplex problémamegoldó gondolkodás fejlődése. In: Molnár Gyöngyvér és Korom Erzsébet (szerk.): *Az iskolai sikerességet befolyásoló kognitív és affektív tényezők értékelése.* Nemzedékek Tudása Tankönyvkiadó, Budapest. 161–180.
- Molnár Gyöngyvér (2013b): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában.* Gondolat Kiadó, Budapest.
- Molnár Gyöngyvér (2015a): Az óvoda és iskola feladatai az értelmi képességek fejlesztése terén. In: Kónyáné Tóth Mária és Molnár Csaba (szerk.): *Tartalmi és szervezeti változások a köznevelésben.* Suliszerviz Oktatási és Szakértői Iroda, Suliszerviz Pedagógiai Intézet, Debrecen. 179–190.
- Molnár Gyöngyvér (2015b): A képességmérés dilemmái: a diagnosztikus mérések (eDia) szerepe és helye a magyar közoktatásban. *Génius Műhely Kiadványok*, 2. sz. 16–29.
- Molnár, G., Greiff, S., Wüstenberg, S. és Fischer, A. (2016): Empirical study of computer based assessment of domain-general dynamic problem solving skills. In: Csapó, B., Funke, J. és Schleicher, A. (szerk.): *On the nature of problem solving.* OECD, Paris. 123–143.
- Molnár Gyöngyvér, Papp Zoltán, Makay Géza és Ancsin Gábor (2015): *eDia 2.3 Online mérési platform – feladatfelviteli kézikönyv.* SZTE Oktatáselméleti Kutatócsoport, Szeged.

- Nagy József (2000): *XXI. század és nevelés*. Osiris Kiadó, Budapest.
- Nagy József (2007): *Kompetenciaalapú kritériumorientált pedagógia*. Mozaik Kiadó, Szeged.
- Obama, B. (2016): Giving every student an opportunity to learn through computer science for all. <https://www.whitehouse.gov/photos-and-video/video/2016/01/29/weekly-address-giving-every-student-opportunity-learn-through-comp>
- OECD (2010a): *Are the new millennium learners making the grade? Technology use and educational performance in PISA 2006*. <http://www.oecd.org/edu/cei/45053490.pdf>
- OECD (2012): *Connected minds: Technology and today's learners, educational research and innovation*, OECD Publishing.
- OECD (2013d): *PISA 2012 Results: What makes schools successful? Resources, policies and practices (Volume IV)*. OECD, Paris.
- OECD (2013e): *PISA 2012 Results. What 15-year-olds know and what they can do with what they know (Volume I)*. OECD, Paris.
- OECD (2013a): *PISA 2012 Assessment and analytical framework: mathematics, reading, Science, problem solving and financial literacy*. OECD Paris.
- OECD (2013b): *Draft mathematics framework*. OECD publishing. <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Mathematics%20Framework%20.pdf>.
- OECD (2013c): *Draft reading literacy framework*. OECD publishing. <http://www.oecd.org/pisa/pisaproducts/Draft%20PISA%202015%20Reading%20Framework%20.pdf>
- OECD (2014): *PISA 2012 results: Creative problem solving. Students' skills in tackling real-life problems (Volume V)*. OECD, Paris.
- Pachler, N., Daly, C., Mor, Y. és Mellar, H. (2010): Formative e-assessment: Practitioner cases. *Computers and Education*, **54**. 3. sz. 715–721.
- Peak, P. (2005): Recent trends in comparability studies. Pearson educational measurement. http://www.pearsonassessments.com/NR/rdonlyres/5FC04F5A-E79D-45FE-8484-07AACAE2DA75/0/TrendsCompStudies_rr0505.pdf.
- Quellmalz, E. S., Timms, M. J., Buckley, B. C., Davenport, J., Loveland, M. és Silberglitt, M. D. (2012): 21st Century dynamic assessment. In: Mayrath, M. C., Clarke-Midura, J., Robinson, D. H. és Schraw, G. (szerk.): *Technology-based assessments for 21st century skills: Theoretical and practical implications from modern research*. Information Age Publishing, Charlotte, NC. 55–90.
- Ridgway, J. és McCusker, S. (2003): Using computers to assess new educational goals. *Assessment in Education*, **10**. 3. sz. 309–328.
- RTTT (2009): Race to the Top Program Executive Summary. <https://www2.ed.gov/programs/racetothetop/executive-summary.pdf>.
- 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.
- Sim, G. és Horton, M. (2005): Performance and attitude of children in computer based versus paper based testing. In: Kommers, P. és Richards, G. (szerk.): *Proceedings of world conference on educational multimedia, hypermedia and telecommunications 2005*. VA: AACE, Chesapeake. 3610–3614.

- Snow, C. E. és Van Hemel, S. B. (2008, szerk.): *Early childhood assessment: Why, what, and how*. The National Academies Press, Washington, DC.
- Snow, K. L. (2006): Measuring school readiness: Conceptual and practical considerations. *Early Education and Development*, **17**. 1. sz. 7–41.
- Steklács János, Molnár Gyöngyvér és Csapó Benő (2015): Az olvasás-szövegértés online diagnosztikus mérések tartalmi kereteinek elméleti háttere. In: Csapó Benő, Steklács János és Molnár Gyöngyvér (szerk.): *Az olvasás-szövegértés online diagnosztikus értékelésének tartalmi keretei*. Oktatókutató és Fejlesztő Intézet, Budapest. 13–28.
- Strain-Seymour, E., Way, W. D. és Dolan, R. P. (2009): *Strategies and processes for developing innovative items in large-scale assessments*. Pearson Education, Iowa City, IA.
- 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*. International Educational Data Mining Society. 225–228
- Trilling B. és Fadel. C. (2009): *21st Century skills. Learning for life in our time*. Jossey-Bass A Wiley Imprint. San Francisco.
- Tschirgi, J. E. (1980): Sensible reasoning: A hypothesis about hypotheses. *Child Development*, **51**. 1–10.
- UNESCO (2014): Information and communication technology (ICT) in education in Asia. A comparative analysis of ICT integration and e-readiness in schools across Asia. UNESCO, <http://www.uis.unesco.org/Communication/Documents/ICT-asia-en.pdf>.
- Valenti, S., Neri, F. és Cucchiarelli, A. (2003): An overview of current research on automated essay grading. *Journal of Information Technology Education: Research*, **2**. 1. sz. 319–330.
- Wüstenberg, S., Greiff, S., Molnár, G. és Funke, J. (2014): Determinants of cross-national gender differences in complex problem solving competency. *Learning and Individual Differences*, **29**. 18–29.

A DISSZERTÁCIÓBAN KÖZVETLENÜL FELHASZNÁLT SAJÁT KÖZLEMÉNYEK JEGYZÉKE

- Csapó Benő, Molnár Gyöngyvér és Nagy József (2015): A DIFER tesztek online változatával végzett mérések tapasztalatai. In: Csapó Benő és Zsolnai Anikó (szerk.): *Online diagnosztikus mérések az iskola kezdő szakaszában*. Oktatókutató és Fejlesztő Intézet, Budapest. 163–182.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papíralapú teszektől a számítógépes adaptív tesztelésig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, **18**. 3-4. sz. 3–16.
- Csapó, B., Molnár, Gy. és Nagy, J. (2014): Computer-based assessment of school readiness and early reasoning. *Journal of Educational Psychology*, **106**. 2. sz. 639–650.

- Molnár Gyöngyvér (2004): Hátrányos helyzetű diákok problémamegoldó gondolkodásának fejlettsége. *Magyar Pedagógia*, **104**. 3. sz. 319–337.
- Molnár Gyöngyvér (2005): Az objektív mérés megvalósításának lehetősége: a Rasch-modell. *Iskolakultúra*, **15**. 3. sz. 71–80.
- Molnár Gyöngyvér (2006): A tudáskonceptió változása és annak megjelenése a PISA 2003 vizsgálat komplex problémamegoldás moduljában. *Új Pedagógiai Szemle*, **56**. 1. sz. 75–86.
- Molnár Gyöngyvér (2006): A Rasch-modell alkalmazása a társadalomtudományi kutatásokban. *Iskolakultúra*, **16**. 12. sz. 99–113.
- Molnár Gyöngyvér (2006): *Tudástranszfer és komplex problémamegoldás*. Műszaki Kiadó. Budapest.
- Molnár Gyöngyvér (2007): Hátrányos helyzetű diákok problémamegoldó gondolkodásának longitudinális követése. *Magyar Pedagógia*, **107**. 4. sz. 277–293.
- 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 (2010): Technológiaalapú mérés-értékelés hazai és nemzetközi implementációi. *Iskolakultúra*, **20**. 7–8. sz. 22–34.
- Molnár Gyöngyvér (2011): Az információs-kommunikációs technológiák hatása a tanulásra és oktatásra. *Magyar Tudomány*, **172**. 9. sz. 1038–1047.
- Molnár Gyöngyvér (2011): Számítógépes játék-alapú képességfejlesztés: egy pilot vizsgálat eredményei. *Iskolakultúra*, **21**. 6–7. sz. 3–11.
- Molnár Gyöngyvér (2012): A problémamegoldó gondolkodás fejlődése: az intelligencia és szocioökonómiai háttér befolyásoló hatása 3-11. évfolyamon. *Magyar Pedagógia*, **112**. 1. sz. 41–58.
- Molnár Gyöngyvér (2013): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában*. Gondolat Kiadó, Budapest.
- Molnár Gyöngyvér (2013): Mindennapi helyzetekben alkalmazott problémamegoldó stratégiák változása. *Iskolakultúra*, **23**. 7–8. sz. 31–43.
- Molnár Gyöngyvér (2013): Terület-specifikus komplex problémamegoldó gondolkodás fejlődése. In: Molnár Gyöngyvér és Korom Erzsébet (szerk.): *Az iskolai sikerességet befolyásoló kognitív és affektív tényezők értékelése*. Nemzedékek Tudása Tankönyvkiadó, Budapest. 161–180.
- Molnár Gyöngyvér (2015): Az óvoda és iskola feladatai az értelmi képességek fejlesztése terén. In: Kónyáné Tóth Mária és Molnár Csaba (szerk.): *Tartalmi és szervezeti változások a köznevelésben*. Suliszerviz Oktatási és Szakértői Iroda, Suliszerviz Pedagógiai Intézet, Debrecen. 179–190.
- Molnár Gyöngyvér és Csapó Benő (2003): A képességek fejlődésének logisztikus modellje. *Iskolakultúra*, **13**. 2. sz. 57–69.
- Molnár Gyöngyvér és Csapó Benő (2013): Az eDia online diagnosztikus mérési rendszer. XI. Pedagógiai Értékelési Konferencia. Szeged, 2012. április 11-13. 82.
- Molnár Gyöngyvér és Magyar Andrea (2015): A számítógép alapú tesztelés elfogadottsága pedagógusok és diákok körében. *Magyar Pedagógia*, **115**. 1. sz. 49–66.

- Molnár Gyöngyvér és Pásztor Attila (2015): A számítógép alapú mérések megvalósíthatósága kisiskolás diákok körében: első évfolyamos diákok egér- és billentyűzet-használati képességeinek fejlettségi szintje. *Magyar Pedagógia*, **115**. 3. sz. 237–252.
- Molnár Gyöngyvér és Pásztor-Kovács Anita (2015): A problémamegoldó képesség mérése online tesztkörnyezetben. In: Csapó Benő és Zsolnai Anikó (szerk.): *Online diagnosztikus mérések az iskola kezdő szakaszában*. Oktatókutató és Fejlesztő Intézet, Budapest. 279–300.
- Molnár Gyöngyvér és Pásztor-Kovács Anita (2015): A számítógépes vizsgáztatás infrastrukturális kérdései: az iskolák eszközparkjának helyzete és a változás tendenciái. *Iskolakultúra*, **25**. 4. sz. 49–61.
- Molnár Gyöngyvér, Tongori Ágota és Pluhár Zsuzsa (2015): Az informatikai műveltség online mérése. In: Csapó Benő és Zsolnai Anikó (szerk.): *Online diagnosztikus mérések az iskola kezdő szakaszában*. Oktatókutató és Fejlesztő Intézet, Budapest. 241–260.
- Molnár, Gy., Greiff, S., Csapó, B. (2013): Inductive reasoning, domain specific and complex problem solving: relations and development. *Thinking skills and Creativity*, **9**. 8. sz. 35–45.
- Molnár, Gy., Greiff, S., Wüstenberg, S. és Fischer, A. (2014): Empirical study of computer based assessment of domain-general dynamic problem solving skills. In: Csapó, B. Funke, J. és Schleicher, A. (2016, szerk.): *On the nature of problem solving*. OECD, Paris. 123–143.
- R. 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.