

TARTALOM

TANULMÁNYOK

Golnhofer Erzsébet és Szabolcs Éva: Lázár György és a magyar pedológia – mítosz és valóság	133
Nagy Zsuzsanna: A fogalmazásértékelés megbízhatósága két független bíráló értéktételeinek elemzése alapján	153
Magyar Andrea és Molnár Gyöngyvér: Számítógép alapú adaptív és rögzített formátumú tesztelés összehasonlító hatékonyságvizsgálata	181

LÁZÁR GYÖRGY ÉS A MAGYAR PEDOLÓGIA – MÍTOSZ ÉS VALÓSÁG

Golnhofer Erzsébet és Szabolcs Éva

ELTE Pedagógiai és Pszichológiai Kar Neveléstudományi Intézet

A II. világháború utáni évek hazai pedagógiai világának történeti szempontú kutatása egyre több eredményt mutat fel. Megszűnni látszanak azok a félelmek, amelyek a közelmúlt megítélése kapcsán jó ideig visszatartották a kutatókat az 1945 és 1989 közötti időszak vizsgálatától. Politikai szempontú állásfoglalások helyett történeti alaposágú munkák születnek, melyek reflektálnak a közelmúlt kutatásának speciális problémáira is, például arra, hogy a személyes érintettségből adódó visszaemlékezések, interjúk hogyan és milyen mértékig képezhetik a korszak történeti forrásbázisát (Golnhofer, 2004, 2006; Kéri, 2003; Sáska, 2006; Szabolcs, 2006; Golnhofer és Szabolcs, 2008; Németh, 2009). Különösen a korszak eleje, a totális állam kiépülésének időszaka, a pedagógia és a politika szinte kibogozhatatlan összefonódása került a kutatók érdeklődésének homlokterébe (pl. Sáska, 2006; Szabolcs, 2006). Pedagógia és politika kapcsolatát korábban már 1988-ban elemezte Knausz Imre az Országos Neveléstudományi Intézet működését tárgyalva a szakszerűség és a politikum összefonódásának jegyében (Knausz, 1988).

A szakmai közvéleményt ma is élénken foglalkoztatja azoknak a pedagógiai szakembereknek és/vagy oktatáspolitikusoknak a pályafutása, akik az úgynevezett szocialista korszakban meghatározó szerepet tölthettek be a pedagógiai tudományos közéletben és az oktatáspolitikában. A mindennapi élet, az iskoláztatás, a tudományosság átpolitizáltságának különböző fokozatai voltak jelen az időszakban, és ez, a korszak pedagógiájában aktív szerepet játszott aktorok életpályájában is megmutatkozott valamilyen módon. Tudjuk, hogy a közoktatás-politika mindenkor értékvilága és a professzionizálódott neveléstudomány kapcsolata a modern társadalmakban mindig jelen van. Látható azonban, hogy a totalitárius rendszerekben ez a kapcsolat szélsőségesen szorossá válik, hiszen az uralmon lévők saját kizárólagos értékrendjük követését kívánják az iskoláztatás alakításában résztvevőktől. Az akkori életpályák vizsgálatakor számos kérdés merülhet fel: Hogyan élték meg tudományterületünk szereplői a pedagógia és a politika 1945 után egyre jobban összefonódó viszonyát? Milyen lehetőségei voltak az egyénnek a kialakuló egypártrendszer ideológiai egyneműségében, a hatalmi szempontoknak tudatosan alávetett elitcsere időszakában a tudományos meggyőződésnek, általában a tudományos megközelítés szabadságának a vállalására? Mennyiben volt alárendelve a személyiség a korszak politikájának, a külső kényszereknek? Hogyan befolyásolták a személyiségjellem-

zók, az egyéni habitusok a döntéseket?¹ Ezek a kérdések különböző súllyal jelentek meg az 1945 és 1989 közötti korszak egészében, hiszen politikatörténeti szempontból nem homogén időszakról beszélünk. Magyarország szovjetizálása, a Rákosi-korszak más korlátokat állított, mint a kádári Magyarország egyes periódusai.

A tanulmányban az 1945 utáni koalíciós évek, majd a kommunista hatalomátvétel idején és a Rákosi-korszakban szerepet játszott *Lázár György (Lám Leó)* személyét mutatjuk be, élet- és szakmai pályáját az adott korszakok keretei között értelmezzük, hangsúlyosan elemezve a magyar pedagógia perében betöltött szerepét.

A korszak pedagógiai történéseinek értelmezésében érdemes felidézni, hogy 1945 után rövid ideig egy demokratikus társadalom kibontakozásának reményében folyt a szakmai diskurzus a pedagógia, a közoktatás kérdéseiről. 1947 második felétől az erősödő kommunista párt hatalmi és ideológiai befolyása egyre jobban meghatározta a nevelés kérdéseiről való gondolkodást, és 1948–1949-re kizárólagossá vált a vulgármarxista ideológia alapján művelt pedagógia. Az 1949 nyarára befejeződött kommunista hatalomátvétel nemcsak ideológiai váltást jelentett, hanem magával hozta a párthű káderek pozícióba juttatását és a szakmai intézményrendszer teljes átalakulását is (*Golnhofer, 2004; Romsics, 1999*).

Lázár György neve a politikai diktatúra pedagógiai megszemélyesítőjeként, a pedagógia támadójaként él a szakmai mitológiában. Neve összefonódott az 1950-ben közzétett, militáns hangvételű, a korszak ideológiai frazeológiájához igencsak illeszkedő Mérei-kritikájával (*Lázár, 1950*). Ki volt ő? Hogyan került a pedagógia közelébe, sőt fontos politikai és tudományos pozíciókba? Életútja e szakaszának végigkísérése közelebb vihet annak megválaszolásához is, hogy a vizsgált korszakban milyen jellemző karrier-utak közül „választhattak” a szakma képviselői. Érzékelhető ugyanis, hogy 1945 után a különböző generációkhoz tartozó pedagógiai szakemberek más-más módon váltak aktív szereplőivé vagy elszenvedőivé a gyorsan változó politikai, szakmapolitikai és tudományos életnek. Ez attól függött, mi módon reagáltak a korszak által felvetett politikai, ideológiai, erkölcsi, illetve az ezek által erőteljesen meghatározott szakmai problémákra, másképpen szólva hogyan látta, hogyan kezelte ezeket a reakciókat a hatalom (*Petriné és Szabolcs, 1998; Dessewffy, 1999; Golnhofer, 2004; Németh, 2009*).

Lázár neve manapság szinte kizárólag *Mérei Ferenc* és az általa képviselt pszichológiai és pedagógiai megközelítés kritikájával kapcsolatban merül fel, és az eddigi pedagógiatörténeti kutatások is erre a problematikára koncentrálnak, nem tárgyalják a kritikát író személyét, életútjából adódó szerepét e történetben. Az 1997-ben kiadott Pedagógiai Lexikon is szükséztlenül nyilatkozik róla, akárcsak a Magyar Életrajzi Lexikon. Ezért tartjuk fontosnak e szinte ismeretlen életút feltárását és személyes mozzanatainak értelmezését.

¹ *Dessewffy* (1999) hasonló kérdésekkel foglalkozik a tudásszociológia szemszögéből.

Szemponatok az életút megértéséhez

Lázár György azokhoz a pedagógiai közszereplőkhöz tartozott, akik 1945 után szinte előzmények nélkül, nagyon fiatalon jelentek meg a közoktatás-politikában és a neveléstudomány művelői között. A róla fellelhető családi, baráti, tanítványi visszaemlékezések, életrajzok mind kiemelik, hogy az 1921-ben született Lázár tehetséges fiatalember volt (Kéri, 2003). Szakmai és társadalomalakító tettvégya, becsvágya, tenni akarása az 1945 utáni, politikailag radikális változásokat hozó években jól nyomon követhető. Munkánkban életútját két kritikus eseményhez kapcsolódóan követjük. Az 1949 végén készített önéletrajza és a feltárt levéltári források mellett a felesége emlékét megidéző kötet (Kéri, 2003) alapján mutatjuk be életének alakulását 1950-ig, a Mérei-kritika megjelenéséig, és kevésbé részletesen, de kitekintünk pályájának alakulására 1956 őszéig. A forrásokból megismert narratív, elbeszélő momentumok és „száraz” adatok, tények bemutatásával a politikai, a hatalmi, a tudományos és a személyközi kontextusok érzékeltetésére is törekszünk.

Lázár György tevékenységéről részletesen beszámol a már említett, 1949. november 18-án a Magyar Tudományos Akadémia Magyar Tudományos Tanácsa (MTA MTT) számára készült önéletrajza, ami a kor politikai kívánalmainak megfelelően hangsúlyoz történéseket az életéből:

„Apám Lám Béla gépészmérnök orosz hadifogságban nősült; anyám neve *Lebedinszkij Mária*. Anyám is értelmiségi származású. 1921-ben születtem a Szovjet-Unióban, Nikolszk-Usszurijszk városban. Még ugyanabban az évben Romániába, Kolozsvárra kerültem. Itt nőttem fel. Elemi iskoláimat a Kolozsvári Ág. Hiv. Evangélikus Elemiben végeztem, ahol a tanítás nyelve német volt. Jól megtanultam németül. Középiskoláimat a Kolozsvári Református Fiúfőgimnáziumban végeztem, és ez idő alatt románul és franciául tanultam meg. 1939-ben érettségiztem, 8-as általános eredménnyel a kolozsvári kisebbségi gimnáziumokat vizsgáló kormánybizottság listáján a 3. helyezést nyertem. Ez alatt az idő alatt a családban az ismerősök és barátok körében és részben az iskolában is liberális kommunista nevelésben részesültem. Erősen vallásos voltam. 16-17 éves koromban átestem egy narodnyik-lázon. Közvetlenül politikai jellegű megmozdulásokban, szervezkedésekben azonban ekkor még nem vettem részt. Magatartásom úgyszólván teljesen apolitikus volt./ Érettségi után a Kolozsvári Református Teológiai Fakultásra iratkoztam be. Itteni tanulmányaimat még abban a tanévben félbeszakítottam. Az úgynevezett bécsi döntés után Budapestre jöttem, ahol felvételt nyertem az Eötvös Kollégiumba. A Pázmány Péter Tudományegyetem bölcsészkarára iratkoztam be, görög-latin szakra. Egyetemi éveim alatt szaktárgyaim mellett főleg pszichológiával és kultúrfilozófiával foglalkoztam. Alapvizsgát 1942-ben tettem, mindkét főtárgyból dicséretes, magyar irodalomból jó eredménnyel. 1942 június - júliusában német ösztöndíjjal egyhónapos tanulmányúton voltam Münchenben. 1944-ben lettem szakvizsgámat, kitűnő-dicséretes eredménnyel. Az Eötvös Kollégiumban folytatott tanulmányaim hatására liberális-polgári magatartásom még határozottabb jelleget öltött. Vallásos világnézetemet rövid idő alatt teljesen levettem. Később, részben müncheni élményeim hatására is, apolitikus intellektuel magatartásom fokozatosan polgári baloldali magatartássá alakult át, amely a kollégiumban és másutt lefolytatott politikai vitákban nyilvánult.”²

² MTA Levéltár, II. Filozófiai és Történettudományi Osztály, 199/4.

Az idézett kéziratból jól látható, hogy *Lázár* egy politikai fejlődési ívet próbál bemutatni, melyből egyértelműen kiderül, hogy az önéletrajzíró a háború végére a polgári, vallásos családi indíttatásától baloldali magatartásig jutott el. A szövegrészben egyértelmű az a törekvés, hogy az önéletrajz hangsúlyai az elvárt világnézeti „tökéletesedés” irányába mutassanak. Ugyanakkor a szövegben nehezen értelmezhetőek a különféle világnézetekre, magatartásokra való utalások, így az iskolaihoz, családi neveléséhez kötődő liberális, kommunista nevelés, illetve a liberális polgári magatartás és a müncheni tanulmányút közötti összefüggések.

Az írásból a továbbiakban azt is megtudhatjuk, hogy a háború alatt *Lázár* kapcsolatba került a „*földalatti felszabadító mozgalommal*”, vélelmezése szerint annak szociáldemokrata csoportjával, illetve képviselőjével, *Szigethy Miklóssal*. Részletes, ám kissé zavaros leírását adja annak, hogyan próbált átállni a parasztpárti *Kovács Imre* vezetésével egy kisebb csoportban a szovjet csapatokhoz. Hadifogságba is került, de elmondása szerint átállt a „*demokratikus magyar hadseregbe*”. Harcokban már nem vett részt, és közlése szerint ebben az időszakban, 1945 nyarán kezdte a klasszikus marxista-leninista műveket tanulmányozni, a „*Párt egyéb kiadványait, vezető politikusaik munkáit és a magyar és szovjet tudósok szakmámba vágó műveit*” rendszeresen olvasni.

E közlésekből is látható, hogy *Lázár* öndefiníciójában igyekezett minél korábbra időzíteni kapcsolatát részben a baloldallal, részben a kommunista párttal, részben a marxizmussal, de kommunista párttagsága csak 1945. augusztus 21-ére datálódik. A kommunista párthoz való kötődését a párt különböző szerveinél betöltött tisztségeinek, mozgalmi tevékenységeinek felsorolásával és egyes személyeihez (*Rákosi Mátyás, Lukács György*) való kapcsolatának említésével teszi egyértelművé. A korra jellemző tények tekintjük, hogy az általa említett ellenállók – *Szigethy, Mikó, Kovács* – az önéletrajz írásakor a kommunista párt szemében már árulóknak vagy bukott embereknek tekinthetők. *Kovács Imre* a magyar politikai fejleményekkel egyet nem értve 1947-ben Zürichbe távozott, *Mikó Zoltánt* 1945. augusztus 15-én kivégezte a szovjet Belügyi Népbiztosság (NKVD) (*Sári, 1996; Hovanyec, 1984*). E kapcsolatok említése nemcsak a baloldaliság bizonyítását szolgálta, hanem – feltehetően – a helyzet szülte kényszer őszinteséget és átláthatóságot is jelentette. Részben az utóbbi érzékelhető még az önéletrajzban említett, de általunk nem idézett fegyelmi ügyének³ vagy a hatalom által már nem támogatott NÉKOSZ-beli tevékenységének említésénél. A korszakhoz való igazodást a szöveg stílusában néhány erős kifejezés használata is jelzi, például „*áruló*”, „*kétkulacsos áruló*”, „*súlyosan vétettem*”.

Az önéletrajzból kiolvasható narratívák kulturális és ideológiai érvelés alapján igazolják baloldaliságát. A bemutatott kognitív elemek arra utalnak, hogy olyan egyéneket, szervezeteket, tevékenységeket jelölt meg, akik, amik számára fontos társadalmi környezetet biztosítanak, valamint további karrierlehetőséget hordoznak magukban. Ez a célszerűség adja önéletrajzának koherenciáját, amit azonban gyengítenek a többféleképpen értelmezhető világnézeti utalások és életesemények.

³ MTA Levéltár. II. Filozófiai és Történettudományi Osztály 199. doboz. Nyilatkozat *André Kaminsky* Magyarországra jövetelével kapcsolatban.

A háború utáni tevékenységét és munkahelyeit sokféleség jellemezte mind szakmai, mind politikai vonalon. 1945 szeptemberében letette a tanári vizsgát, görög és latin szakos tanári oklevelet szerzett. Októbertől a Székesfővárosi Könyvtárban (1946-tól Székesfővárosi Szabó Ervin Könyvtárban) dolgozott. Közben német nyelvből lefordította *Lukács György* Goethe és kora című munkáját. A könyvtárban 1946 májusától 1947 szeptemberéig a kommunista pártcsoport titkáráként is működött. Ezzel párhuzamosan 1945. szeptember 6-ától a Magyar Kommunista Párt (MKP) Pártközpontja Országos Propaganda Osztályán volt ideiglenes alkalmazott, ahonnan 1947 márciusában a Miniszterelnökségre került *Rákosi Mátvás* miniszterelnök-helyettes titkárságára miniszteri titkárként.⁴ 1948 decemberében summa cum laude minősítéssel doktori szigorlatot tett filozófia, esztétika és pszichológia szaktárgyakból. *Lázár* 1948-ban és 1949-ben elsősorban politikai munkát végzett, de nevéhez fűződik 1948-ban a NÉKOSZ keretei között az Apáczai Cseri János Kollégium és Gimnázium megszervezése – az Eötvös Kollégium védőszárnya alatt – és igazgatása. 1948 végétől a Magyar Dolgozók Pártja (MDP) Országos Központjában a Kultúrpolitikai Osztályon dolgozott, emellett 1949 elejétől a Pedagógus Szakszervezet Országos Központjának (PSZOK) is munkatársa volt, ugyanakkor 1949. április 1-jétől tudományos státusban tevékenykedő káderként külön juttatásban is részesült.

Az önéletrajz megírásának idején, 1949 novemberében már a tudományok hatalmi, politikai megszállását célzó, elvégző MTA MTT Társadalomtudományok Szakosztály Titkárságának pedagógiai és filozófiai szakmai referense volt, tudományos munkatársa és egyúttal a Pázmány Péter Tudományegyetem megbízott előadója, ahol neveléstant tanított. Jól megfigyelhető, hogy az egyetemi státuszának állandósítására irányuló kommunista párti törekvés egybeesik a régi, „káros” pedagógia képviselőjének, *Prohászka Lajosnak* a nyugdíjazásával. Nyilvánvaló, hogy *Lázárt* az új idők káderének tekintették. Az utóbbit támasztja alá az a vita, ami az MTA MTT Társadalomtudományok Szakosztályának 1949. májusi ülésén bontakozott ki döntően *Lukács György* és *Mérei Ferenc* között arról, hogy *Lázár* kinek az intézetéhez kerüljön, illetve, hogy az elméleti téren égető marxista káderhiányt *Lázár* a filozófia vagy a pedagógia területén segítse megoldani.⁵

1949. december végén a bölcsészettudományi kar dékánja, *Tamás Lajos* a kari tanács ülésén bejelentette, hogy *Ortutay Gyula* Vallás- és Közoktatásügyi miniszter *Prohászka Lajost* nyilvános rendes tanári szolgálata alól felmentette és intézkedett nyugdíjazása ügyében.⁶ Néhány nappal később *Lukács György* javaslata alapján a bölcsészettudományi kar kérte, hogy „az 1950-es költségvetésben szereplő nyilvános rendes tanári állás közül be nem töltött 12 állás terhére dr. *Lázár György* intézeti tanári kinevezését az egyik üres Pedagógiai Tanszék terhére szíveskedjék megerősíteni.”⁷ Ideiglenes kinevezése után véglegesítése 1950. október 26-ára datálódik.⁸ Kiméretlen politikai eszközöket

⁴ MOL XIX-A83-a 215/13; 231/2.

⁵ MTA Levéltár Magyar Tudományos Tanács 23/1.

⁶ ELTE Levéltár BTK Tanácsülési jegyzőkönyvek, 1949. december 12. ELTE Lt8/a 289.

⁷ ELTE Levéltár BTK Dékáni Hivatal iratai 264/1949/50.

⁸ ELTE Levéltár BTK Dékáni Hivatal iratai, 775/1949–50.

alkalmazó eljárások árnyékában került *Lázár* az egyetemre, amit nem tehet semmissé, hogy több visszaemlékező szerint jó kapcsolatot ápolt a diákokkal, többen rajongtak előadásaiért (*Zrinszky*, 2006).

Lázár az MTA MTT titkárságának munkatársaként különböző ügyekkel foglalkozott. Tagja volt a neveléstudomány helyzetét elemző és ötéves tervét kialakító öttagú bizottságnak,⁹ részt vett a bölcsészkarokon leváltandó egyetemi tanárok megítélésével és a főiskolai tanárok kinevezésével foglalkozó üléseken,¹⁰ tudományos és pártbeosztásából adódóan kapcsolatot tartott egyetemekkel, tudományos intézetekkel, szerkesztőkkel; kéziratokat, cikkeket, ösztöndíjakat véleményezett, bírált, mindezzel szakmai egzisztenciákat befolyásolt. Látszólag komoly döntési jogkörrel rendelkezett, például 1950. április 20-án egy neveléstörténeti könyv tervezetre azt írta, hogy nem adható ki.¹¹ Ennek ellenére nem döntéshozóként, hanem inkább – *Huszár Tibor* (1995) kifejezése szerint – beavatottként tekinthetünk rá, hiszen az MTA MTT-n belül a tényleges, a végső döntéseket az MTT Pártkollégiuma hozta meg az MDP központi pártszerveinek jóváhagyása, kezdeményezése alapján.¹²

*Lázár*ra jellemző volt a közügyek iránti érdeklődés, mely részben családi indíttatásból fakadt. Apja, *Lám Béla* Kolozsvár megbecsült polgára, író, költő, mérnök volt; *Áprily Lajos* és *Reményik Sándor* jó barátja (*Kéri*, 2003). Édesapja – az önéletrajz születéséhez képest majdnem 20 évvel később keletkezett levelében – így látta a fiát ért sokféle hatást:

„*Jofkát* Anyám nevelte 'úrifiúnak', *Járosi Andor* papnak, *Reményik Sándor* buzgó hazafinak, én hű férjnek és családapának – és megmaradt minden túlzásra hajlamos orosz forradalmárnak – mint anyja és nagyapja.” (*Kéri*, 2003. 355. o.).

Ez a társadalmi problémák iránti érdeklődés az 1940-es és az 1950-es évek fordulóján a diktatúra keretei között nyilvánosan csak a kommunista ideológiában ölthetett testet, és *Lázár* ebben láthatóan megtalálta magát.

Személyes viszonyrendszere is a kommunista ideológia irányába mozdította el, hiszen kapcsolata a kommunista *Lakatos Imrével* és rajta keresztül – ebben az időben hivatalos pártpolitikusnak tekinthető – *Lukács Györggyel*, valamint a *Lukács-körrel* kimutatható (*Congdom*, 1998; *Mihályi*, 1998). Visszaemlékezések szerint rendszeresen találkoztak a budai Majestic Szállóban kommunista barátaikkal, többek között *Lakatos Imrével*, *Király Józseffel*, *Szigeti Józseffel*, s késő éjszakákba nyúlóan vitatkoztak a kommunista üdvtanról. Feltehető, hogy ezek a viták erőteljesen hozzájárultak *Lázár* politikai identitásának alakulásához és karrierjének építéséhez is. Még az Eötvös Kollégiumban kötött barátságot *Szigeti Józseffel*, a későbbi marxista filozófussal,¹³ aki a háború utáni évekről, párttagságukról így emlékezett:

⁹ MTA Levéltár MTT 1/6.

¹⁰ MTA Levéltár MTT 1. doboz 5/3; 1/10.; 8/2.

¹¹ MTA Levéltár, MTT Titkárság iratai, 8/2.

¹² Az MTA MTT-ről és a döntési mechanizmusokról részletesen l. *Huszár* (1995).

¹³ *Szigeti József* 1957 és 1959 között művelődésügyi miniszterhelyettesként részt vett az 1956-os forradalom utáni tudományos élet forradalmat támogató szereplőinek leváltásában, menesztésében (*Staar*, 2007).

„Tiszta szívvel és teljes meggyőződéssel csatlakoztunk a Magyar Kommunista Párthoz, amelynek politikai tevékenységét lenyűgözőnek tartottuk.” (Szigeti, 2000. 346. o.)

Családja és kortársai tehetségét nem vitatták; személyével foglalkozva gyakran említették kimagasló intellektusát. *Tomasz Jenő* tanár így értékelte *Lázár* munkáját az Eötvös Kollégiumban:

„*Lám Leó* (latin–görög): általános műveltségével az egész csoportból kimagaslik. Franciaul és németül könnyen olvas.”¹⁴ Egy későbbi dokumentumban „finom, mélyenjáró elmének” nevezte.¹⁵ Más tanárok, szeniorok mint *Balázs János* és *Szabó Miklós* is hasonlóan vélekedtek róla, kiemelkedő tehetségnek, tudományos pályára valónak tartották. Azonban mindhárman felfigyeltek pályáját is befolyásoló egyéb tulajdonságaira: „Könnyen szétszóródik a figyelme [...]”¹⁶; „Sokrétű érdeklődése van, talán maga sem tudja, hol kellene kikötni, tartani lehet attól, hogy elmélyedő munkát nem tud végezni. Gyenge idegalkatú”.¹⁷ „[...] némi szertelenségre hajlik. Munkabírása nagyon korlátozott, s félő, hogy meg fog torpanni.”¹⁸

Családtagjai jól látták összetett személyiségét, kimagasló tehetsége és sokoldalúsága mellett indulatait, szenvedélyességét, fellángolásait is tapasztalhatták. Elvált felesége, *Sós Júlia* 1954-ben így írt róla:

„Ő nekem azonnal teljes szenvedélyességgel kezdett udvarolni, én pedig ellentmondásos érzelmekkel viseltem iránta. Tetszett is, riasztott is. Savanarola-szerű magatartása, merevsége, egyszempontúsága, végtelen indulatossága volt az, amit nem szerettem benne. Kulturáltsága, mély érzelmei, tehetsége, lelki gazdagsága volt az, ami vonzott.” (Kéri, 2003. 149. o.)

Hasonlóan látta fiának ellentmondásos személyiségét *Lám Béla*:

„*Jofka* fiam hozzám tartozik erényeivel és hibáival. Fáj ítélkezniem fölötte, de meg kell tennem. Kétségtelenül okos ember, kis túlzással azt is mondhatnánk, hogy zseniális, de ugyanakkor könnyen befolyásolható. Realitásokban nincs meg a józan ítélőképessége, szexuális téren megbízhatatlan. Józan ítélőképességű, konzervatív lénynek, akinek megvan a maga többé kevésbé kialakult világnézete, a hozzá tartozás, tőle függés – súlyos probléma. [...] *Jofka* anya nélkül nevelkedett, és nem érezte anyja hiányát sohasem.” (Kéri, 2003. 145–146. o.)

Azok a kortársai, akikkel együtt dolgozott a pártban és az oktatásügyben, szintén elismerték tehetségét, de náluk a megítélés alapja az volt, hogy *Lázár* miképpen tud működni a párt kívánalmai szerint.

¹⁴ MTA *Tomasz Jenő* 1941. május 29-i jelentése a b. EJC I–IV. latin szakos tagjainak tanulmányi előmeneteléről az 1940/1941. tanév második felében. 53. doboz. 101/8. dosszié 101/8/a.

¹⁵ MTA *Tomasz Jenő* 1944. április 17-én kelt jelentése a b. EJC I–IV. éves latin szakos növendékeinek tanulmányi előmeneteléről az 1943/1944. tanév második felében. 53. doboz. 101/9. dosszié.

¹⁶ MTA *Szabó Miklós* 1942. május 28-i jelentése az 1941/1942. tanévben a görög szakvezetésben végzett munkáról. 53. doboz. 101/8/b.

¹⁷ MTA *Tomasz Jenő* 1943. május 20-i jelentése a b. EJC I–IV. éves latin szakos növendékeinek tanulmányi előmeneteléről az 1942/1943. tanév második felében. 53. doboz. 101/9/a.

¹⁸ MTA *Szabó Miklós* 1944. április 17-én kelt jelentése az 1943/1944-es tanév I–II. tanév félévéről a b. EJC görög csoportjainak munkájáról. 53. doboz. 101/9. dosszié. 101/9/b.

A húszas éveinek második felében járó *Lázár György* 1945 utáni tevékenységi körét nehéz röviden összefoglalni, hiszen láttuk, könyvtárosi, fordítói, tanári, igazgatói tevékenysége mellett politikai (hivatali) és mozgalmi munkát végzett, leginkább kettőt, hármat egyidejűleg. Ezekben az években nem jelent meg pedagógiai témájú tanulmánya, nem folytatott nyilvánosságot is kapó tudományos tevékenységet. Ugyanakkor kommunista párti beágyazottsága következtében 1949 második felében – amint már említettük – a tudományos életet meghatározó legfőbb intézet, az MTA MTT pedagógiai, filozófiai referensévé vált, ahol a kiépült pártállam képviselőjeként működött. Gyakori munkahelyváltásai egyaránt utalhatnak nyughatatlan természetére, tenni akarására, mozgalmi hevületére, emberi, szakmai konfliktusaira és a politikai eseményekben bővelkedő kornak káderhiányára, valamint a káderekkel kapcsolatos mozgósító jellegű akcióira egyaránt. Személyiségének és kapcsolatrendszerének ismeretében állítjuk, hogy fogékony volt a kommunista pártra, párttagokra jellemző, az ideológia, a politika elsődlegessége, meghatározó jellege alapján egybemosódó szakmai és politikai tevékenységre. Önéletrajzából is látható, hogy énképe megegyezett a külső hatalmi elvárásokkal, egyszerre tekintette magát pedagógiai szakembernek és politikai munkát végző, elkötelezett kommunistának, s ez a kép megfelelt a kommunista párt akkori elvárásainak.

Ugyanakkor a személyes dokumentumok azt jelzik, hogy a politikai és szakmai felfutás zaklatott életszakaszt takar. Elvált felesége, *Sós Júlia* 1954-ben így írt volt apósának:

„nagyon sok izgalom volt, ekkor oszlatták fel *Jofka* iskoláját, s őt is sok nyílt és alátomos támadás érte...[1948–51] közben *Jofka* erősen túlhajszolt volt, agyonelgondozott, agyonizgatott. Rengeteg kellemetlensége volt. A kultúrpolitika terén egyesek hajszát indítottak ellene, sokan éppen ezért gyűlöltek, mert művelt, önálló gondolkodású, mélyre látó ember. *Jofka* teljes erejével belevetette magát a kultúrpolitikai harcokba. Lelkesedett, gyűlölte, támadott...Hiába próbáltam meggyőzni arról, hogy az emberek nagy része csak ostoba, önző karrierista, és nem fizetett beépített ellensége a népi demokráciának. Mindaz, ami *Jofkában* jó, és mindaz, ami rossz, hozzájárult ahhoz, hogy rengeteg ellensége legyen. Azt hiszem alig van ember, akinek olyan sok ellensége volna, mint neki. Ki azért gyűlölte, mert művelt, lelkes és tehetséges – ki pedig azért, mert túlzottan szenvedélyes, merev és erőszakos. S a kultúra frontján aztán sok szeméttel lehet találkozni, s szegény *Jofkának* volt is benne része bőven.” (Kéri, 2003. 150–151. o.)

Lázár és Mérei

A pártkáder Lázár

A pedagógia, a neveléstudomány területén döntően az 1948-tól külső hatásra bekövetkező ideológiai váltást, a kommunista párt ideológiájától eltérő nézetek kiszorításának folyamatát több pedagógiatörténeti feldolgozás is részletesen bemutatta (Knausz, 1986; Golnhofer, 2004; Sáska, 2006). *Lázár György* (1950) hírhedt *Mérei* kritikáját e folyamat részeként, betetőzéseként említi az elemzések többsége. A mi kérdésünk tanulmányunkban inkább a következő: Hogyan illeszkedik ez az írás *Lázár* akkori helyzetéhez és tevékenységéhez, szakmai kommunikációjához?

1945-től 1949-ig *Lázár György* és *Mérei Ferenc* sok szempontból hasonló múlttal rendelkezett, például tagjai voltak a kommunista pártnak; számos politikai szervnél dolgoztak, különböző tudományos, szakmai szervezetekben képviselték a kommunista pártot, érvényesítették a párt döntéseit; szakmai tevékenységük mellett intenzív mozgalmi életet éltek; lelkes hívei voltak a népi kollégiumi mozgalomnak és bár különbözőképpen, de mindketten aktívan részt vettek a NÉKOSZ munkájában. A kommunista pártot támogató baloldali értelmiségként mindkettőjüknél összemosódott a politikus és a szakember, a tudós szerepe, ugyanakkor jelentős különbségek is fellelhetők életútjukban. Az életkori és eltérő tapasztalati, tehát generációs különbségből adódóan (*Somlai, 2011*) *Mérei* régebben kapcsolódott a párthoz, szakmai múltja sokkal jelentősebb volt, szakmai és politikai tisztségei révén 1945 után jelentősebb döntések meghozatalában vagy közvetítésében vett részt, mint a fiatal *Lázár*, tehát mind szakmai, mind politikai szempontból *Lázárnál* fajsúlyosabb szereplő volt. 1948 második felétől *Mérei* publikációiban, az Országos Neveléstudományi Intézet (ONI) igazgatójaként vállalta az MDP által meghirdetett ideológiai fordulat támogatását. Főszerepet játszott az MDP-ben Az ideológiai harc a pedagógia területén című téma kidolgozásában.¹⁹ 1948 őszétől jól dokumentálhatók írásában a burzsoá pedagógiának nevezett gyermektanulmányi elemek elleni fellépés, a polgári pszichológiától való eltávolodás (*Mérei, 1948, 1949; Knausz, 1986; Golnhofer, 2004*).

1949-ben még mindketten a hatalom oldalán álltak, tevékeny részesei voltak – bár nem azonos módon – a pluralizmust felszámoló ideológiai harcnak, az MTA MTT tagjaként részt vettek a tudósok minősítésében, a tudósok „kijelölésében,” kirekesztésében (l. pl. *Bognár Cecil, Faragó László, Prohászka Lajos, Szabolcsi Miklós* sorsát). Kettőjük helyzetének, tevékenységének értelmezése szempontjából fontos tényező, hogy *Mérei* politikai és szakmai státusza az ideológiai és kultúrharcban megrendült, a *Rajk*-per után belekerült az elkötelezett kommunista értelmiségek ellen meginduló, szimbolikus és személyes elemekkel átszőtt szövevényes harcba, leszámolásba (*K. Horváth, 2006*). A felgyorsuló szelekciós, gyakran kontraszelekciós káderpolitika *Lázárt* továbbra is megtartotta, így ő megmarad a „kivívott” politikai és ezáltal szakmai pozícióiban.

Kettőjük kapcsolatának árnyalásához idézzünk fel néhány eseményt. Más vonatkozásban már bemutatunk egy dokumentumot, melyben kifejezésre jut *Mérei* pozitív véleménye *Lázár György* *kvalitásairól*. 1949. május 9-én *Mérei* az MTA MTT Társadalomtudományok Szakosztályának ülésén azt javasolta, hogy *Lázár* kerüljön az ONI-ba mint „rendkívül értékes elméleti káder.”²⁰ *Lázár* nem lett munkatársa az intézetnek, de *Mérei* javaslata – a pozitív megítélés mellett – azt is tükrözte, hogy a kommunista párt által elfogadott káder esetleg erősítheti az intézetet és benne *Mérei* pozícióit.

A kettőjük közötti viszony változására utal több dokumentum, levélváltásuk 1949 őszétől. Ezekben a levelekben *Mérei* nehezményezte egy tudóspótdíj és egy ösztöndíj elutasítását, valamint a tankönyvkészítés tudományos munkaként való minősítésének elvetését. *Mérei* azt kérte *Lázártól*, hogy véleménykülönbségüket tisztázzák, ha szükséges, a

¹⁹ MOL 276.f.86.cs. 16. ö.e.

²⁰ MTA Levéltár Magyar Tudományos Tanács 23/1.

Párt előtt is.²¹ Az ONI elleni támadások erősödésével 1950. február 14-én *Mérei* levelet írt *Lázár*nak, az MTA munkatársának, közreműködését kérve az intézet problémáinak megoldásához. Úgy tűnhet, mintha a levél elébe akarna menni a problémáknak, és azzal, hogy aktív fellépésre kéri *Lázárt*, illetve mulasztásaira figyelmezteti, *Mérei* ismét csak saját és intézete helyzetének stabilizálására törekedett. A levél témája az MTA és az ONI kapcsolata volt.²²

„Fel fogom itt sorolni konkrétan azokat az ügyeket, amelyek világosan megmutatják, hogy az Akadémia a Neveléstudományi Intézet ügyeit elhanyagolja. Jelentőségéhez méltóan nem foglalkozik azokkal és munkánkat nem megkönnyíti, hanem megnehezíti.”

Érdeemes szó szerint felidézni a *Mérei*-levél zárószóit:

„Kedves Lázár Elvtárs! E pontokat azért sorolom fel, hogy Téged személy szerint meggyőzzek arról, hogy mint az Akadémia pedagógiai ügyeinek előadója rossz úton jársz, ahelyett, hogy elfogulatlanul minden erőddel a Neveléstudományi Intézet erősítését, munkájának javítását tűznéd ki célul, szembeállsz vele és tevékenységeddel nehezíted munkáját.

Nem térek ki itt a pletykák halmazára, mert meg vagyok győződve arról, hogy azoknak legnagyobb része torzított és intrikus természetű emberek kártevő szándékaiból fakadnak. Csak tárgyilag fontos a kérdés, vagyis az, hogy a Neveléstudományi Intézet és az Akadémia pedagógiai előadója közt valamilyen hátsó gondolatoktól mentes kapcsolat létesüljön. Hogy e kapcsolat konkrét tervekre vonatkozzék, hogy az Akadémia pedagógiai előadója a pedagógiával foglalkozó legnagyobb intézettel együtt tudjon működni. Ha ez közvetlenül nem valósítható meg, akkor az egész kérdésben Pártunkhoz kell fordulni [...]”.

Mérei éles fogalmazása olyan szakember benyomását kelti e levél alapján, mint aki saját meggyőződésének igazában biztos abban, hogy kifogásai meghallgatásra találnak, és ezáltal elintéződnék a problémák. A korszellemhez alkalmazkodva a megoldás érdekében a hatalomhoz („Pártunkhoz”) való fordulást is kilátásba helyezte, de lehetőséget is adott a megegyezésre, hiszen csak személyes véleményét fejezte ki, még nem terelte hivatalos útra az ügyeket. Nincs olyan dokumentum a birtokunkban, amely arra utalna, hogy *Lázár* megválaszolta volna a levelet. A bemutatott levelek tartalma és stílusa már nem a *Lukács Györggyel Lázárért* vetélkedő, *Lázár* megbecsülését kifejező *Mérei* hangja.

Az események azt mutatták, hogy *Mérei* szakmai és/vagy ideológiai lépései a szocialista pedagógia megteremtése érdekében, a polgári pedagógia bírálata, majd önkritikája, nem jártak eredménnyel. 1950. március 29-én az MDP Központi Vezetőség Titkársága elmarasztaló határozatot hozott a Vallás- és Közoktatásügyi Minisztérium munkájával kapcsolatban. Az elmarasztalás *Mérei*nek és az Országos Neveléstudományi Intézetnek is szólt, és április 29-én minisztertanácsi rendelet megszüntette az ONI-t (*Knausz*, 1986). *Lázár* hírhedt írása, ami a kommunista párt elméleti-politikai fórumában, a Társadalmi Szemlében jelent meg, ezekhez az eseményekhez kapcsolódott, mintegy indoklásul, de

²¹ MTA levéltár II. Filozófiai és Történettudományi Osztály, 199/4. *Mérei* 1949. október 12-én és 22-én kelt levelei.

²² MTA Levéltár, II. Filozófiai és Történettudományi Osztály, 199/4. *Mérei* levele *Lázár Györgynek*.

főként határozott iránymutatásként és szigorú figyelmeztetésként a szakmai közvélemény számára (l. Lázár, 1950).

A Lázár-tanulmány tartalma, gondolatmenete arra épült, hogy *Mérei*nek a Magyar Pedagógiában megjelent önbírálatát (*Mérei*, 1949) ízekre szedje, és bebizonyítsa, önbírálatnak kevés, további érveket kell felsorakoztatni a pedológia ártalmassága mellett. *Knausz* (1986) elemzése szerint Lázár tanulmánya elsősorban az egyén-társadalom ket-tősség különböző megítélése alapján érvelt a pedológia ellen. A pedológia ártalmasságát Lázár a szovjet pedagógia nyomán abban látta, hogy a gyermeki fejlődés egyéni sajátosságait hangsúlyozva nem ad kellő teret a társadalmi (értsd szocialista szellemű) ráhatásoknak, és sokat bíz a spontán nevelődésre.

Érdeemes egy további szempontot is szemügyre venni Lázár cikkében: miként fonódik össze a szaktudomány és a politikai ideológia az érvelésben. Lázár többször hivatkozott mérvadó szovjet dokumentumokra, melyekben a „tudományellenes burzsoá pedológia” kifejezés már eleve utalt a pedagógia és a politika összekapcsolódására. Lázár hosszan idézte a szovjet kommunista párt 1936. júliusi határozatát, melyben a „pedológiai eltévelyedésekről” esett szó. A mai (fiatalabb nemzedékhez tartozó) olvasónak már az is szembeűnő, hogy egy párthatározat hoz normatív, kötelezően követendő állásfoglalást tudományos kérdéstről, ami szintén jelzi az említett összefonódást:

„A Szovjetunió Kommunista (bolsevik) Pártja Központi Bizottsága elítéli a jelenlegi, úgynevezett pedológia elméletét és gyakorlatát. Megállapítja, hogy az úgynevezett pedológia elmélete és gyakorlata áltudományos, marxista-ellenes tételeken alapszik. Ilyen tétel mindenekelőtt a divatos pedológia főtörvénye, amely szerint a gyermek sorsát biológiai és szociális tényezők, az átöröklés és valamiféle meg nem változtatható miliő hatásai végzetesen meghatározzák. Ez a mélységesen reakciós törvény kiáltó elmentmondásban van a marxizmussal és a szocialista építés egész gyakorlatával, amely sikerrel neveli át az embereket a szocializmus szellemében, felszámolja a kapitalizmus maradványait a gazdasági rendszerben és az emberek tudatában.” (Lázár, 1950. 253. o.)

Lázár még tovább ment a pedagógia és politika együttjárásának elfogadásában, amikor a „*mi pártszerű tudományunk*” kifejezést használta. Gondolatmenetéből jól látszik, hogy tudatosan gondolkodott pedagógia és politika összefonódásáról:

„szükséges, hogy a pedológia politikai tartalmán túl kitérjünk a pedológia főbb sajátságaira, feltárjuk elvi-világnézeti alapjait, s leleplezzük néhány legfontosabb tételét, szembesítve a marxista-leninista neveléstudomány felfogásával.” (Lázár, 1950. 261. o.)

Ez az akkori ideológiai légkörben természetesnek számított, hiszen a tudományok minden területén a sztálinizmus, a sztálini ismeretelmélet megkérdőjelezhetetlen „pártossága” volt irányadó. Ennek alátámasztására Lázár további mérvadó szovjet forrásokhoz fordult:

„A burzsoázia számára a pártyszerűség és az elméleti tárgyilagosság – egymással összeegyeztethetetlen dolgok. A proletáriátus számára a pártyszerűség és a tudományos megismerés – belsőleg összeforrtak.” (Csagin, 1949 idézi Lázár, 1950. 261. o.)

Így fonódott össze a pedagógiában – ezáltal a Lázár képviselte érvrendszerben – a tudomány és az ideológia.

Lázár frazeológiája és érvelése átpolitizált tudományos gondolatmenetével a szovjet pedagógiára támaszkodva próbálta elhíttetni a szakmával, hogy *Mérei* ekkorra már felhígított pszichológiai alapú pedagógiája káros, és a szocialista rendszer számára nem követhető, mert abban csak egy „igaz” tudomány létezik. Az alábbi, vádló hangvételű bekezdés jól példázza ezt:

„*Mérei* és társai természetesen arra törekedtek, hogy a pedológia minél nagyobb mértékben érvényesüljön az egyetemi és főiskolai képzésben. Szomorú tanulsággal szolgálnak ezzel kapcsolatban azok az egyetemi és főiskolai tantervek és vizsgai követelmények, amelyek még 1949 szeptemberében is érvényben voltak. Egyetemeink és főiskoláink az áltudományos imperialista nevelésmélet problematikájának alapos ismeretét kívánták meg pedagógus-jelöltjeinktől. Az akkori tantervek és vizsgarendek csak úgy hemzsegnek a különböző imperialista 'szaktekintélyek' neveitől. A marxista-leninista klasszikusoknak, vagy a nagy szovjet pedagógusoknak nevét hiába keresné az ember. Csak egy 1949 folyamán a vizsgai követelményekhez csatolt kiegészítés követeli meg néhány szovjet pedagógiai műnek és a szovjet iskolarendszernek ismeretét.” (Lázár, 1950. 275. o.)

A tanulmány osztályharcos szemléletét jól tükrözik azok a kifejezések, amelyek szinte háborús állapotokat írnak le a pedagógiában (kiemelések a tanulmány szerzőitől): a pedológiára mint a burzsoá-imperialista neveléstudomány *harc területére* utal; a burzsoá-imperialista nevelési elméletek *feladják tarthatatlanná lett hadállásaikat*; a pedológia a munkásosztálynak és a dolgozó parasztságnak *esküdt ellensége*; *Mérei* munkássága *belülről jött támadás* a szocializmus ellen (kiemelés az eredetiben); *Mérei* önbírálat nem a *hadállások teljes és végleges feladása*; *Mérei* *fegyverletételéről* beszél (kiemelés az eredetiben); *Mérei* *nem adta fel a harcot, csak visszavonul* (kiemelés az eredetiben); a pedológia maradványait *ki kell irtani*. Az írás formai kellékei közé sorolható a durva, sértő szóhasználat is: *Mérei* revizionizmusa gyalázatos és sületlen elméletekre vezet; *Mérei* társadalomlélektani gondolatai kapcsán Lázár megátalkodott porhintésről, kártevő pedológiáról ír; *Mérei* pedológiája a burzsoá erkölcs és szellemi rothadás szálláscsinálójának dicstelen szerepét játszotta.

Az osztályharc mesterséges élezése az ellenségkonstruálást is szolgálta. Kifejezte azt az ideológiát, melyben ki kell jelölni a barátot és az ellenséget, melyben fel kell adni az autonómiát, csak az lehet a hatalom partnere, aki a hatalom ideológiáját és politikáját elfogadja (Németh, 2009). A jó és a rossz, a barát és az ellenség sematikus szembeállításában Lázár ellenségként határozta meg *Méreit*. Lázár stílusa illeszkedett a korszak félelemkeltő légköréhez, illetve maga is generálta azt. Feltehetően a tartalom mellett a stílus is hozzájárult ahhoz, hogy ez a cikk kiemelkedett a többi, hasonló célú ideológiai bírálat közül.

Fontos lenne tudnunk, hogy ez az írás mennyiben tükrözi Lázár azonosulását a helyzettel és a szovjet pedagógiával, illetve mennyiben állja meg a helyét az a nézet, hogy megrendelésre, pártutasításra készült a tanulmány. A megírás körülményeiről a volt feleség egyik levelét érdemes idézni:

„A *Mérei* cikket apa elolvasta. Ami abban elmélet – elvi megállapítás – azt mélyen helyeslem, de ugyanakkor a gögös, személyeskedő kijelentéseket utáltam. És ehhez még nekem keserűséget okozott az is, hogy noha *Jofka* és *Mérei* ellentéte jelentékeny és reá-

lis volt, s régi keletű, de őt *Jofka* személy szerint nem tartotta sem ártalmas sem hitvány embernek. Azon fáradozott, hogy saját elvei helytelenségéről meggyőzze – sőt még baráti viszonyban is volt vele. Csak hosszas rábeszélés után jött *Jofka* arra a gondolatra, hogy *Mérei* valóban tudatosan ártalmas s ekkor írta meg a cikket.” – írta *Sós Júlia* 1954 márciusában apósának (*Kéri*, 2003. 151. o.).

A volt férjét reálisan látó, de mindig védelmező *Sós Júlia* visszatekintése mellett szükséges lenne még forrásokat találni e kérdés megválaszolására. Vélelmezhetjük *Lázár* addigi pályafutásából, személyiségének jellegzetességeiből, hogy kommunista elkötelezettségét komolyan gondolta, ezzel együtt igyekezett megfelelni közvetlen környezetének szakmai-politikai elvárásainak, hiszen ez egzisztenciális kérdés is volt számára.²³

Nyilvánosan nem jelent meg több írás *Lázártól Mérei*, illetve a pedológia ellen, de 1950 októberében a budapesti egyetem Pedagógiai Intézetének általa készített munkatervében újabb ítéletet fogalmazott meg:

„A Mérei-klikk elérkezettnek látta az időt arra, hogy ezt az Intézetet is a maga céljainak a szolgálatába állítsa. [...] *Mérei* célja: megakadályozni a lélektani káderek kiképzését, illetve megújítását a marxizmus-leninizmus szellemében...Pártunk ébersége azonban keresztül húzta a Mérei-klikknek ezt a tervét is.”²⁴

Azt gondolhatnánk, hogy a hatalmat kiszolgáló cikk után töretlenül felfelé ívelt *Lázár György* karrierje. De nem így történt. Bár 1950-ben és 1951-ben a budapesti egyetemen a Pedológia Intézet vezetője volt, tagja az MTA Pedagógiai Főbizottságának és az 1951-ben induló Pedagógiai Szemle szerkesztőbizottságának, de 1952-ben kikerült az egyetemről, könyvtárosként dolgozott, megszűntek a szakmai szervezetekben betöltött tagságai is, nem írt pedagógiai műveket.²⁵ Magánéletében sem volt sikeres. 1950-ben *Sós Júlia* szembefordult a *Mérei*-vitában férjével. „*Ez a probléma [Mérei] még szinte nyomtalanul elmúlt*”, de 1951-től kezdte bírálni férje magatartását, tartóssá váltak a veszekedések és 1954-ben elváltak (*Kéri*, 2003. 129. és 151. o.).

Az önkritikus Lázár

Lázár cikke a kortársak szemében vádiratként híresült el a szakmában, majd elhallgatás kísérte. A kommunikatív és kulturális emlékezet lassan elengedi a róla szóló, a pedagógiában a „gonoszt” megtestesítő elbeszélést, és vélhetően ennek helyét átveszi a tudományos, történeti megközelítés vagy a lassú felejtés.

Az írás megítélése először 1954-ben változott egy szűkebb szakmai nyilvánosságban. Az MTA Pedagógiai Főbizottsága 1954. június 12-én²⁶ az 1950. március 29-i párthatározat kapcsán megfogalmazta, hogy 1950-től megindult a reakciós, burzsoá pedagógia bírálata, megismerkedtek a szovjet pedagógiával, s ezzel együtt nagy arányokat öltött a

²³ Korabeli mércével mérve jelentős volt a fizetése, de azt is tudjuk, hogy nagyvonalúan jegyezte a Békekölcsonnt.

²⁴ MTA Levéltár 199. doboz.

²⁵ Ekkor dolgozott már szerkesztőként, szöveggondozóként *Apáczai Csere János* (1959, 1961) műveinek kritikai kiadásán.

²⁶ MTA Levéltár, Pedagógiai Főbizottság, 202. doboz.

burzsoá pedagógiai nézetek leleplezése, ám viszonylag kevés eredményt hozott az elméleti munka. Nem teremtődtek meg a pedagógiai kutatómunka feltételei, nem gondoskodtak a pedagógiai káderek kiválogatásáról. A fenti helyzet kialakulásához nagymértékben hozzájárult *Lázár György* működése. A bizottság összegzése fent idézett levéltári forrás alapján:

„A valóságtól elrugaszkodott, önkényes spekulációi jogos aggodalmat keltettek és igazolni látszottak azokat, akik a pedagógia elméleti kérdéseivel való foglalkozást meddőnek tartották.” A vélemény változásában érzékelhető volt a *Sztálin* halálát követő desztalinizáció, a *Nagy Imre*-féle reformkísérlet hatása.

Lázár később, 1956-ban arra utalt, hogy már 1954-ben felismerte tanulmányának ártalmasságát. Ez mindenképpen figyelmet érdemel, mert egybeesik az említett *Nagy Imre*-féle reformkísérletekkel. Akkor, 1954-ben már túl volt munkahelyén, az Egyetemi Könyvtárban egy erkölcsi természetű fegyelmi tárgyaláson.²⁷

A második eljárás, amely 1955-ben zajlott, politikai természetű volt, ahogy az alábbi jegyzőkönyvből megtudható:

„A most megindított fegyelmi eljárás során *Lázár György* beismerése és *Molnár József*, valamint *Szilágyi István* tanuk nyilatkozata alapján megállapítást nyert, hogy *Lázár György* az Egyetemi Könyvtár dolgozója előtt több ízben destruktív nézeteket hangoztatott. Legutóbb augusztus havában jelentette ki, hogy a Központi Vezetőség és Rákosi Elvtárs jelenlegi politikájával nem ért egyet, és Rákosi Elvtárs gazember, de gazember a pártbizottság is, amely kiszolgálja ezt a politikát. *Lázár György* azzal mentegette magát, hogy kijelentéseit erős felindulásban mondta el. Hivatkozott arra is, hogy túlfeszített idegállapota miatt nem minden kijelentését lehet beszámíthatónak felfogni. Egyben sajnálatát fejezte ki amiatt, hogy ezek a kijelentések elhangzottak. Ennek ellenére megállapítható, hogy *Lázár György* tevékenysége egyetemünk légkörére nézve káros, hiszen ilyen és ehhez hasonló nézeteket már több ízben hangoztatott. Miként a Fegyelmi Bizottság megállapította, a Rákosi Elvtárs személyére nézve hangoztatott destruktív kijelentések utólagos visszavonása nem tekinthető teljes értékűnek, és ezért sem szolgálnak enyhítő körülményként, mert *Lázár György* – miként ezt a fegyelmi tárgyaláson is beismerte – Pártunk politikájával ma sem ért egyet.”²⁸

Vélhetően a terhére felrótt destruktív nézeteket már korábban is hangoztatta, de ez a per beilleszkedett az oldódásra reflektáló resztalinizáció légkörébe, amikor *Nagy Imre* és hívei háttérbe szorultak.

A szakma figyelme 1956-ban irányult újra *Lázár Györgyre*. A resztalinizációt követő politikai változashoz igazodó általános szakmai mozgásban két alkalommal is önkritikát gyakorolt *Mérei* kritikájával kapcsolatban. A politikai széljárás változásával, a politikai és az ideológiai nyomás enyhülésével napirendre került *Mérei* tevékenységének enyhébb megítélése, *Lázár* szigorú, kíméletlen írásművének felülbírálata, a dogmatikus pedagógia elleni fellépés.

Az első önkritika az MTA Pedagógiai Főbizottság 1956. júniusi üléséhez köthető. Ekkor az MTA Pedagógiai Főbizottsága a „XX. kongresszus útmutatásai nyomán” a

²⁷ MOL XIX-J-2-f-14/855-L-5/1954.

²⁸ MOL XIX-J-2-f 40. doboz.

személyi kultusz és a dogmatizmus kérdéseit tárgyalta, a neveléstudomány tartalmi és személyi kérdései kerültek napirendre a vitában. Többen önkritikával illették korábbi szakmai tevékenységüket, s *Mérei* személye, szakmai tevékenysége szinte minden felszólalásban helyet kapott. *Lázár György* önkritikája jól illeszkedett ebbe a közegbe. Az 1950-es tanulmányát, mely zömmel politikai állásfoglalásokat tartalmazott, károsnak tartotta. Kijelentette, hogy *Mérei* jelentős képviselője az akkori szocialista magyar pedagógiának, és azt is elmondta, hogy az ominózus tanulmányban leírt tudományos nézeteit „revideálni” kell, s azt is kiemelte, hogy soha többé nem szabad előfordulnia, hogy valakit véleménye miatt eltávolítsanak munkahelyéről, ellenségnek bélyegezzék. Felszólalásának azonban csak kisebbik része foglalkozott *Méreivel*, az őt bíráló cikkével; java-részt a népi kollégiumokkal kapcsolatos vádakát igyekezett cáfolni.

A Főbizottság ülésén hozzászólók egy része nevesítve reagált *Lázár György* önkritikájára. *Jóború Magda* megjegyezte, hogy *Lázár* későbbi években kényszerítő körülmények nélkül is követelte a *Mérei* téves nézetei elleni harcot²⁹. *Ravasz János* nem tartotta megnyugtatónak *Lázár* önkritikáját. Többen általánosságban név nélkül ítélték el a *Mérei*-vita stílusát és következményeit az érintett személyek és a tudomány szempontjából. Voltak olyan résztvevők is, akik tágabban vetették fel a felelősség kérdését, például *Kontra György* így szólalt fel:

„Mindenki bűnös, aki arra, ami történt, nem mondta, hogy helytelen. Annak idején hozza nem értő, tapasztalan fiatalok ítélkeztek felelős, tudományos képzettséggel rendelkező emberek felett. Az ilyenekkel könnyű volt bármit elhítenni a pedagógiáról.”³⁰ Mai szemmel olvasva ezt a vitát, nem szabadulhatunk attól a gondolatától, hogy a megszólaló szakemberek ugyanúgy megfelelési kényszernek engedelmessé váltak, mint korábban, csak akkor szakmailag más, kollégákat egzisztenciálisan is súlyosan érintő nézeteket hirdettek a diktatúrát szolgáló politikai ideológia szellemében.

A második önkritika a Petőfi Kör pedagógusvitájában történt. Az olvadás légkörében a Demokratikus Ifjúsági Szövetség (DISZ) égisze alatt szerveződött vitában 1956 őszén *Lázár* ismét a népi kollégiumok jelentőségéről beszélt, de egy neki feltett kérdés kapcsán részletesen kitért elhíresült tanulmányára is (*Hegedűs és Rainer*, 1992. 67–68. o.). Kijelentette, hogy cikke „eszközévé vált annak, hogy *Mérei* elvtársat hat esztendőre eltiltották, elűtötték a tudományos munkásságtól [...] megsértettem, ha jóhiszeműen is, a pedagógia szabadságát.” (*Hegedűs és Rainer*, 1992. 67. o.)

Érdeemes felfigyelni *Lázár* önkritikájára abból a szempontból is, hogy a tudomány és a politikai ideológia ebben is érintkezik, de már nem fonódnak össze, ahogy korábban, hiszen ekkor a politikai elvárás részben az volt, hogy önkritikusan szemléljék a pedagógia szereplői saját korábbi szakmai, politikai tevékenységüket:

²⁹ A Magyar Tudományos Akadémia Pedagógiai Főbizottságának 1956. június 30-i ülése. A magyar pedagógia a kibontakozás útján. A gyorsírást jegyzőkönyv alapján készült kivonat. *Pedagógiai Szemle*, 6. 4. sz. 289–316.

³⁰ A Magyar Tudományos Akadémia Pedagógiai Főbizottságának 1956. június 30-i ülése. A magyar pedagógia a kibontakozás útján. A gyorsírást jegyzőkönyv alapján készült kivonat. *Pedagógiai Szemle*, 6. 4. sz. 301. o.

„Meg kell, hogy mondjam, hogy 1954-ben rájöttem arra, hogy cikkemnek *Mérei* elvtárs politikai magatartására vonatkozó része rágalom. Ezt önkritikailag megüzentem *Mérei* elvtársnak. Később arra is rájöttem, hogy a cikk egy sereg vonatkozásban tévesen, hamisan, egyoldalúan világítja meg a pedagógia kérdését tudományos szempontból is, tehát tudományos szempontból sem minden tekintetben kielégítő.” (*Pedagógusvita*, 1992. 67–68.)

Bár elismerte, hogy tévedett, de mentegette is magát és így beszélt félrevezetéséről:

1954-ig „magam is meg voltam győződve, hogy ez a lépés helyes volt. Pedig a pedagógiai szabadságnak a pedagógiai kutatás szabadságává is kell válnia. Ezzel a cikkel megsértettem, ha jóhiszeműen is, a pedagógia szabadságát. Ez olyan jóhiszeműség, amely megbocsáthatatlan. Jóhiszeműségem mindenestre onnan ered, hogy felelős pártfunkcionáriusok [...] *Fazekas Erzsébet* elvtársnő, valamint *Lengyel Klára*³¹ elvtársnő megmagyarázták nekem azt, hogy *Mérei* elvtárs ellen olyan vizsgálat folyik, amely már eredményeket mutatott fel és amely kétségtelenné teszi azt, hogy *Mérei* elvtárs működésének ellenséges tendenciája volt. Én ezt elhittem.” *Lázár* válaszában azt is megfogalmazta, hogy igazságot kell tenni, hogy ennek érdekében elemző cikket fog írni, s biztos abban, hogy még együtt fog dolgozni *Mérei Ferenc*cel.³²

Fontos látnunk, hogy a vitában felszólalók mindegyike, a kirekesztők és a kirekesztettek egyaránt megmaradtak a marxista terminológia keretei között, nem léptek ki ebből a világnézetből, csak szélsőségeivel akartak szakítani. Mind *Lázár*, mind a szakma Petőfi Kör-beli önkritikájának értelmezéséhez érdemes felidézni, hogy 1956 szeptemberében a párt rehabilitálta *Méreit*, így kevesebb kockázattal lehetett szólni a dogmatikus pedagógia felülvizsgálatáról, az egzisztenciálisan hátrányosan érintett személyekről és nézeteik hibás megítéléséről.

Zárógondolatok

Meddig és hogyan maradt meg a kommunikatív és a kulturális emlékezetben *Lázár* támadása *Mérei* és a pedagógia ellen? Megkockáztathatjuk azt az állítást, hogy *Mérei* személyének jelentősége volt az, ami *Lázár* nevét mitikussá növesztette. Bár *Mérei* megítélése ambivalens maradt a börtönből való 1963-as szabadulása után, mégis informális hatalommá vált a pszichológiában, és ez hatott a pedagógiára is. Azt is feltételezhetjük, hogy az „ördögi” *Lázár* mítosza nemcsak *Mérei* nagyságából következett, hanem *Lázár*-nak a kortársak szemében megszállottnak, indulatosnak tartott személyiségéből. Elítélését felerősíthette, hogy a tudományban tehetsége ellenére nem volt igazán elismerhető teljesítménye. Azt is állíthatjuk, hogy *Lázár* nem használta ki szakmai lehetőségeit az adott keretek között, pedig a politika magasra repítette és felkínálta az érvényesülést számára.

³¹ *Gerőné Fazekas Erzsébet* a Petőfi Kör vitája idején az ELTE Új és Legújabbkori Történeti Tanszékének a vezetője, *Lengyel Klára* a Lenin Intézet tanára.

³² Ez azonban nem valósult meg, részben *Mérei Ferenc* forradalom utáni bebörtönzése, részben *Lázár György* pedagógiai kutatásából való kiszakadása következtében.

Tanulmányunkkal – szándékaink szerint – közelebb kerültünk a *Lázár*-mítosz megértéséhez. Forrásaink elsősorban abban segítettek, hogy a pedagógia és a politika *Lázár* személyében megtestesülő összefonódását érzékeltessük. A még meglevő nyitott kérdések és feltáratlan források további kutatásra ösztönözhetik a kor és pedagógiája iránt érdeklődő szakembereket.

Köszönetnyilvánítás

Köszönetet mondunk *Garai Imre* doktorandusz hallgatónak, amiért az Eötvös Kollégium levéltári anyagában található vonatkozó forrásokra felhívta a figyelmünket.

A tanulmány az OTKA K100496 számú, „A magyar neveléstudomány története a szakmai folyóiratok tükrében (1945–1989) – tudományos kommunikáció, szakmai diskurzusok” című kutatásának keretében készült.

Irodalom

- Apáczai Csere János (1959): *Magyar Enciklopédia I.* Akadémia Kiadó, Budapest.
- Apáczai Csere János (1961): *Magyar Enciklopédia II.* Akadémia Kiadó, Budapest.
- Congdom, L. (1998): Bűn és büntetlenség. Az ismeretlen Lakatos Imre. *Replika*, 29. sz. <http://www.c3.hu/scripta/scripta0/replika/honlap>. Utolsó letöltés: 2012. március 13.
- Csagin, B. A. (1949): *A filozófia pártszerűsége és a polgári objektivizmus elleni harc.* Szikra, Budapest.
- Dessewffy Tibor (1999): *Iskola a hegyoldalban.* Új Mandátum Könyvkiadó, Budapest.
- Golnhofér Erzsébet (2004): *Hazai pedagógiai nézetek 1945–49.* Iskolakultúra könyvek, Pécs.
- Golnhofér Erzsébet (2006): Rendszerváltások és egyéni élettörténetek. *Educatio*, 15. 3. sz. 539–552.
- Golnhofér Erzsébet és Szabolcs Éva (2008): Múltidézés egy megtalált forrás segítségével. In: Golnhofér Erzsébet és Szabolcs Éva (szerk.): *Iskola és történeti emlékezet. Felmérés 1947-48-ból.* Gondolat Kiadó, Budapest. 11–32.
- Hegedűs B. András és Rainer M. János (1992, szerk.): *A Petőfi Kör vitái hiteles jegyzőkönyvek alapján VI. Pedagógusvita.* Múzsák – 1956-os Intézet, Budapest.
- Hovanyec László (1984): Kovács Imre. Politikuspályák. In: Sánta Ilona (szerk.): *Politikuspályák.* Kossuth Könyvkiadó, Budapest. 202–215.
- Huszár Tibor (1995): *A hatalom rejtett dimenziói. Magyar Tudományos Tanács 1948-49.* Akadémiai Kiadó, Budapest.
- Kéri Piroska (2003): *Sugárkoszorú. Sós Júlia és köre.* Abacus, Budapest.
- K. Horváth Zsolt (2006): Szexuál-lélektani szubkultúra az 50-es évektől. Vázlat a Törzs keletkezéséről és politikai szocializációjáról. In: Borgos Anna, Erős Ferenc és Litván György (szerk.): *Mérei életmű.* Új Mandátum Kiadó, Budapest. 37–55.
- Knausz Imre (1986): A magyar „pedológia” pere 1948-1950. *Pedagógiai Szemle*, 36. 11. sz. 1087–1102.
- Knausz Imre (1988): Szakszerűség és politikum az Országos Neveléstudományi Intézetben. *Pedagógiai Szemle*, 38. 11. sz. 1042–1047.
- Lázár György (1950): A magyar pedológia visszavonulási taktikája. *Társadalmi Szemle*, 5. 3–4. sz. 250–276.

- Mérei Ferenc (1948): A falakon belül. A neveléstudomány feladata. *Köznevelés*, 4. 22. sz. 549–552.
- Mérei Ferenc (1949): A burzsoá objektivizmus a gyermeklélektanban. *Magyar Pedagógia*, 57. 1. sz. 68–78.
- Mihályi Gábor (1998): Egy életkudarc története. Lakatos Imre feltámadása. *Replika*, 30. sz.
<http://www.c3.hu/scripta/scripta0/replika/honlap>. Utolsó letöltés: 2012. március 13.
- Németh András (2009): A magyar egyetemi neveléstudomány szocialista paradigmájának kialakulása az 1945–1953 közötti időszakban. In: Németh András és Biró Zsuzsanna Hanna (szerk.): *A magyar neveléstudomány a XX. század második felében*. Gondolat Kiadó, Budapest. 26–56.
- Petriné Feyér Judit és Szabolcs Éva (1998): *Nagy Sándor 1916–1994*. OPKM, Budapest.
- Romsics Ignác (1999): *Magyarország története a XX. században*. Osiris Kiadó, Budapest.
- Sári Katalin (1996): *Mikó Zoltán emlékezete (1910 – 1945). A nemzeti ellenállás mártírja*. Püski Kiadó, Budapest.
- Sáska Géza (2006): A reformpedagógia alakváltozása az 1945-ös „kis” és az 1947 utáni „nagy” rendszerváltást követő időszakban. *Magyar Pedagógia*, 106. 4. sz. 263–285.
- Somlai Péter (2011): *Nemzedéki konfliktusok és kötelek*. www.szmi.hu/download.php?file_id=1023. Utolsó letöltés: 2012. február 1.
- Staar Gyula (2007): *Tudóssors Közép-Európában*. Valóság, 50. 11. sz. 93–113.
- Szabolcs Éva (2006, szerk.): *Pedagógia és politika a XX. század második felében Magyarországon*. Eötvös József Könyvkiadó, Budapest.
- Szigeti József (2000): *Intellektuális önéletrajzom*. Eötvös József Könyvkiadó, Budapest.
- Zrinszky László (2006): *Válaszok feltett és feltehetően kérdésekre*. Kézirat.

ABSTRACT

ERZSÉBET GOLNHOFFER AND ÉVA SZABOLCS: GYÖRGY LÁZÁR AND HUNGARIAN PEDOLOGY:
MYTH AND REALITY

This paper covers a highly sensitive period in Hungarian history: the years after World War II and up to the Revolution of 1956 that saw the emergence of the communist party as the country's ruling power. In the mythology of the history of education in post-war Hungary, *Lázár* is seen as the quintessential political representative of the communist dictatorship and has been demonized in the cultural and communicative memory. His name is associated with an attack against Hungary's researchers in pedology – an early trend in psychology and pedagogy that studied children's behaviour and development. His career both as a communist party bureaucrat and as an educational researcher is analysed. Based on archival sources, we attempted to understand how the repressive domestic environment, Soviet pressure and his own personal attitudes determined his professional choices. Elements of personal narrative and information on contemporary politics and research policy have guided us in interpreting the complex background of *Lázár's* decisions. Data on his personal and professional life, including his career as a party bureaucrat, are presented as well as information on his retreat from the militant ideological and professional position he took during these years. We conclude that the demonic, mythological character of his personality was due to three reasons: his attacks were directed against renowned researchers, he was obsessive in character, and he lacked substantial scholarly achievements.

Magyar Pedagógia, **113**. Number 3. 133–151. (2013)

Levelezési cím / Address for correspondence: Golnhofer Erzsébet és Szabolcs Éva, ELTE Pedagógiai és Pszichológiai Kar Neveléstudományi Intézet, H-1075 Budapest, Kazinczy u. 23–27.

A FOGALMAZÁSÉRTÉKELÉS MEGBÍZHATÓSÁGA KÉT FÜGGETLEN BÍRÁLÓ ÉRTÉKÍTÉLETEINEK ELEMZÉSE ALAPJÁN

Nagy Zsuzsanna

Szegedi Tudományegyetem, Oktatáselméleti Kutatócsoport

Az anyanyelvi kommunikációs képességek, köztük a megfelelően fejlett írásbeli szövegalkotás olyan eszköztudás, amit az élet minden területén alkalmazunk, és amelynek minősége alapvetően meghatározza tevékenységeink sikerességét. A külföldi neveléstudományi kutatásoknak (l. pl. *MacArthur, Graham és Fitzgerald, 2008; Hidi és Boscolo, 2007; Torrance, van Waes és Galbraith, 2007; Mäki, 2002*) erős és egyre hangsúlyosabbá váló, dinamikusan fejlődő iránya a fogalmazási folyamatoknak, illetve körülményeiknek, a fogalmazásképeségnek, valamint az azt meghatározó tényezőknek a vizsgálata. Ez a megállapítás érvényes mind az anyanyelvi, mind az idegen nyelvi szövegalkotás kutatása terén. A hazai fogalmazáskutatás helyzetéről megállapítható, hogy bár egyre több hazai mérés (l. pl. *Molnár E. K., 2000, 2002, 2003, 2009; Szilassy, 2012; Pintér, 2009; Nagy, 2009*) foglalkozik e képesség vizsgálatával, még ma is elhanyagolt területe az írásbeli szövegalkotás kutatása a magyar neveléstudományon belül. A fogalmazásképeség fejlettségének neveléstudományi mérések során megvalósuló, megbízható értékelése hozzájárul ahhoz, hogy objektív módszerekkel feltárjuk és megismerjük tanulóink szövegalkotási képességének szintjét, és megértsük, milyen tényezők befolyásolják a diákok szövegeinek minőségét.

Míg a tudásszintmérő és sok egyéb képesség fejlettségét vizsgáló tesztek esetében az értékelés megbízhatósága statisztikai módszerekkel biztosítható, az objektivitás megfelelő javítókulcsok segítségével elérhető, sőt számítógépes értékelés esetén az értékelők szubjektivitásából fakadó hibafaktor ki is küszöbölhető, addig a tanulói szövegek értékelése a körültekintően kidolgozott értékelési skálák ellenére is a feladatjavítók eltérő szigorúságából, illetve a skálapontok definícióinak eltérő értelmezéséből eredő értékelési hibáktól terhelt. Ennek következtében a fogalmazáskutatásokban az eredmények érvényességének biztosításában az alkalmazott értékelési szempontrendszer megbízható működése, a szövegeket bíráló szakértők megfelelő képzése központi szerepet játszik.

A tanulmányban a fogalmazásképeség értékelésének kérdéseivel foglalkozunk. A szakirodalom elemzése alapján bemutatjuk a papír alapú szövegértékelés korábbi hazai és külföldi kutatásokban alkalmazott módszereit és eszközeit, rávilágítunk az értékelés megbízhatósága kapcsán felmerülő problémákra. Ezt követően egy hazai és nemzetközi mintákon alapuló, saját fejlesztésű fogalmazás-értékelési szempontrendszer működésének vizsgálatára irányuló kutatás eredményeit mutatjuk be. Célunk az, hogy az értékelési

skálák működését (1) két független bíráló értékítéleteinek vizsgálatával, (2) a klasszikus és a valószínűségi tesztelmélet által kínált módszerekkel elemezzük, valamint (3) a bírálók értékelései közötti hasonlóságok és különbségek megismerése révén a szempontrendszer működésére vonatkozóan átfogó kép birtokába jussunk.

A kutatás elméleti háttere

A fogalmazásképesség meghatározásai

Az írott szöveg megalkotásának képességét az IEA 1980-as években zajlott nemzetközi fogalmazásvizsgálata (*Kádárné*, 1990; *Gorman, Purves és Degenhart*, 1988) mint „a nyelvi, gondolkodási és együttműködési képességnek, valamint olvasás- és írástechnikai készségeknek” (*Kádárné*, 1990. 19. o.) kommunikációs célok megvalósítása érdekében működő együttesként írta le. *Kádárné* idézett meghatározása alapján a fogalmazásképesség egyrészt számos készség együttes működése által valósítja meg aktivitását, másrészt több más képességgel való egyidejű, egymást kiegészítő működése révén fejti ki hatását. Ennek megfelelően mind a hazai, mind a külföldi szakirodalomban fellelhető, hogy a fogalmazásképesség szerkezetét, működését modellező elméleti megközelítések összetett rendszerként tekintenek az írásbeli szövegalkotásért felelős képességre.

Nagy József (2002) személyiségmodelljében annak fejlődése szempontjából kiemelt szerepet kapnak a kommunikációs képességek, így a fogalmazás is. *Nagy* (1996) – összhangban a nemzetközi vizsgálatok értelmezésével – a fogalmazásképességet a kognitív kommunikációs képesség, ezen belül az írásképesség egyik – különféle képességekből, készségekből és egyéb komponensekből szerveződő – egységeként, „összefüggő gondolatrendszerek írásbeli közlésének szabályrendszereként” (64. o.) értelmezi. Ebben a definícióba a szöveg megírásához szükséges anyaggyűjtéstől, rendszerezéstől, az adott műfajhoz és közlési helyzethez illeszkedő megformáláson át, az elkészült szöveg végleges- sé alakításáig minden kognitív művelet beletartozik (*Molnár E. K.*, 2003).

A kognitív pszichológia klasszikus fogalmazásmodelljei a szövegalkotási képesség működését, fejlődését írják le (*Molnár E. K.*, 1996, 2003; *Eysenck és Keane*, 1997; *Beaugrande*, 1984; *Flower és Hayes*, 1980; *Hayes és Flower*, 1980; *Bereiter*, 1980; *Bereiter és Scardamalia*, 1987a, 1987b). *Molnár E. K.* (1996, 2003) kiemeli, hogy a klasszikus gondolkodás-lélektani modellek – az interaktív (*Beaugrande*, 1984), a rekurzív (*Flower és Hayes*, 1980; *Hayes és Flower*, 1980) és a képességintegrációs (*Bereiter*, 1980; *Bereiter és Scardamalia*, 1987a, 1987b) – közös vonása, hogy a jól írók (*expert, mature*) fogalmazási folyamatainak azonosításával a gyakorlatlan írók (*novice, immature*) problémáira világítanak rá. A kész szövegre fókuszáló fogalmazásszemlélettel szemben a kognitív pszichológiai irányzatok a szövegalkotási folyamatot állítják modelljük középpontjába, és azt összetett, többösszetevős rendszerként írják le, ahol a szöveg megalkotásának folyamata nem elszigetelten, hanem a feladatkörnyezettel (*Flower és Hayes*, 1980 idézi *Molnár E. K.*, 1996; *Hayes*, 1996) való folyamatos interakció mel-

lett zajlik, feltételezve például különböző szociális készségek és képességek működését is a fogalmazás közben.

A fogalmazások értékelésének szempontjai

Azt, hogy az írásművek értékelésekor milyen szempontokat érvényesítünk, meghatározza az, hogy a fogalmazásképeségnek és a fogalmazási folyamatnak milyen elméleti modelljét, meghatározását fogadjuk el és tekintjük érvényesnek. A fogalmazásképeséget, illetve a szöveg létrehozásának folyamatát modellező elméleti keretek elemzése révén a fogalmazási képesség szerkezetében azonosított összetevők, részképességek a fogalmazásértékelés során külön szempontokban minősíthetők.

Az iskolai osztályozás során megvalósuló fogalmazásértékelés jellemzően a fogalmazásképeségnek csak néhány aspektusát minősíti. A pedagógusok általában egy vagy több – alapvetően két – szempontból, tartalmi, valamint nyelvtani-stilisztikai-helyesírási tekintetben, a hagyományos ötfokú skála mentén minősítik a tanulók szövegeit (Molnár E. K., 2000), illetve egyre nagyobb arányban eltérő hosszúságú és részletezettségű szöveges értékelést adnak a diákok munkáiról (Nagy, 2011). Az ilyen típusú értékelés a fogalmazás két nagyobb, egymástól eltérő részképesség-csoportjának fejlettségét jellemzi: egyrészt a fogalmi kidolgozást, másrészt a különböző nyelvi konvenciók ismeretét (Molnár E. K., 2000).

A fogalmazásképeség értékelésére irányuló neveléstudományi kutatások szintén ezt a kétirányú felosztást követik, ugyanakkor jellemzően nem egyszerűen e két szempont szerint minősítik a fogalmazásokat, hanem több részképességet értékelnek eltérő számú szempont mentén (Molnár E. K., 2000). A fogalmazáskutatások szempontrendszerének kidolgozása komoly körültekintést igényel. El kell döntenünk, tárgyunk mely aspektusát kívánjuk feltárni, mire fektetjük a hangsúlyt az értékelésnél, és ennek megfelelően kell összeállítanunk a feladatokat. Szempontunk lehet, hogy a tanulók által alkotott szövegek mennyire hatékonyak, de vizsgálhatjuk azt is, hogy a felmérés résztvevői mennyire tudják a művelt köznyelvnek, a kért műfaj sajátosságainak megfelelő módon megfogalmazni gondolataikat. A hazánkban zajlott országos reprezentatív fogalmazáskutatások elsősorban ez utóbbira voltak kíváncsiak és ezt egészítették ki egyéb szempontokkal (Molnár E. K., 2003).

A szövegalkotási képesség mérésére irányuló vizsgálatok eszközei lehetnek a különböző kontextusba helyezett fogalmazásfeladatok, melyek megoldásakor a tanulóknak meghatározott szövegtípusoknak megfelelő szöveget vagy szövegeket kell alkotniuk meghatározott idő alatt. Ezeket – a vizsgálatok függő változójának, a szövegalkotási képesség mérésére készült – eszközöket egészíthetik ki a háttéradatokat, illetve más, az eredményeket feltehetően befolyásoló tényezők feltérképezésére készített kérdőívek, tesztek (Kádárné, 1990). A fogalmazási képesség méréséhez használt feladatok kiválasztásakor figyelembe kell vennünk, hogy a választott módszerek jelentősen befolyásolják azt, hogy ennek az összetett képességnek mely komponensei kerülnek a felszínre, mi válik valójában mérhetővé a feladatok értékelése által (Molnár E. K., 2000). Emellett Purves (1992) alapján Molnár Edit Katalin (2000) felhívja a figyelmet arra is, hogy a feladatmegoldás körülményei szintén hatást gyakorolnak arra, hogy a szövegalkotási ké-

pesség ki tud-e bontakozni a maga teljes valójában, a tanuló számot tud-e adni írásbeli kommunikációjának fejlettségéről, hiszen a fogalmazás folyamata, üteme egyéni, amely nem feltétlenül működik a diákra egyébként jellemző módon a mérési alkalom egy-két órája során.

A kutatások során összegyűjtött tanulói szövegek értékelésére a szervezők különböző szempontrendszereket dolgoznak ki, melyek segítségével egységesen pontozhatják a született dolgozatokat. A következőkben néhány hazai és nemzetközi példán mutatjuk be a fogalmazásértékelési szempontrendszerek típusait.

Az IEA 1980-as években lebonyolított vizsgálatánál (Written Composition Study 1984–1985) például *Takala* (1988, magyarul ismerteti *Kádárné*, 1990; *Molnár E. K.*, 2003) írásbeli kifejezőképesség szerveződését bemutató elméleti modelljére alapozott szempontrendszerrel értékelték a tanulói fogalmazásokat. *Takala* háromszintű modelljében a szövegtervezésben, illetve a szövegszerkesztésben való jártasság két-két készség-csoportra – az intellektuális készségek és a társas-együtműködési készségek, valamint a nyelvi készségek és az írás- és szövegrendezés készség körére – tagolódik. Ezek az ismeretek hat csoportját foglalják magukban: (1) fogalmak, gondolkodásmódok ismerete és szókincs; (2) anyagszervezési és -szerkesztési elvek; (3) kommunikációs normák, kifejezőmódok, stílus eszközök ismerete; (4) nyelvtani, nyelvhelyességi, írásjel-használati ismertek (5) helyesírási szabályok; (6) írásjegyek, jelrendszerek, írásmódok, formaszabályok ismerete (*Kádárné*, 1990). A fogalmazásvizsgálat eszközrendszerének kidolgozása-kor az ismeretek e hat köréhez rendelték értékelési kritériumokat. Ennek alapján az IEA-vizsgálat szempontrendszere hat analitikus és egy holisztikus szempontot tartalmaz, vagyis (1) a tartalom, (2) a felépítés/szerkezet, (3) a stílus, (4) a nyelvhelyesség, (5) a helyesírás és (6) a külső alak mellett a bírálók egy összbenyomás-osztályzattal is minősítették a tanulók írásait. Munkájukat a nemzetközi etalonskálák, mintadolgozatok is segítették. Az elkészült tanulói szövegeket minden esetben két bíráló értékelt. A végleges osztályzatokat zsűrimódszerrel alakították ki, kettejük folyamatos egyeztetése révén konszenzusos osztályzatokat állapítottak meg (*Kádárné*, 1990).

Orosz (1972) 18 szempont alapján értékelt a vizsgálata eredményeként készült szövegeket. A tartalmat négy további nézőpontból elemezte, így figyelt az anyaggyűjtésre, a fogalomválasztásra, valamint az ítéletalkotásra és -kapcsolásra is. A szerkesztést minősítő hét osztályzatban a szöveg globális megalkotástól egészen a szavak szintjéig vizsgálta és értékelt az írásokat, a stilizálás minősítésének pedig szintén hét szempontot szentelt.

Horváth (1998) 32 szempontot alkalmazott mérésében; a tartalom, szerkezet és nyelvi megformálás mellett az összbenyomás, a helyesírás, a külalak, a hatékonyság és a feladatspecifikáció szerint is értékelt a fogalmazásokat. A tartalmat minősítő tíz szempont között külön szerepel például a gondolatok és az álláspont kifejtettségének, az érvek alkalmazásának, a hitelességnek és a hatékonyságnak az értékelése. A szerkezeti jegyek között a gondolategységek elkülönítése mellett az egész szöveg, illetve a kisebb szöveg-egységek felépítését értékelő szempontok is megjelennek. A nyelvi megformáláson belül a hangnem, a stílus, a szókincs és a nyelvhelyesség megfelelését minősítő szempontok találhatóak.

Molnár Edit Katalin (2000, 2002, 2003) vizsgálatának fogalmazásértékelési szempontrendszere részben az IEA-mérés (*Kádárné*, 1990) skáláiból, részben egy más terüle-

teket – alapvetően írás- és helyesírási készséget – mérő vizsgálat (*Vidakovich*, 1990) szempontjaiból merített. A fogalmazások értékelése egy holisztikus és több analitikus szempont mentén történt. Míg a holisztikus szempont a tartalmi és a formai jegyek szétválasztása nélküli minősítette a tanulók munkáit, addig az analitikus szempontok egyes szövegjellemzőkre fókuszáltak. A tartalom, a szerkezet/műfaj/felépítés, illetve a stílus értékelésekor az IEA kritériumait vette figyelembe, a nyelvhelyességen belül a tanuló nyelvhasználatát, annak szabadosságát, világosságát, egyértelműségét jellemezte, a diákok munkáinak külalakját pedig a jelek elrendezettsége, a szöveg megjelenítésének szépsége szerint minősítette *Vidakovich* (1990) kategóriái alapján.

A bemutatott szövegvizsgálatok egymáshoz hasonló szempontrendszerrel dolgoznak. A különbségek abból adódnak, hogy a szempontokat különböző alszempontokra bontják, és így a három központi értékelési egység – a tartalom, a szerkezet és a stílus – egyes műveleteit külön osztályozzák (*Molnár E. K.*, 2003).

A nemzetközi szakirodalomban alkalmazott fogalmazás-értékelési szempontrendszer három csoportját különíthetjük el, melyeket az alábbiakban néhány példa bemutatásával szemléltetünk. A vizsgálatok egy része egy globális osztályzattal minősíti a mérés során nyert tanulói fogalmazásokat. Ezek az egyszerű értékelési rendszerek, hasonlóan a hagyományos iskolai értékeléshez, a szövegeket egy értékkel helyezik el a teljes mintában, ezzel fejezve ki a közöttük megfigyelhető minőségi különbségeket. *Davis* (2005) a nyelvtani tesztelés szövegalkotási teljesítményre gyakorolt hatását elemezte, főiskolás hallgatók szövegeit négyfokú holisztikus értékeléssel minősítette. *Chai* (2006) elsősorban a vizsgált tanulók szövegalkotás előtt készített vázlatainak minőségére, annak hatá-saira figyelt mérésében, összefüggést keresett a vázlatok és az elkészült végleges szövegek minősége között. Ebben a mérésben a fogalmazások értékelése egy holisztikus osztályzattal történt. *Gelati és Boscolo* (2009) négyfokú holisztikus értékelést alkalmazott általános iskolás tanulók körében végzett kísérletében.

Több vizsgálat dolgozik az IEA-mérés ismertített szempontrendszeréhez hasonló értékelési skálákkal. Ezekre a fogalmazáskutatásokra az jellemző, hogy szempontrendszerükben az IEA-szempontok között is megtalálható kritériumok jelennek meg, azokkal részben vagy egészében megegyező szempontokat tartalmaznak, esetleg finomítva, alszempontokra bontva egyes elemeket. *Engelhardt, Gordon és Gabrielson* (1991) 18 írásfeladatot értékelt tartalom és szerkezet, stílus, mondatforma, szóhasználat és nyelvtan szerint. *Zhang és Vukelich* (1998) az írás előtti tevékenységek (*prewriting*) hatásait elemző kutatásában egy holisztikus és öt analitikus szempontból értékelte a fogalmazásokat. Háromfokú skálán minősítette a mondat szerkezetet, a nyelvtant, a szóhasználatot, a kifejtést és az elrendezést. *Popp, Ryan, Thompson és Behrens* (2005) általános iskolás tanulók körében végzett mérésükben az ötletek, az elrendezés, a hangnem, a szóhasználat, a mondatok gördülékenysége és a nyelvtani szabályok betartása alapján jellemezték a diákok fogalmazásainak minőségét. *Barkai* (2007) szintén az IEA szempontjaihoz hasonló értékeléssel dolgozott, egy holisztikus és öt analitikus – tartalom, szerkezet, nyelvtan, írástechnika, stílus – szempont szerint jellemezte a szövegeket. *Crawford és Smolkowski* (2008) a fogalmazásokat három aspektusból – stílus és gördülékenység, tartalom és szerkezet, illetve nyelvhasználat szerint – négyfokú skálán minősítette.

A vizsgálatok egy másik hányada olyan szempontrendszert alkalmaz, amely vagy speciális szövegtípusok értékelését, vagy a szövegek árnyaltabb jellemzését teszik lehetővé újabb szempontok bevonásával. *Gearhart, Herman, Novak és Wolf* (1995) a nagymintás fogalmazásmérések értékelésrendszerének oktatási hatékonyságát vizsgálva az IEA-mérés rendszeréből is ismerős szempontok (fókusz/elrendezés, kidolgozás, írástechnika), valamint általános kompetencia és a *Writing What You Read* (WWYR) többszintű keretrendszer szempontjai alapján értékelték a tanulók szövegeit. A WWYR-rendszer egy holisztikus és öt analitikus dimenziót tartalmaz a tanulók narratív szövegeiben megjelenő téma, karakter, elrendezés, cselekmény és kommunikációbeli fejlődésének hatfokú értékelésére. A szövegalkotás médiumának, többek között, a szövegek minőségére gyakorolt hatásait vizsgálva *Whithaus, Harrison és Midyette* (2008) főiskolás hallgatók fogalmazásait négyfokú skálán értékeltette a vizsgálatban részt vevő bírálókkal a téma, az érvelés, a mondat szerkezet és a szóhasználat, logikai konzisztencia és a nyelvtani hibáktól való mentesség szerint.

Ransdell, Levy és Kellogg (2002), illetve *Ransdell és Levy* (1996) a fogalmazások értékelésére két független bírálót és komplex értékelési rendszert alkalmazott. A hat alcsoportos minőségi skála (*Six-Subgroup Quality Scale*) a szóhasználat és elrendezés, a technikai minőség, a tartalom, a szándék/közönség/hangnem, a szerkesztés és kidolgozás, valamint a stílus szempontjából minősíti a szövegeket, és ezeket az alcsoportokat bontja tovább, összesen 13 szemponttá a fogalmazások különböző aspektusainak minősítése érdekében. *Segev-Miller* (2004) tanárjelöltek fogalmazással kapcsolatos metakognitív stratégiáinak vizsgálatakor a résztvevők fogalmazásainak értékeléséhez több szempontot alkalmazott. Értékelte a témát, a makropropozíciót, a kidolgozást, a megfelelő retorikai struktúrát, az explicit kohéziót, a nyelvi megformálást, az értelmezést és az idézést.

Ezek az utóbbi vizsgálatok szempontrendszereikkel részletesebb fogalmazásértékelést valósítanak meg, illetve egy-egy kiválasztott szövegtípus vizsgálatát teszik lehetővé. A hagyományos iskolai fogalmazásértékelés során többnyire a tanulói szövegek holisztikus értékelése érvényesül és mellőzött az egyes szövegjellemzők minősítése. Ezen egyszerű és a fogalmazáskutatások egy része által is használt skáláknak az az előnye, hogy a szövegek gyors értékelését teszik lehetővé, és megvalósítják azt a kritériumot, hogy az egyes munkákat minőségük alapján elhelyezzük a teljes szövegkorpuszban. Ugyanakkor a globális értékelés nem biztosítja a fogalmazások különböző aspektusokból történő értékelését. A holisztikus értékelés nem képes arra sem, hogy a fogalmazásképesség fejlettségét az egyes szövegjellemzők mentén külön minősítse, ezáltal rávilágítson arra, hogy a tanuló a szövegalkotás közben egyes vizsgált tényezőket jobban, míg másokat gyengébben kezel. Kívánatos tehát a fogalmazások többszempontú minősítése az árnyalt szövegértékelés érdekében.

A fogalmazás-értékelési szempontrendszerek megbízhatósága és működésük statisztikai jellemzése

A fogalmazásvizsgálatokban különös problémát jelent a megbízhatóság biztosítása. A szövegértékelés eredményeinek érvényessége nagymértékben függ az alkalmazott értékelési skálapontok egyértelmű definiálásától, ugyanakkor ez nem biztosítja maradékta-

lanul a szempontrendszer megbízható működését. A skálapontok körültekintő megfogalmazása ellenére jellemző, hogy a bírálók eltérően értelmezik az értékkategóriákat. Ebből következik, hogy időnként ugyanazt a szöveget a különböző értékelők eltérően minősítik. *Vígh* (2010) hasonló problémákra hívja el a figyelmet az idegen nyelvi íráskészség pontozásának kérdéseit elemezve. Több – *Vígh* (2010) által is hivatkozott – nemzetközi kutatás (*Weigle*, 1998; *Engelhard és Myford*, 2003; *Eckes*, 2005, 2008; *Schoonen*, 2005) eredménye alapján a vizsgázók szövegeinek értékelésekor a bírálók több területen – például szigorúságukban, az értékelési szempontrendszer alkalmazásában, valamint a gyengébb és jobb teljesítményű diákok teljesítménye pontozásának következetességében – eltéréseket mutatnak. Az említett jelenségek hátterében több ok is állhat. *Szilassy* (2012) *Hillocksra* (1986) hivatkozva kiemeli, egy szöveg minősége nehezen fordítható le számokkal kifejezett értékké. A fogalmazások bírálóinak gondolkodása kulturálisan meghatározott. Emellett saját személyiségjegyeik, tulajdonságaik (*Molnár E. K.*, 2000; *Szilassy*, 2012), valamint bizonyos mértékben az értékelendő szöveg tartalmára, az általa megvalósított kommunikációs helyzetre vagy általában a szövegalkotásra, a szövegek minőségére vonatkozó tudásuk (*Nagy*, 2009), sőt a szerzőről való korábbi ismereteik (*Horváth*, 1998; *Szilassy*, 2012) is befolyással bírnak mások szövegeiről alkotott értéktételeikre. Ennek következtében a fogalmazásértékelés mindig valamilyen mértékű bírálói szubjektivitástól terhelt.

A fogalmazás-értékelési szempontrendszerek működését jellemző egyik mutató a reliabilitást kifejező *Cronbach- α* érték, ami a fogalmazásvizsgálatokban 0,75 felett már elfogadható. A holisztikus szempontra mint függő változóra végzett regresszióanalízis segítségével megadható az analitikus szempontok által megmagyarázott összes variancia, vagyis megállapítható, hogy az értékelők szövegről alkotott globális ítéletét mennyiben határozzák meg az analitikus szempontokban kifejezett szövegjellemzők, illetve értéktétele mennyiben tulajdonítható más, külső tényezőknek.

A fogalmazáskutatásokra jellemző, hogy a különböző szempontokhoz tartozó skálapontok pontos definiálása mellett az egyes tanulói dolgozatok több bíráló által történő értékelése révén igyekeznek növelni a mérés megbízhatóságát. Ilyen módon lehetőség nyílik arra, hogy az értékeléshez használt szempontrendszereket a *Cronbach- α* értékek és a regresszióelemzés mellett az értékelők ítéletei közötti korrelációkkal, illetve a *Kendall-féle* konkordanciaelemzés eredményeivel (pl. *Beyreli és Ari*, 2009) jellemezzék. A konkordanciaértékek az utóbbi elemzésekben 0 és 1 közé eső számok, melyek az értékelők közötti egyetértés mértékét fejezik ki. Minél nagyobb ez az érték, annál inkább azonos ítéleteket alkotnak a bírálók az egyes szempontok szerint.

Az értékelők ítéletei közötti korrelációkat (*r*) a különböző bírálók azonos szempontokra adott értéktételeinek összefüggései jelentik, és minél erősebbek ezek a korrelációs együtthatók, annál nagyobb az összhang az értékelők ítéletei között. *Isonio* (1991) vizsgálatában például ez az érték 0,76, *Beyreli és Ari* (2009) mérésben 0,66–0,83 között volt. Az IEA 1980-as évek végén végzett nemzetközi fogalmazáskutatásában is elemezték az értékelők ítéletei közötti korrelációk erősségét, melyek az összbemérés osztályzatok esetén 0,61 és 0,82 közötti értékeket mutattak, az összefüggések közepesek, illetve erősek voltak (*Gorman, Purves és Degenhart*, 1988; *Kádárné*, 1990).

A valószínűségi tesztelmélet széles körű elemzési lehetőségeket kínál a képességkutatások, így a fogalmazásvizsgálatok számára is (Molnár Gy., 2003, 2005, 2006, 2008, 2013; Vigh, 2010). Ennek ellenére csak kevés olyan hazai kutatást ismerünk, amely ilyen típusú vizsgálatokra vállalkozott volna a nyelvi képességek terén (pl. Vigh, 2008, 2010; Kontra, 2009; Molnár és Józsa, 2006; Dávid, 2008). Az írásbeli szövegalkotás vizsgálatait tekintve is megállapítható, hogy a parciáliskredit-modell nyújtotta lehetőségek segítségével a korábbinál szélesebb eszköztár áll rendelkezésünkre mind a fogalmazásképesség, mind az ennek értékelését lehetővé tevő szempontrendszerek működésének jellemzésére. Ugyanakkor nem tudunk olyan hazai kutatásról, amely a parciáliskredit-moddellel történő elemzések előnyeit kihasználva vizsgálta volna a tanulók szövegalkotási képességét, vagy tesztelte volna az alkalmazott mérőeszköz megbízhatóságát. Ugyanakkor a nemzetközi anyanyelven vagy idegen nyelven írt szövegek íratása által vizsgálódó fogalmazáskutatásokban (pl. Engelhard, 1994; Gyagenda és Engelhard, 1998; Griffin és Anh, 2005; Sugita, 2009; Barkaoui, 2011; Sudweeks, Reeve és Bradshaw, 2004; Wiseman, 2012) az értékelők szigorúságát, a szempontrendszerek működését az eredmények Rasch-moddellel történő elemzése révén is vizsgálják.

A nem dichotóm adatok, így a többfokú skálákon történő fogalmazásértékelések eredményeinek elemzésére alkalmas parciáliskredit-modell lehetővé teszi, hogy a szövegek minőségének jellemzéséhez használt értékelési szempontokat, skáláik működését jellemezzük (Molnár Gy., 2008, 2013). Mivel a parciáliskredit-modell közös skálán helyezi el a tanulók képességszintjét és az egyes szempontok nehézségi paramétereit, a vizsgált tanulók képességszintje alapján határozhatjuk meg a szempontok átlagos nehézségét. Ugyancsak vizsgálható az értékelési szempontok modellilleszkedése, a skálapontok egymástól való elkülönülése, a mintának való megfelelése, illetve az értékelők skálahasználata is. A Rasch-moddellel történő elemzések több értékelő munkájának jellemzését, a többaspektusú modell segítségével szigorúságuk összehasonlítását szintén lehetővé teszik (Vigh, 2010).

A vizsgálat módszerei és eszközei

Kutatásunk során 8. évfolyamos tanulóktól (N=429) az elbeszélés műfajában kértünk szövegalkotást. Az adatfelvételt a 2010–2011-es tanév első félévében került sor. Minden tanulói szöveget két független bíráló értékelésének vetettük alá. A fogalmazások értékelői a szövegek minősítését megelőzően pedagógiai mérés-értékelés területén szereztek képzettséget, rendelkeztek értékelői, feladatjavítói tapasztalattal, illetve tájékoztatást kaptak az alkalmazott fogalmazás-értékelési szempontrendszerről, próbaértékelés során tanulmányozták annak skáláit.

A szövegek minőségének megítéléséhez egy tíz szempontból álló, saját fejlesztésű értékelési rendszert használtunk, ami egy holisztikus és kilenc analitikus szempontból épült fel (Gorman, Purves és Degenhart, 1988; Kádárné, 1990; Molnár E. K., 2003; Ransdell és Levy, 1996). Az értékelőktől azt kértük, hogy hazai és nemzetközi mintákon alapuló szempontrendszerünk mentén, ötfokú skálán – (1) tartalom, (2) szövegtípus, il-

letve (3) hangnem szerinti feladattartás, (4) szerkezet és kidolgozás, (5) stílus, (6) érthetőség, (7) nyelvhelyesség, (8) helyesírás, központosítás, (9) külalak és olvashatóság – szerint minősítsék a tanulók munkáit. Mindemellett szintén ötfokú, globális értéklettel fejezzék ki az egyes dolgozatokról alkotott általános ítéletüket.

A tartalom értékletben az értékelők az anyaggyűjtést, a mondanivalót, a hitelességet, relevanciát, a tartalom mélységét és gazdagságát minősítették. A feladattartás szempontokban került sor arra, hogy a tanulóktól kért szövegtípust és a feladatnak megfelelő hangnemet vizsgáljuk. A szerkezet és kidolgozás skálája nem csak azt vizsgálta, hogy milyen a szöveg makroszintű elrendezése, a szöveg felépítésének logikája, hanem értékelte a szöveg belső koherenciáját, a bekezdések felépítését és egymáshoz való kapcsolódását is. A stílussal a nyelvi kifejezést, megformálást, a választékosságot minősítettük, de elválasztottuk ettől a feladathoz illeszkedő hangnem kiválasztását és megtartását. Az érthetőség szempontban a szöveg megértésre való előkészítettségét, a megfogalmazás világosságát vizsgáltuk. A nyelvhelyesség szempontban az írott köznyelv grammatikai szabályainak betartását értékeltük. A helyesírást és a központosítást egy szemponton belül értékelték. A külalakkal egy jegyben vizsgáltuk az olvashatóságot, itt néztük meg az íráskép rendezettségét, a jelek elrendezését, a szöveg megjelenítésének szépségét, a szó-távolságok és betűnagyságok arányosságát, egyenletességét. Valamennyi szemponthoz ötfokú értékelési skálát dolgoztunk ki, melyben röviden leírtuk azokat a szöveggeljellemzőket, amelyek a tanulók által létrehozott szövegekben megfigyelhetők.

A mérőeszköz értékelési skáláit korábbi, 2010 tavaszán 4. és 8. évfolyamos tanulók körében megvalósított fogalmazásvizsgálatunk (Nagy, 2010) alkalmával egy bíráló értékeléleteinek klasszikus tesztelméleti módszerekkel történő elemzése által már teszteltük. A kvantitatív elemzés eredményei szerint a szempontrendszer reliabilitása mindkét vizsgált évfolyamon megfelelő (Cronbach- $\alpha_1=0,94$, Cronbach- $\alpha_2=0,96$). Az összbenyomás értékletre mint függő változóra végzett regresszióanalízis eredménye szerint az analitikus szempontok körülbelül 91%-ban magyarázták a szövegek minőségéről átfogó jellemzést adó holisztikus osztályzatot, és valamennyi szemponton nyújtott teljesítmény szignifikáns ($p<0,01$) összefüggést mutatott az összbenyomás-jeggyel és a többi szempont szerint tapasztalt teljesítménnyel is. Eredményeink alapján a szempontrendszer működését kielégítőnek, az értékelési skálákat további használatra alkalmasnak találtuk (Nagy, 2010).

Jelen kutatásunkban a szempontrendszer működését további vizsgálatnak vetettük alá. Mivel ugyanazon tanulói minta fogalmazásainak két független bíráló által történő értékeléseivel rendelkezünk, lehetőségünk nyílt arra, hogy az értékelők ítéleteinek összevetésével további következtetéseket fogalmazzunk meg az alkalmazott fogalmazásértékelési rendszer megbízhatóságára vonatkozóan. Vizsgálatunkban a klasszikus tesztelmélet módszerei mellett alkalmaztuk a parciáliskredit-modell nyújtotta eszköztárat is. Az adatok elemzését az SPSS-programcsomag és a ConQuest (Wu, Adams és Wilson, 1998) elemzőszoftver segítségével végeztük el.

A vizsgálat eredményei

A szempontrendszer működésének vizsgálatára irányuló kutatás eredményeinek bemutatásakor előbb a klasszikus, majd a valószínűségi tesztelmélet módszereivel végzett statisztikai elemzéseket ismertetjük. A klasszikus tesztelmélet eszközszerére alapozó analízisek között kitérünk a két értékelő ítéletei alapján külön számított reliabilitásértékekre, majd bemutatjuk az egyes szempontokra adott értékek belső összefüggésrendszerét. Ezt követően a két értékelő azonos szempontokból alkotott ítéleteinek korrelációit közöljük, végül az összbenyomás osztályzatokra végzett regresszióanalízisek eredményeit elemezzük. Az ezt követő fejezetben a valószínűségi tesztelméleti módszerekkel végzett vizsgálatok, így a parciáliskredit-moddellel, illetve a többspektrális modellel folytatott elemzéseink eredményeit foglaljuk össze.

A szempontrendszer működése a klasszikus tesztelmélet modelljei alapján

A szempontrendszer működésének klasszikus tesztelméleti eszközökkel történő vizsgálatát a reliabilitásmutató meghatározásával kezdtük. Az összbenyomás szempont az értékelők szövegről alkotott általános benyomását fejezi ki, és ezt – ahogyan az a regresszióanalízisek eredményeinek bemutatásakor látható – az analitikus szempontokban kifejezett szövegjellemzők határozzák meg leginkább. Emiatt a globális skálát kihagytuk a reliabilitás vizsgálatából.

Az analitikus szempontok elemzése alapján a két bírálóra vonatkozóan két Cronbach- α értéket állapítottunk meg, melyek között csak néhány ezrednyi a különbség. A szempontrendszer megbízhatósági mutatója mindkét bíráló esetén megfelelőnek bizonyult, mindkét értékelőnél a Cronbach- α_1 0,95.

Az 1. és a 2. táblázat a bírálók értékítéleteink szempontonkénti korrelációit foglalja össze. Az eredmények alapján – mind az első, mind a második bíráló értékeit vizsgálva – valamennyi értékelési szempont szignifikánsan ($p < 0,001$) korrelált minden más értékelési szemponttal. A korrelációs együtthatók erősségének mintázata a két értékelő esetén hasonló. A tartalom, a feladattartás: szövegtípus, illetve hangnem, a szerkezet és kidolgozás, a stílus, az érthetőség és a nyelvhelyesség szempontok egymással való összefüggései bizonyultak erősebbnek, míg a helyesírás, központozás, valamint a külalak és olvashatóság osztályzatok gyengébben korreláltak a többi szemponttal. Az első bírálónál az utóbbi két jegy 0,37 és 0,65, a másodiknál 0,53 és 0,76 közötti erősségű összefüggéseket mutatott a szempontrendszer többi skálájával, míg az egyéb szempontok korrelációs együtthatói 0,73 és 0,95, illetve 0,69 és 0,84 közötti értékeket vettek fel. Ez a két szempont a korrelációk elemzése alapján a többtől némiképp függetlenebbnek bizonyult.

A helyesírás, központozás, illetve a külalak és olvashatóság skála az értékelők holisztikus ítéletével is a többi szempontnál gyengébb összefüggést mutatott. Az első bírálónál ezek az osztályzatok 0,60 és 0,40, a másodiknál 0,68 és 0,56 erősséggel korreláltak az összbenyomással. Ugyanakkor a többi szempont esetén mindkét értékelőnél magasabb, 0,78 és 0,91, illetve 0,81 és 0,88 közötti korrelációs értékeket tapasztaltunk.

A fogalmazásértékelés megbízhatósága két független bíráló értéktételeinek elemzése alapján

1. táblázat. Az első bíráló értéktételeinek szempontonkénti összefüggései

Szempontok	T	Fsz	Fh	SzK	S	É	Ny	HK	KO
Tartalom	–								
Feladattartás: szövegtípus	0,77	–							
Feladattartás: hangnem	0,85	0,88	–						
Szerkezet és kidolgozás	0,85	0,86	0,95	–					
Stílus	0,78	0,78	0,82	0,81	–				
Érthetőség	0,82	0,81	0,85	0,86	0,85	–			
Nyelvhelyesség	0,75	0,73	0,79	0,79	0,91	0,83	–		
Helyesírás, központosítás	0,57	0,57	0,59	0,61	0,62	0,63	0,65	–	
Külsőalak és olvashatóság	0,37	0,40	0,39	0,39	0,45	0,43	0,44	0,42	–
Összbenyomás	0,91	0,80	0,87	0,88	0,81	0,84	0,78	0,60	0,40

Megjegyzés: minden összefüggés $p < 0,001$ szinten szignifikáns; T=tartalom; Fsz=szövegtípus szerinti feladattartás; Fh=hangnem szerinti feladattartás; SzK=szerkezet és kidolgozás; S=stílus; É=érthetőség; Ny=nyelvhelyesség; HK=helyesírás és központosítás; KO=külsőalak és olvashatóság.

2. táblázat. A második bíráló értéktételeinek szempontonkénti összefüggései

Szempontok	T	Fsz	Fh	SzK	S	É	Ny	HK	KO
Tartalom	–								
Feladattartás: szövegtípus	0,78	–							
Feladattartás: hangnem	0,78	0,84	–						
Szerkezet és kidolgozás	0,80	0,84	0,82	–					
Stílus	0,77	0,76	0,79	0,78	–				
Érthetőség	0,80	0,75	0,77	0,80	0,83	–			
Nyelvhelyesség	0,72	0,71	0,76	0,69	0,81	0,74	–		
Helyesírás, központosítás	0,62	0,60	0,66	0,60	0,64	0,64	0,76	–	
Külsőalak és olvashatóság	0,51	0,48	0,53	0,50	0,54	0,54	0,58	0,53	–
Összbenyomás	0,86	0,85	0,88	0,87	0,86	0,86	0,81	0,68	0,56

Megjegyzés: minden összefüggés $p < 0,001$ szinten szignifikáns; T=tartalom; Fsz=szövegtípus szerinti feladattartás; Fh=hangnem szerinti feladattartás; SzK=szerkezet és kidolgozás; S=stílus; É=érthetőség; Ny=nyelvhelyesség; HK=helyesírás és központosítás; KO=külsőalak és olvashatóság.

Az első és a második bíráló által azonos szempontokra adott értékletek összefüggéseit a 3. táblázatban közölt korrelációs együtthatók alapján követhetjük nyomon. Valamennyi analitikus szempont és az összbenyomás esetén is a két bíráló ítéletei erősen ($0,85 < r < 0,93$) és szignifikánsan korrelálnak egymással. A magas r -értékek arra utalnak,

hogy a bírálók közel azonosan ítélik meg a tanulói szövegeket, bírálataik szoros összefüggésben állnak egymással valamennyi vizsgált jellemző mentén.

3. táblázat. A két bíráló azonos szempontokra adott értékítéleteinek összefüggései

Szempontok	Összefüggések (r)
Tartalom	0,85
Feladattartás: szövegtípus	0,86
Feladattartás: hangnem	0,89
Szerkezet és kidolgozás	0,87
Stílus	0,88
Érthetőség	0,85
Nyelvhelyesség	0,88
Helyesírás, központosítás	0,93
Külalak és olvashatóság	0,86
Összbenyomás	0,85

Megjegyzés: minden összefüggés szignifikáns $p < 0,001$ szinten.

Azonban az összbenyomás-teljesítményre mint független változóra végzett regresszióelemzések eredményei (4. táblázat) már különbségeket mutatnak a két bíráló értékelési teljesítményében. Bár az analitikus szempontok 88, illetve 90%-ban magyarázzák a holisztikus értékletek varianciáját, az egyes szempontok hozzájárulása az összbenyomáshoz igen eltérő. Ha megvizsgáljuk az első értékelő értékítéleteinek jellemzőit, akkor szembetűnő, hogy az összbenyomás varianciájának csaknem felét a tartalom határozza meg, és a többi szempont közül csak a szerkezet és kidolgozás hozzájárulása jelentősebb, a többié elenyésző. Az értékek szignifikanciáját vizsgálva szintén azt tapasztaltuk, hogy statisztikailag releváns mértékben csak ez a két szempont járul hozzá az értékelő szövegről alkotott általános benyomásához.

A második értékelő esetén (5. táblázat) sokkal kiegyenlítettebbek az értékek. A bíráló ítéletét legnagyobb mértékben a feladattartás: hangnem határozza meg, viszont a többi szempont magyarázóereje is kielégítő, a helyesírás, központosítás, illetve a külalak és olvashatóság kivételével szignifikáns is.

A klasszikus tesztelméleti módszerekkel végzett elemzések eredményeinek összegzéseként megállapítható, hogy bár mindkét bíráló esetén az egyes szempontok erős, illetve közepesen erős összefüggését tapasztaltuk, illetve az azonos szempontokra adott ítéletek között erős korrelációkat találtunk, a két értékelő bírálói gondolkodása nem ekvivalens. Az összbenyomásra végzett regresszióelemzések eredményei a két értékelő munkája között jelentős különbségeket jeleztek, az analitikus szempontok globális ítélethez való hozzájárulásának igen nagy különbségeit mutatták ki.

A fogalmazásértékelés megbízhatósága két független bíráló értékítéleteinek elemzése alapján

4. táblázat. Az összbenyomás osztályzatra mint függő változóra végzett regresszióanalízis eredménye az első bíráló esetén

Szemponatok	<i>r</i>	β	$r*\beta*100$
Tartalom	0,91	0,51	45,70*
Feladattartás: szövegtípus	0,80	0,05	3,84
Feladattartás: hangnem	0,87	-0,07	-5,85
Szerkezet és kidolgozás	0,88	0,32	28,39*
Stílus	0,81	0,08	6,37
Érthetőség	0,84	0,06	5,05
Nyelvhelyesség	0,78	0,04	2,81
Helyesírás, központosítás	0,60	0,02	1,16
Külalak és olvashatóság	0,40	0,01	0,18
Összes megmagyarázott variancia			87,65

Megjegyzés: Minden *r*-érték $p < 0,001$ szinten szignifikáns. *A β -érték $p < 0,001$ szinten szignifikáns.

5. táblázat. Az összbenyomás osztályzatra mint függő változóra végzett regresszióanalízis eredménye a második bíráló esetén

Szemponatok	<i>r</i>	β	$r*\beta*100$
Tartalom	0,86	0,18	15,13**
Feladattartás: szövegtípus	0,85	0,10	8,43*
Feladattartás: hangnem	0,88	0,20	17,54**
Szerkezet és kidolgozás	0,87	0,16	14,19**
Stílus	0,86	0,14	12,27**
Érthetőség	0,86	0,15	12,44**
Nyelvhelyesség	0,81	0,10	7,75*
Helyesírás, központosítás	0,68	0,02	1,28
Külalak és olvashatóság	0,56	0,02	0,94
Összes megmagyarázott variancia			89,96

Megjegyzés: Minden *r* érték szignifikáns $p < 0,001$ szinten.

*A β érték szignifikáns $p < 0,05$ szinten. **A β szempont magyarázó ereje szignifikáns $p < 0,001$ szinten.

A szempontrendszer működése a valószínűségi tesztelmélet modelljei alapján

Mint ahogy azt a reliabilitásmutatók elemzésekor is kifejtettük, az értékelők globális benyomását kifejező holisztikus szempont alapvetően másként viselkedik, mint az anali-

tikus skálák, amelyek valamilyen mértékben mind szerepet játszanak az általános ítélet kialakításában. Ezért a valószínűségi tesztelmélet módszereivel történő elemzésekbe is csak az analitikus szempontokat vontuk be. A vizsgálatból kizártuk a szövegek tartalmi és formai jegyeinek szétválasztása nélküli, a szöveg minőségéről alkotott általános képet reprezentáló összbenyomás szempontját.

A következőkben a valószínűségi tesztelmélet eszközeivel lefuttatott elemzések között a többaspektusú modellel és a parciális kreditmodellel történt számítások eredményeit és az ezekből levont következtetéseinket összegezzük. Bemutatjuk az értékelők szigorúsági paramétereit; a két bíráló ítéleteinek egymástól független elemzése alapján bemutatjuk az alkalmazott szempontok nehézségét és modellilleszkedését; majd elemezzük az értékelők skálahasználatában mutatkozó különbségeket.

Az értékelők szigorúsága

Az értékelők szigorúságának összehasonlítását a többaspektusú (*multifaceted*) modell teszi lehetővé. A vizsgálat során a modell a szempontokra adott ítéleteket elemezve felállítja az egyes bírálók szigorúsági paramétereit. Több bíráló esetén a legkevésbé és a leginkább szigorúan osztályozó értékelő, két bíráló esetén e kettő paramétereinek különbségét a szórással összevetve alkothatunk képet arról, milyen mértékű az értékelők szigorúsága közötti eltérés.

6. táblázat. Az értékelők szigorúsági paramétereit

Bírálók	Szigorúsági paraméter	Infit paraméter
Első	0,28	1,04
Második	-0,28	1,12

Az elemzés eredményei szerint (6. táblázat) a szövegtörzs értékelését végző szakértők szigorúsági paramétereit között mindössze 0,56 logitegységnyi a különbség, ami a szórással (2,5) összehasonlítva alacsonynak tekinthető, vagyis az értékelők szigorúságában csak kicsi az eltérés.

Az értékelési szempontok nehézségi indexe és modellilleszkedése

Az értékelési szempontok nehézségi paramétereit a két értékelő esetén két külön parciáliskredit-modell felállításával vizsgáltuk, így az eredmények közvetlenül hasonlíthatók össze. Nem az volt a célunk, hogy az egyes szempontokat összehasonlítsuk, hanem az, hogy az értékelők munkáját külön jellemezzük. A 7. és a 8. táblázatban közölt adatok alapján a szempontok nehézségét kifejező középérték-logitok az első bíráló értékelésében a helyesírás és a külalak kivételével az átlagos képességszint (0 logit) feletti, a második bírálónál a tartalom, a nyelvhelyesség és a külalak nehézsége helyezkedik el az átlagos képességszint alatt. Mindkét bíráló a külalak és olvashatóság szempontot ítéli meg a legenyhébben és a szerkezet és kidolgozást a legszigorúbben.

A fogalmazásértékelés megbízhatósága két független bíráló értéktételeinek elemzése alapján

7. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése az első bíráló értéktételei alapján

Szempontok	Középtérték-logitok	Standard hiba	Infit paraméter
Tartalom	0,30	0,07	0,80
Feladattartás: szövegtípus	0,26	0,07	0,90
Feladattartás: hangnem	0,46	0,07	0,60
Szerkezet és kidolgozás	0,59	0,07	0,60
Stílus	0,18	0,07	0,80
Érthetőség	0,42	0,07	0,70
Nyelvhelyesség	0,03	0,07	0,80
Helyesírás, központosítás	-0,12	0,07	2,00
Külsőalak és olvashatóság	-2,14	0,20	2,80

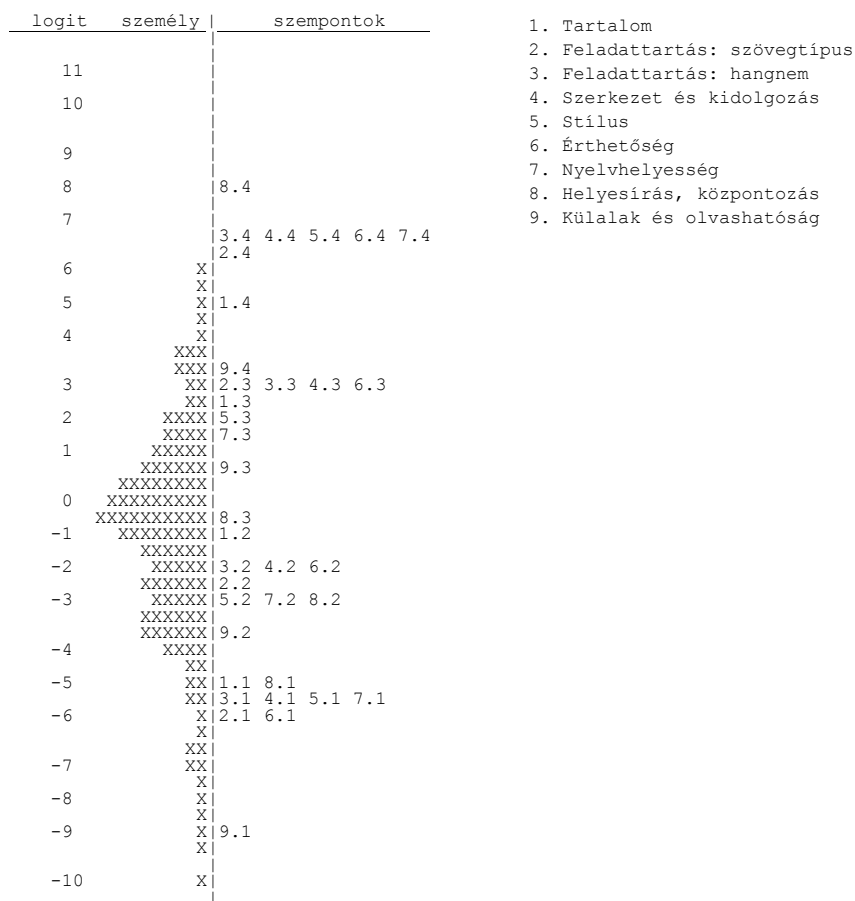
8. táblázat. Az értékelési szempontok nehézségi indexe és modellilleszkedése a második bíráló értéktételei alapján

Szempontok	Középtérték-logitok	Standard hiba	Infit paraméter
Tartalom	-0,21	0,07	0,90
Feladattartás: szövegtípus	0,02	0,07	0,90
Feladattartás: hangnem	0,37	0,07	0,80
Szerkezet és kidolgozás	0,52	0,07	0,80
Stílus	0,38	0,07	0,80
Érthetőség	0,19	0,07	0,80
Nyelvhelyesség	-0,04	0,07	1,00
Helyesírás, központosítás	0,07	0,07	1,60
Külsőalak és olvashatóság	-1,32	0,19	2,00

A szempontok modellilleszkedését vizsgálva kutatásunkban *Park* (2004 idézi *Vigh*, 2010) nyomán az elfogadható értékek sávját 0,8 és 1,2 logit között határoztuk meg. Ennek alapján az első bíráló esetén a feladattartás: hangnem, a szerkezet és kidolgozás, az érthetőség, valamint a helyesírás, központosítás és a külsőalak és olvashatóság sem illeszkedik jól a modellhez. A második bíráló esetén nem illeszkedő szempontként a helyesírás, központosítást és a külsőalak és olvashatóságot kell kiemelnünk. E két szempont esetén az infit paraméter mindkét értékelőt tekintve 1,2 logit felett (2,0, és 2,8, illetve 1,6 és 2,0) helyezkedik el, vagyis ezeket a szempontokat többdimenzionalitás jellemzi, a többi szemponthoz képest mást mérnek, tőlük igen függetlenek, illetve értékelésükbe más szempontok is közrejátszanak (*Vigh*, 2010).

Az 1. és a 2. ábrán a két értékelő esetén külön lefuttatott parciáliskredit-modell outputjaként kirajzolt személy-szempont térképek láthatók, melyek közös skálán ábrázolják

a minta fogalmazásképeségének fejlettségét és az értékelési skálapontok *Thrustone*-küszöbértékeit. Ez utóbbi értékek azokat a képességszinteket jelölik, amelyek mellett a tanulók 50%-os valószínűséggel kapják munkájukra adott szempontból az adott értéketet (*Molnár Gy.*, 2008, 2013; *Vígh*, 2010). Az ábrák bal oldalán elhelyezkedő 11 és -10 közötti számsor (logitskála) értékei az átlaga alatti, átlagos és átlag feletti képességszinteket mutatják (*Molnár Gy.*, 2008, 2013; *Vígh*, 2010). Az ábrák alapján a tanulók képességszintje nagyon széles skálán, 21 logitegységnyi területen helyezkedik el. A 90 fokban elforgatott eloszlásgörbére emlékeztető ábra X-ei mindkét ábrán körülbelül 3 tanulót jelölnek, akik a logitskála adott pontján helyezkednek el képességparamétereik alapján. Az ábrák jobb oldalán elhelyezkedő számok az egyes szempontok (1–9. magyarázatukat l. az 1. és 2. ábra jobb oldalán) esetén az egyes skálapontok küszöbértékeit jelentik (*Molnár Gy.*, 2008, 2013; *Vígh*, 2010).

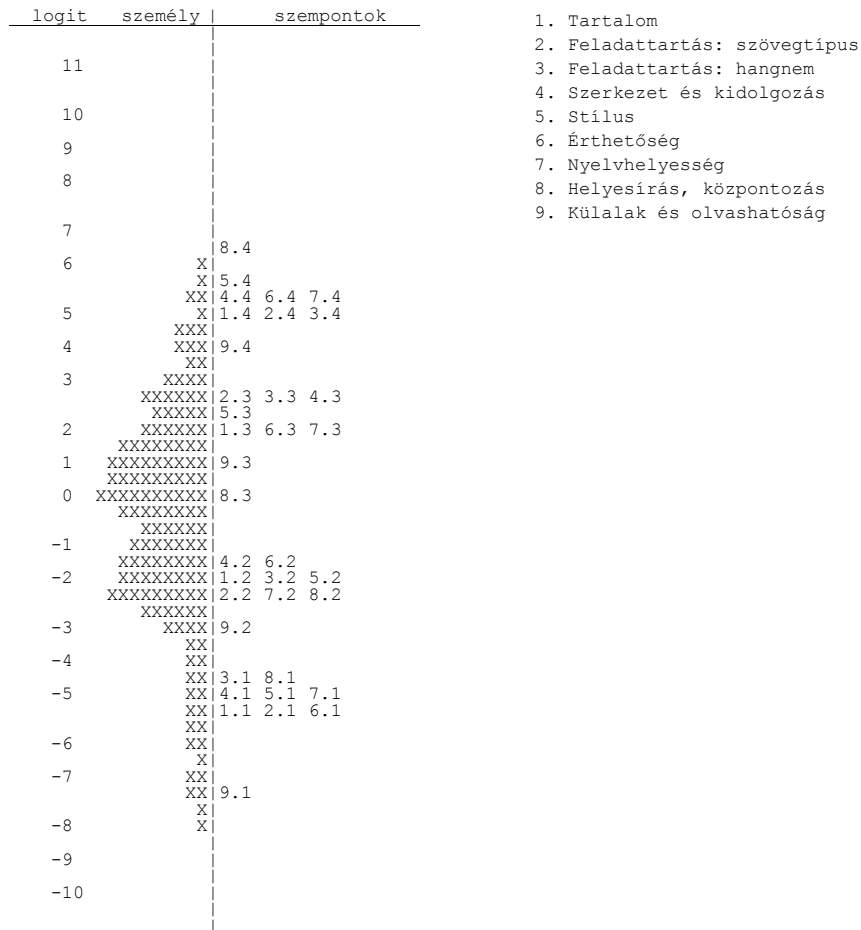


1. Tartalom
2. Feladattartás: szövegtípus
3. Feladattartás: hangnem
4. Szerkezet és kidolgozás
5. Stílus
6. Érthetőség
7. Nyelvhelyesség
8. Helyesírás, központozás
9. Külalak és olvashatóság

1. ábra

Az értékelési szempontrendszer személy-szempont térképe az első bíráló értékítéletei alapján ($X \approx 3$ tanuló)

Az 1. és 2. ábráról leolvasható, hogy a szempontrendszer skálapontjai mennyire jól illeszkednek a vizsgált tanulók fogalmazásképségeinek fejlettségi szintjéhez. Az értékelési skálák legmagasabb foka nem működik megfelelően. A tanulók nagyon kis arányát jellemzik a legmagasabb értékek, különösen az első bíráló ítéletei szerint. Az átlagos képességszint alatt mindkét bíráló esetén több tanuló helyezkedik el. A szempontrendszer működését jellemzi, hogy a rendszer a vizsgált mintához nem illeszkedik tökéletesen, nem fedi le kellő mértékben a tanulók képességszintjét, ugyanis a skála legmagasabb foka – különösen az első bíráló ítéletei szerint – csak a tanulók kis arányát jellemzi, a minta többsége a 2-es és a 3-as értéket kifejező szintre került. Mindezek alapján a szempontrendszer a vizsgált mintánál jobb fogalmazásképséggel bíró tanulóknál esetén működik jól.



1. Tartalom
2. Feladattartás: szövegtípus
3. Feladattartás: hangnem
4. Szerkezet és kidolgozás
5. Stílus
6. Érthetőség
7. Nyelvhelyesség
8. Helyesírás, központozás
9. Külalak és olvashatóság

2. ábra

Az értékelési szempontrendszer személy-szempont térképe a második bíráló értékítéletei alapján ($X \approx 3$ tanuló)

Az értékelők skálahasználata

Az értékelők skálahasználatának jellemzését az itemek karakterisztikus görbéinek vizsgálata teszi lehetővé. A vízszintes tengelyen a képességszinteket, a függőlegesen a pontszámokhoz tartozás valószínűségét jelenítjük meg. A görbék az értékelési skálák egyes pontjainak karakterisztikus görbéi, annak valószínűségét jelenítik meg, hogy a különböző képességszintek mellett adott szempontból az adott skálapontra sorolják az értékelők a tanulót (Molnár Gy., 2008, 2013; Vigh, 2010).

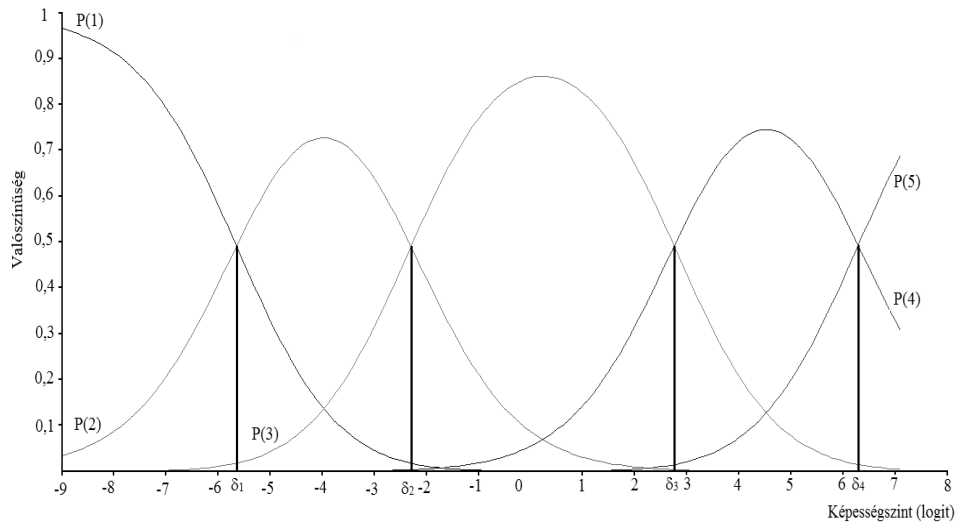
Az értékelők skálahasználatának jellemzőiről további információkat adnak a görbék metszéspontjaiban elhelyezkedő δ -paraméterek. Ezek az értékek az egyes skálapontok közötti átmenetet jelölik; esetünkben azt mutatják meg, hogy a fogalmazásképeség fejlettségének mely logitértékben kifejezett szintjén azonos annak a valószínűsége, hogy két szomszédos skálapont egyikére sorolják a bírálók az adott tanulót. A δ -paraméterek két szempontból teszik lehetővé az értékelési skálák jellemzését. Távoltságuk összehasonlításával megítélhetjük, milyen különbségek vannak az egyes skálapontok megszerzésének lehetőségében, hiszen minél nagyobb két δ -paraméter távolága, annál nagyobb a valószínűsége annak, hogy az értékelők a tanulót az adott skálaponthez rendelik. Másfelől érdemes vizsgálni a δ -paraméterek elrendeződését, vagyis azt, hogy nem cserélődnek-e fel a különböző értékek, ez ugyanis a bíráló skálahasználatának problémáit jeleznék (Molnár Gy., 2008, 2013; Vigh, 2010).

A következőkben csak néhány, jellemző működésű szempont görbét mutatjuk be, illetve jelezzük, melyek azok a szempontok, amelyek hasonlóan viselkednek. A 3. ábrán ideális skálaműködés látható. Amellett, hogy a δ -paraméterek megfelelően emelkednek, legnagyobb valószínűsége annak van, hogy a tanuló a hármaskategóriába kerül, közel azonos valószínűsége van a 2-es és a 4-es értéket megszerzésének. Így működik az értékelési skála az első bíráló ítéletei alapján a feladattartás: szövegtípus mellett, a feladattartás: hangnem és az érthetőség szempontokon; valamint a második bíráló tartalom, szerkezet és kidolgozás, feladattartás: szövegtípus, feladattartás: hangnem, stílus, érthetőség, nyelvhelyesség, illetve külalak és olvashatóság szerint.

Az ideálistól eltérő, az értékelésben problémákat jelző itemkarakterisztikus görbékre ad példát a 4. és az 5. ábra. A 4. ábra alapján legnagyobb valószínűsége a 4-es értéket megszerzésének van, ennél sokkal kisebb a 2-esé, és a 3-asé. A legtöbb tanuló tehát 4-es osztályzatot kap az így viselkedő szempontoknál, és alacsonnyá válik az ennél kisebb értéket megszerzők aránya. Mindkét értékelő így használja például a helyesíráskálát, valamint az első bírálónál így működik a tartalom és a nyelvhelyesség szempont is.

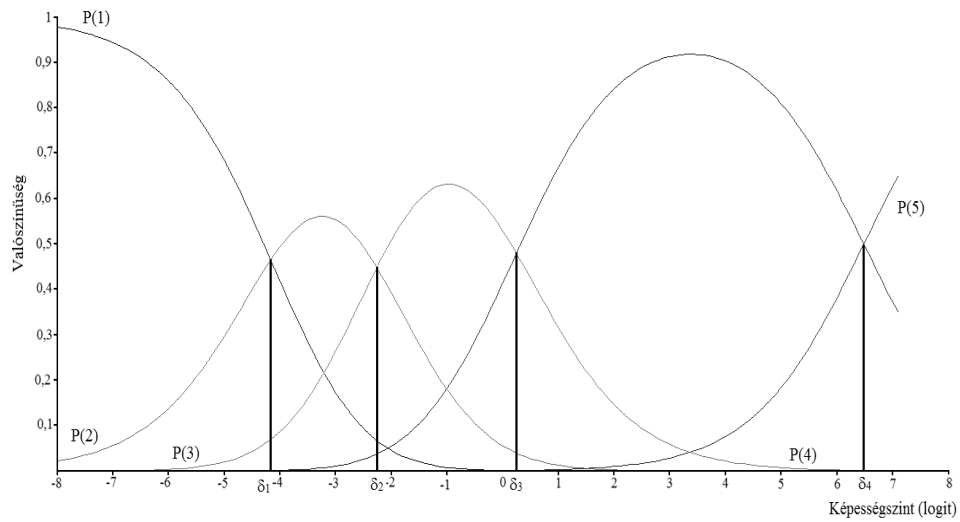
A rosszul működő értékelési skálára adnak példát az 5. ábrán megjelenített itemkarakterisztikus görbék. Megállapítható, hogy az első bíráló értékelésében a külalak és olvashatóság szempont esetén az 1-es skálapont gyakorlatilag nem működik, legnagyobb esélye a 2-esnek, valamivel kisebb a 3-asnak és még kisebb a 4-esnek van.

A fogalmazásértékelés megbízhatósága két független bíráló értéktételeinek elemzése alapján



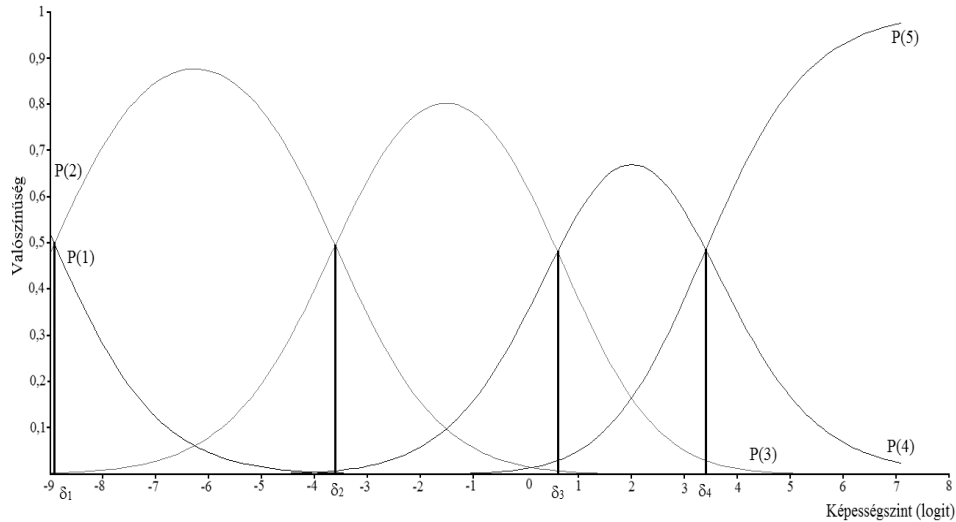
3. ábra

A feladattartás: szövegtípus szempont itemkarakterisztikus görbéi az első bíráló értéktételei alapján



4. ábra

A helyesírás és központozás szempont itemkarakterisztikus görbéi a második bíráló értéktételei alapján



5. ábra

A külalak és olvashatóság szempont itemkarakterisztikus görbéi az első bíráló értéktételei alapján

Összegezve a valószínűségi tesztelmélet eszközeivel végzett vizsgálataink eredményeit: a parciáliskredit-moddellel történő elemzéseink szerint a szigorúsági paraméterek kis eltérése ellenére sem tekinthetjük egyenértékűnek a két bíráló ítéleteit. Az értékelőket eltérő skálahasználat jellemezte, illetve az első bíráló esetén a szempontok modellilleszkedésében is problémákat azonosítottunk.

Összegzés és következtetések

Tanulmányunkban a fogalmazásképeség értékelésének kérdéseit vizsgáltuk. A szakirodalom összegzése alapján a szövegek minősítésének eszközeivel, módszereivel, valamint az értékelés megbízhatóságának problémáival foglalkoztunk. Ezt követően két független bíráló azonos tanulói szövegkorpusz minőségére adott ítéleteit a klasszikus és a valószínűségi tesztelmélet eszközeivel elemezve egy fogalmazásértékelési szempontrendszer működésére vonatkozóan fogalmaztunk meg következtetéseket. Empirikus vizsgálatunk célja az volt, hogy a korábbi, egy bíráló ítéleteivel dolgozó kutatásaink során megbízhatónak ítélt értékelési skálák működését két értékelő által adott osztályzatok további elemzésének alapján vizsgáljuk, a szempontrendszer használhatóságának érvényességére vonatkozóan árnyaltabb kijelentéseket fogalmazzunk meg. Hazai kontextusban kutatásunk újszerűségét az adta, hogy adatainkat a klasszikus tesztelmélet eszköztárába tartozó reliabilitásvizsgálatok, regresszióanalízisek és összefüggés-vizsgálatok mellett a valószí-

núségi tesztelmélet módszereivel történő elemzésnek is alávetettük és ugyanazon tanulói szövegekre vonatkozóan rendelkezünk két egymástól független értéklettel.

Mindkét bíráló esetén szignifikáns összefüggéseket tapasztaltunk az egyes szempontokra adott értékletek között, és a két értékelő azonos szempontokból megállapított minősítései szintén erős, szignifikáns kapcsolatban álltak egymással. Értékelőink szigorúságában nem találtunk nagy különbséget. Ennek ellenére jelentős eltéréseket mutattunk ki a bírálók értékelői teljesítményében. Míg a második értékelő munkáját jellemezve megállapítottuk, hogy az analitikus szempontok a szöveg külső megjelenítését és a helyesírási szabályok ismeretét kifejező osztályzatok kivételével kellő mértékben járultak hozzá az összbenyomás jegyéhez, illetve az említett két szempontot kivéve megfelelő modellilleszkedést mutattak, addig az első bíráló esetén számos problémát regisztráltunk. Láthattuk, hogy az ő értékelésében szignifikáns mértékben csak a tartalom, illetve a szerkezet és kidolgozás osztályzatok magyarázták a globális értéklet varianciáját. Több szempont esetén mutattunk ki az elfogadható intervallumon kívül elhelyezkedő inflatív paramétereket is, melyek a szövegértékelési skálák rossz modellilleszkedését jelzik.

Az értékelők skálahasználatának jellemzésére irányuló vizsgálataink alapján ugyancsak különbségeket találtunk a bírálók munkájában. A személy-szempont térképek alapján a szempontrendszer mintához való illeszkedése egyik bíráló esetén sem volt tökéletes, hiszen a legmagasabb képességszintre alig kerültek diákok. Ugyanakkor a szempontok itemkarakterisztikus görbéi és az egyes skálapontokra kerülés valószínűségét jelző δ -távolságok több esetben különböztek. Emellett az első értékelő munkájában találtunk rosszul működő szempontot is: a külalak és olvashatóság legalacsonyabb skálapontját gyakorlatilag nem használta a bíráló.

Ha az első bíráló munkáját tekintjük, a szempontrendszer működésében problémák mutatkoznak. A második bíráló értékletei alapján mind a klasszikus, mind a valószínűségi tesztelmélet eredményei az értékelési skálák megbízható működését igazolták. Szakirodalmi áttekintésünkben rámutattunk arra, hogy a bírálók személyiségbeli különbségeitől kezdve, a képzettségük eltérésein át számos oka lehet annak, hogy ugyanaz a szöveg a különböző bírálóktól eltérő értékelést kap. Vizsgálatunk során elképzelhető, hogy a szövegértékelés egyik, a szakirodalom elemzésekor is bemutatott problémájával, az egyes értékelési szempontok, illetve a hozzájuk rendelt skálapont-definíciók eltérő értelmezésének jelenségével szembesültünk. Az értékelésekben kimutatott különbségek esetünkben adódhattak a szempontrendszer és az értékelők képzésének elégtelenségéből egyaránt. Így szükségesnek látszik mind az értékelési skálák felülvizsgálata, mind további szövegvizsgálatok esetén az értékelők felkészítésének több részletre kiterjedő lebonyolítása is.

Az alkalmazott értékelési szempontrendszer továbbfejlesztése, illetve a későbbi fogalmazáskutatások módszereinek és eszközeinek megválasztása, kidolgozása szempontjából is fontos kiemelnünk, hogy mind a klasszikus, mind a valószínűségi tesztelmélet eszközeivel végzett elemzéseink rávilágítottak arra, hogy a helyesírás, központosítás, valamint a külalak és olvashatóság osztályzatok viszonylag függetlennek bizonyultak a többi szemponttól (Molnár E. K., 2000). A két bíráló által adott osztályzatok összefüggéseit kutató korrelációs számítások és a regressziós modellek eredményei azt mutatták, hogy ezek azok a szempontok, amelyek legkisebb mértékben korrelálnak a fogalmazás-

képesség egyéb összetevőinek fejlettségét kifejező szempontokon elért teljesítményekkel, illetve a legkevésbé járulnak hozzá az értékelők szövegekről alkotott globális ítéleteinek varianciájához. A helyesírás, központozás, illetve a külalak és olvashatóság szempontok függetlenségét igazolták a parciáliskredit-moddellel lefuttatott számítások is. A szempontok modellilleszkedését jelző infit paraméterek mindkét bíráló esetén az elfogadható intervallum feletti értékeket vettek fel e két skála esetén, ami szintén a helyesírási jellemzők és a szöveg külső megjelenítésének egyéb szempontoktól való függetlenségét, többdimenzionalitását jelzi. Mindezek alapján megfontolandó a helyesírás és a külalak más, a többi szövegjellemzőt vizsgáló skálától független szempontrendszerrel történő értékelése és ezektől történő elválasztása például a *Vidakovich* (1986), illetve *Orosz és Vidakovich* (1988) szempontjai szerint.

További kutatási feladat annak feltárása, milyen egyéb okok húzódnak meg az értékelések különbözősége mögött, befolyásoló tényező-e az értékelők korábbi mérés-értékelési, illetve tanítási tapasztalata. Ennek megállapításához újabb értékelők bevonása, a skálapontok definícióinak közös értelmezése lehet szükséges. A szempontrendszer továbbfejlesztése mind a későbbi fogalmazáskutatások, mind a pedagógusok fogalmazás-értékelési gyakorlatának fejlesztéséhez fontos lehet. Egy kezelhető számú szemponttal, jól értelmezhető skálapontokkal rendelkező, a gyakorlatban megfelelően használható értékelési rendszer kiindulópontul szolgálhat a többszempontú iskolai fogalmazásértékelési kultúra meghonosításához, ezáltal a tanulók fogalmazásképességének fejlesztéséhez is.

Köszönetnyilvánítás

Köszönjük *Molnár Gyöngyvér* és *Vígh Tibor* adatelemzésben nyújtott segítségét. A tanulmányban közölt empirikus eredményeket a X. Pedagógiai Értékelési Konferencián elhangzott előadásban (*Nagy, 2012*) bemutattuk.

A kutatás megvalósítását a TÁMOP 3.1.9-08/1-2009-0001 és a TÁMOP 3.1.9-11/1-2012-0001 támogatta.

Irodalom

- Barkaoui, K. (2007): Rating scale impact on EFL essay marking: A mixed-method study. *Assessing Writing*, **12**. 2. sz. 86–107.
- Barkaoui, K. (2011): Effects of marking method and rater experience on ESL essay scores and rater performance. *Assessment in Education: Principles, Policy and Practice*, **18**. 2. sz. 279–293.
- Beaugrande, R. A. D. (1984): *Text production: Toward a science of text production*. Ablex, Norwood.
- Bereiter, C. (1980): Development in writing. In: Gregg, L. W. és Steinberg, E. R. (szerk.): *Cognitive processes in writing*. L. Erlbaum Associates, Hillsdale. 73–93.

A fogalmazásértékelés megbízhatósága két független bíráló értékítéleteinek elemzése alapján

- Bereiter, C. és Scardamalia, M. (1987a): Knowledge telling and knowledge transforming in written composition. In: Rosenberg, S. (szerk.): *Advances in applied linguistics. Vol. 2. Reading, writing and language learning*. Cambridge University Press, Cambridge. 142–175.
- Bereiter, C. és Scardamalia, M. (1987b): *The psychology of written composition*. Lawrence Erlbaum Associates, Publishers, Hissdale, New Jersey London.
- Beyreli, L. és Ari, G. (2009): The use of analytic rubric in the assessment of writing performance – inter-rater concordance study. *Educational Sciences: Theory and Practice*, **9**. 1. sz. 105–125.
- Chai, C. (2006): Writing plan quality: Relevance to writing scores. *Assessing Writing*, **11**. 198–223.
- Crawford, L. és Smolkowski, K. (2008): When a „sloppy copy” is good enough: Results of a state writing assessment. *Assessing Writing*, **13**. 1. sz. 61–77.
- Dávid Gergely (2008): *Az emelt szintű idegen nyelvi érettségi és az államilag elismert nyelvvizsgák a vizsgázói teljesítmények tükrében. Összegző tanulmány, bővített változat. Nyelvvizsgáztatási Akkreditációs Központ. Budapest.*
- Davis, W. (2005): The effects of grammar testing on the writing quality and reduction of errors in college freshmen’s essays.
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/c1/a7.pdf.
Utolsó letöltés: 2010. április 7.
- Eckes, T. (2008): Rater types in writing performance assessments: a classification approach to rater variability. *Language Testing*, **25**. 2. sz. 155–185.
- Eckes, T. (2005): Examining rater effects in TestDaF writing and speaking performance assessments: a many-facet Rasch analysis. *Language Assessment Quarterly*, **2**. 3. sz. 197–221.
- Engelhard, G. és Myford, C. M. (2003): *Monitoring faculty consultant performance in the Advanced Placement English Literature and Composition Program with a many-faceted Rasch model*. College Entrance Examination Board, New York.
- Engelhard, G. Jr. (1994): Examining rater errors in the assessment of written composition with a many faceted Rasch model. *Journal of Educational Measurement*, **31**. 2. sz. 93–112.
- Engelhard, G. Jr., Gordon, B. és Gabrielson, S. (1991): Writing task and the quality of student writing: Evidence from a statewide assessment of writing. Előadás, Annual Meeting of the American Educational Research Association, Chicago. 1991. április 3–7.
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/22/ef/b7.pdf.
Utolsó letöltés: 2010. április 7.
- Eysenck, M. W. és Keane, M. T. (1997): *Kognitív pszichológia*. Nemzeti Tankönyvkiadó, Budapest.
- Flower, L. és Hayes, J. (1980): The dynamics of composing: making plans and juggling constraints. In: Gregg, L. W. és Steinberg, E. R. (szerk.): *Cognitive processes in writing*. L. Erlbaum Associates, Hillsdale. 31–50.
- Gearhart, M., Herman, J. L., Novak, J. R. és Wolf, S. A. (1995): Toward the instructional utility of large-scale writing assessment: Validation of a new narrative rubric. *Assessing Writing*, **2**. 2. sz. 207–242.
- Gelati, C. és Boscolo, P. (2009): Improving the quality of primary school children’s narration of personal events. An intervention study on the use of evaluation strategies. *L1 – Educational Studies in Language and Literature*, **9**. 3. sz. 1–28.
- Gorman, T. P., Purves, A. C. és Degenhart, R. E. (1988, szerk.): *The IEA study of written composition I. The international writing tasks*. Pergamon Press, Oxford.
- Griffin, P. és Anh, P. N. (2005): Assessment of creative writing in Vietnamese primary education. *Asia Pacific Education Review*, **6**. 1. sz. 72–86.
- Gyagenda, I. S. és Engelhard, G. (1998): Applying the Rasch model to explore rater influences on the assessed quality of students’ writing ability. Előadás, Annual Meeting of the American Educational Research

- Association, San Diego. 1998. április 13–17. <http://www.eric.ed.gov/PDFS/ED422350.pdf>. Utolsó letöltés: 2012. január 20.
- Hayes, J. és Flower, L. (1980): Identifying the organization of writing processes. In: Gregg, L. W. és Steinberg, E. R. (szerk.): *Cognitive processes in writing*. L. Erlbaum Associates, Hillsdale. 3–30.
- Hayes, J. R. (1996): A new framework for understanding cognition and affect in writing. In: Levy, C. M. és Ransdell, S. (szerk.): *The science of writing: Theories, methods, individual differences and applications*. Lawrence Erlbaum Associates, Mahwah, N. J. 1–27.
- Hidi, S. és Boscolo, P. (2007, szerk.): *Writing and motivation*. Elsevier, Amesterdam.
- Hillocks, G. (1986): *Research on written composition*. Clearinghouse on Reading and Communication Skills, Urbana, Illinois.
- Horváth Zsuzsanna (1998): *Anyanyelvi tudástérkép. Középszkolai tantárgyi feladatbankok III.* Országos Közoktatási Intézet, Budapest.
- Isonio, S. (1991): Judgments of placement writingsampels at golden west collage: an evaluation of inter-rater reliability. Kézirat. Golden West College, Huntington Beach. <http://www.eric.ed.gov/PDFS/ED345784.pdf>. Utolsó letöltés: 2012. február 3.
- Kádárné Fülöp Judit (1990): *Hogyan írnak a tizenévesek? – Az IEA-fogalmazásvizsgálat Magyarországon*. Akadémiai Kiadó, Budapest.
- Kontra József (2009): A parciális kredit modell egy alkalmazása. In: Psenáková Ildikó, Mező Ferenc és Viczayová Ildikó (szerk.): *Képzés és gyakorlat II.* Konstantin Filozófus Egyetem, Nyitra. 99–108.
- MacArthur, C. A., Graham, S. és Fitzgearld, J. (2008, szerk.): *Handbook of writing research*. The Guilford Press, New York.
- Mäki, H. (2002): *Elements of spelling and composition studies on predicting and supporting writing skills in primary grades*. Turun Ylioposto, Turku.
- Molnár Edit Katalin (1996): A kognitív pszichológia három fogalmazás-modellje. *Magyar Pedagógia*, **96**. 2. sz. 139–156.
- Molnár Edit Katalin (2000): A fogalmazási képesség fejlődésének mérése. *Iskolakultúra*, **10**. 8. sz. 49–59.
- Molnár Edit Katalin (2002): Az írásbeli szövegalkotás. In: Csapó Benő (szerk.): *Iskolai műveltség*. Osiris, Budapest. 193–216.
- Molnár Edit Katalin (2003): Az írásbeli szövegalkotás fejlődése: vizsgálatok 10–17 éves tanulók körében. Doktori disszertáció. ELTE BTK Neveléstudományi Doktori Iskola, Budapest.
- Molnár Edit Katalin (2009): Az írásbeli szövegalkotás funkciója és hatékonysága magyar egyetemista diákok dolgozatainak szövegeiben. *Anyanyelv-pedagógia*, **2**. 1. sz. <http://www.anyanyelv-pedagogia.hu/cikkek.php?id=138>. Utolsó letöltés: 2013. szeptember 6.
- Molnár Gyöngyvér (2003): Az ismeretek alkalmazásának vizsgálata modern tesztelméleti (IRT) eszközökkel. *Magyar Pedagógia*, **103**. 4. sz. 423–446.
- Molnár Gyöngyvér (2005): Az objektív mérés 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 a parciális kredit modell. *Iskolakultúra*, **18**. 1–2. sz. 66–77.
- Molnár Gyöngyvér (2013): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában. Alapvető elemzések a társadalomtudományi kutatásokban*. Gondolat Kiadó, Budapest.
- Molnár Gyöngyvér és Józsa Krisztián (2006): Az olvasási képesség értékelésének tesztelméleti megközelítései. In: Józsa Krisztián (szerk.): *Az olvasási képesség fejlődése és fejlesztése*. Dinasztia Tankönyvkiadó, Budapest. 155–174.

A fogalmazásértékelés megbízhatósága két független bíráló értéktételeinek elemzése alapján

- Nagy József (1996): *Nevelési kézikönyv: Személyiségfejlesztő pedagógiai programok készítéséhez*. Mozaik Oktatási Stúdió, Szeged.
- Nagy József (2002): *XXI. század és nevelés*. Osiris Kiadó, Budapest.
- Nagy Zsuzsanna (2009): 17 éves tanulók szövegalkotási képessége és szövegekre vonatkozó ítéletei. *Iskolakultúra*, **19**. 11. sz. 19–31.
- Nagy Zsuzsanna (2010): A fogalmazásképesség értékelésének lehetősége egy próbamérés eredményei alapján. Előadás, Országos Neveléstudományi Konferencia, Budapest. 2010. november 4–6. 186.
- Nagy Zsuzsanna (2011): Magyart tanító pedagógusok fogalmazástanítással kapcsolatos meggyőződései és oktatási módszerei. Előadás, XI. Országos Neveléstudományi Konferencia, Budapest. 2011. november 3–5. 171.
- Nagy Zsuzsanna (2012): Fogalmazásértékelési szempontrendszer működésének vizsgálata két független bíráló ítéletei alapján. Előadás, X. Pedagógiai Értékelési Konferencia, Szeged. 2012. április 26–28. 102.
- Orosz Sándor (1972): *A fogalmazástechnika mérésmetodikai problémája és országos színvonala*. Tankönyvkiadó, Budapest.
- Orosz Sándor és Vidákovich Tibor (1988): Anyanyelvi helyesírásvizsgálatok. Eredménymérések néhány módszertani kérdése. *Pedagógiai Technológia*, **2**. sz. 30–40.
- Park, T. (2004): An investigation of an ESL placement test of writing using many-facet Rasch measurement. *Working Papers in TESOL and Applied Linguistics*, **4**. 1. sz. 1–21. o.
- Pintér Henrett (2009): Erkölcsi gondolkodás 9–10 évesek írásbeli szövegalkotásában. *Iskolakultúra*, **19**. 10. sz. 109–116.
- Popp, S. E. O., Ryan, J. M., Thompson, M. S. és Behrens, J. T. (2003): Operationalizing the rubric: The effect of benchmark selection on the assessed quality of writing. Előadás, Annual Meeting of the American Educational Research Association, Chicago. 2003. április 1–25.
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/1b/7c/cf.pdf.
Utolsó letöltés: 2010. április 7.
- Purves, A. (1992): Reflections on research and assessment in written composition. *Research in the Teaching of English*, **26**. 1. sz. 108–22.
- Ransdell, S. E. és Levy, C. M. (1996). Working memory constraints on writing performance. In: Levy, C. M. és Ransdell, S. E. (szerk.): *The science of writing*. Lawrence Erlbaum Associates, Mahwah, New Jersey. 93–101.
- Ransdell, S., Levy, C. M. és Kellogg, R. (2002): The structure of writing processes as revealed by secondary task demands. *L1 – Educational studies in Language and Literature*, **2**. 2. sz. 141–163.
- Schoonen, R. (2005): Generalizability of writing scores: An application of structural equation modeling. *Language Testing*, **22**. 1. sz. 1–30.
- Segev-Miller, R. (2004): Writing from sources: The effect of explicit instruction on college students' processes and products. *L1 – Educational Studies in Language and Literature*, **4**. sz. 5–33.
- Sudweeks, R. R., Reeve, S. és Bradshaw, W. S. (2004): A comparison of generalizability theory and many-facet Rasch measurement in an analysis of college sophomore writing. *Assessing Writing*, **9**. 3. sz. 239–261.
- Sugita, Y. (2009): The development and implementation of task-based writing performance assessment for Japanese learners of English. *Journal of Pan-Pacific Association of Applied Linguistics*, **13**. 2. sz. 77–103.
- Szilassy Eszter (2012): Az írás és fogalmazásjavítás stratégiái. *Anyanyelv-pedagógia*, **4**. 1. sz.
<http://www.anyanyelv-pedagogia.hu/cikkek.php?id=357>. Utolsó letöltés: 2013. szeptember 6.
- Takala, S. (1988): Origins of the international study of writing. In: Gorman, T. P., Purves, A. C. és Degenhart, R. E. (szerk.): *The IEA study of written composition I.: The international writing tasks and scoring scales*. Pergamon, Oxford. 3–14.

- Torrance, M., van Waes, L. és Galbraith, D. (2007, szerk.): *Writing and cognition: research and application*. Elsevier, Amsterdam.
- Vidákovich Tibor (1990): *Diagnosztikus pedagógiai értékelés*. Akadémiai Kiadó, Budapest.
- Vidákovich Tibor (1986): Az íráskészség vizsgálatának néhány mérésmetodikai problémája. *Acta Universitatis Szegediensis de Attila József Nominatae Sectio Paedagogica et Psychologica*, **28**. 117–138.
- Vígh Tibor (2008): Egy IRT-alapú nyelvi feladatban létrehozásának módszertani kérdései. A német érettségi vizsgafeladatok elemzésének eredményei. *Magyar Pedagógia*, **108**. 1. sz. 29–51.
- Vígh Tibor (2010): Az idegen nyelvi érettségi működés és hatása a tanulói teljesítmények és a tanári nézetek tükrében. PhD-értekezés. Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola, Szeged.
- Weigle, S. C. (1998): Using FACETS to model rater training effects. *Language Testing*, **15**. 2. sz. 263–287.
- Whithaus, C., Harrison, S. B. és Midyette, J. (2008): Keyboarding compared with handwriting on a high-stakes writing assessment: Student choice of composing medium, raters' perceptions, and text quality. *Assessing Writing*, **13**. 5. sz. 4–25.
- Wiseman, C. S. (2012): Rater effects: ego engagement in rater decision-making *Assessing Writing*, **17**. 3. sz. 150–173.
- Wu, M., Adams, R. J. és Wilson, M. R. (1998): *ACER ConQuest. Generalised Item Response Modelling Software*. ACER Press, Australia.
- Zhang, L. és Vukelich, C. (1998): Prewriting activities and gender: Influences of the writing quality of male and female students. Előadás, Annual Meeting of the American Educational Research Association, San Diego. 1998. április 13–17.
http://www.eric.ed.gov/ERICDocs/data/ericdocs2sql/content_storage_01/0000019b/80/15/b4/5a.pdf.
Utolsó letöltés: 2010. április 7.

ABSTRACT

ZSUZSANNA NAGY: THE RELIABILITY OF WRITTEN COMPOSITION ASSESSMENT BASED ON AN ANALYSIS OF THE EVALUATIONS OF TWO INDEPENDENT RATERS

The purpose of the present study is to examine the operation of our scoring system, which was developed to assess the quality of schoolchildren's texts and to analyse two independent raters' evaluations. Rater performance was analysed using classical test theory and item response theory. The sample included 429 Hungarian children in Year 8. They were asked to produce a narrative text, and their compositions were scored by two independent raters using our scoring system, which comprises one holistic criterion and nine analytic ones (content, genre, tone, organization and structure, style, readability, lexical and grammatical conventions, spelling and orthographic conventions, handwriting and neatness). The analyses showed high reliability (Cronbach- $\alpha=0.95$) for both raters. There are strong and significant correlations between the ratings ($r=.85-.93$, $p<.01$) and small differences (.56 logit) between the severity parameters of the two raters. The partial credit model analyses revealed differing uses of the scales by the raters. The scores given by the first rater on most scales show a bad model fit. The δ parameters of the scales' characteristic curves indicate that the two raters used the scales differently. Results call attention to the problem of defining scales for written composition scoring systems. The results point out that a precise definition of the scales was unable to guarantee objective and consistent assessment as the two raters still interpreted the scales differently. The findings indicate the need to re-examine the scoring system used and to provide training for raters.

Magyar Pedagógia, **113**. Number 3. 153–179. (2013)

Levelezési cím / Address for correspondence: Nagy Zsuzsanna, SZTE Oktatásméleti Kutatócsoport, H-6722 Szeged, Petőfi S. sgt. 30–34.

SZÁMÍTÓGÉP ALAPÚ ADAPTÍV ÉS RÖGZÍTETT FORMÁTUMÚ TESZTELÉS ÖSSZEHASONLÍTÓ HATÉKONYSÁGVIZSGÁLATA

Magyar Andrea* és Molnár Gyöngyvér**

* SZTE Neveléstudományi Doktori Iskola

** SZTE Neveléstudományi Intézet, MTA-SZTE Képességfejlesztés Kutatócsoport

Az utóbbi két évtizedben a számítógépek rohamos elterjedésével a papír-ceruza alapú tesztek helyét – mind hazai, mind nemzetközi szinten – fokozatosan felváltották, illetve felváltják a számítógép alapú mérések (Molnár, 2011). A tradicionális papír-ceruza teszteléssel szemben számos új lehetőség adódik alkalmazásukkal, például az innovatív, multimédiás elemeket is tartalmazó dinamikusan változó itemek megjelenítése (Wang, 2011; Greiff, Wüstenberg és Funke, 2012; Greiff, 2012) vagy a személyre szabott, adaptív tesztelés megvalósítása (Eggen és Straetmans, 2009).

A hagyományos, többnyire rögzített formátumú tesztek használata esetén minden tesztelt személy ugyanazon sorrendben ugyanazon feladatokat kapja a tesztelés során, függetlenül képességszintjétől és teljesítményétől. Azonban ezzel a technikával csak egy viszonylag szűk képességtartomány mérhető a szükséges pontossággal. Ha a tesztet szélesebb képességtartomány mérésére tesszük alkalmassá, azaz a feladatok nehézségi indexei széles skálán mozognak, akkor minden adatfelvételben részt vevő személy számára csak a teszt néhány feladata jelent kihívást, melyek nehézségi szintjei közel állnak a tesztet megoldó személy képességszintjéhez. A teszt többi feladata, esetlegesen a teszt nagyobb része jóval kisebb mértékben járul hozzá a személy képesség- vagy tudásszintjének pontos meghatározásához. Azok túl könnyűek vagy túl nehezek, azaz vagy nem jelentenek kihívást, vagy frusztrálóak a tesztelt személy számára, ami jelentős mértékben csökkentheti a teszteléssel kapcsolatos érdeklődését, attitűdjét, motivációját (Csapó, Molnár és R. Tóth, 2008).

Az adaptív tesztelési technika alkalmazása során a teszt feladatai nem előre meghatározott fix sorrendben követik egymást, hanem azokat egy feladatbankból választják ki a tesztmegoldó korábbi feladatokon nyújtott teljesítménye alapján. Például feladatszintű adaptivitás esetén (l. később) amennyiben a tanuló meg tudja oldani a teszt egy feladatát, következőleg egy nehezebbet kap, ha nem, akkor könnyebbet. Ezen algoritmus alkalmazásával a tesztelés végén minden tanulóhoz hozzárendelhető egy képességszint, melynél könnyebb feladatokat nagyobb valószínűség mellett old meg helyesen, mint helytelenül. Erre az alapfeltevésre épülnek a tudástér-elméletek is (knowledge theory; l. Tóth, 2005). Míután adaptív tesztelés során a tanulók összességében különböző feladatokból, itemekből összeállított tesztekkel oldanak meg, eredményük a hagyományos, klasszikus tesztelés

mélet eljárásaival nem összehasonlítható. A feladatok közös sajátossága, hogy azok azonos feladatbankból származnak, ami megteremti egyrészt a valószínűségi tesztmodellek alkalmazásának lehetőségét, másrészt a diákok teljesítményének közös képességskálán történő kifejezhetőségét (erről bővebben l. *Molnár*, 2013).

Ez a típusú feladatadás és tesztszűzseállítás a hagyományos, mindenki számára azonos itemeket azonos sorrendben tartalmazó, rögzített formátumú tesztekkel szemben a teljesítmények sokkal finomabb mérését teszi lehetővé (*Linacre*, 2000). Jelentős mértékben megnő a tesztelés során kinyerhető itemekre és személyekre vonatkozó információ nagysága (l. *Molnár*, 2013). Elhanyagolhatóvá válik annak valószínűsége, hogy a tesztelt személyek ugyanazon feladatokat ugyanabban a sorrendben kapják, azaz növekszik a tesztelés biztonsága (*Wainer*, 2000). Mindez új lehetőségeket terem a mérés-értékelés területén. Ha nem törekszünk több információ kinyerésére, azaz megelégszünk a hagyományos tesztelés során elérhető pontossággal, akkor a kiközvetített feladatok száma, vagyis a teszt hossza (*Thompson*, 2007), ezzel párhuzamosan a teszt megoldásához szükséges idő is jelentős mértékben rövidül, utóbbi átlagosan felére csökken (*Frey és Seitz*, 2009, 2011).

Az adaptivitás mértékétől függően különböző típusú adaptív teszteket különböztünk meg (*Al-A'ali*, 2007; *Magyar*, 2012). Feladatsztintú adaptivitás esetén teljes mértékben biztosított, hogy ha a tesztelt személy helytelenül/helyesen oldja meg a teszt egyik feladatát, akkor a teszt következő feladata egy könnyebb/nehezebb feladat lesz. Azonban a feladatsztintú adaptív tesztelés egyik fő problémája, hogy a feladatok paraméterei annak függvényében változnak, milyen feladatok veszik körül az adott feladatot, illetve, az a teszt melyik (elején, közepén, végén) részén helyezkedik el (*Molnár*, 2013). Ezen probléma megoldását kínálja a részteszt szintű, többszakaszos adaptív tesztelés.

A többszakaszos adaptív tesztelés egyesíti a rögzített és az item alapú adaptív tesztek jó tulajdonságait (*Jodoin*, 2006). Mindamelltt, hogy a teszt nehézségi szintjét a tanuló képességsztintjéhez igazítja, a résztesztek szintjén lehetőséget ad a kérdések sorrendjének előzetes meghatározására (*Amstrong*, 2004; *Molnár*, 2013). A feladatonkénti adaptív tesztelés során megismert eljárás előre meghatározott, különböző nehézségű résztesztekkel valósul meg. A többszakaszos adaptív tesztek legalább két szakaszból állnak, egy szakaszon belül legalább két vagy három különböző nehézségi szintű rögzített formátumú résztesztet tartalmaznak. Az alkalmazott algoritmus szerint a tesztelt személy képességsztintje ebben az esetben nem feladatonként, hanem résztesztenként becsülhető meg. Ha a tesztelt személy egy előre meghatározott teljesítmény alatt/felett teljesít az adott részteszten, akkor a teszt következő résztesztjének feladata egy könnyebb/nehezebb részteszt (*Zenisky, Hambleton és Luecht*, 2010), egy számára a korábbinál nagyobb diagnosztikus erővel bíró részteszt lesz.

A számos előnyös tulajdonság a hagyományos rögzített formátumú tesztekkel szemben vonzóvá teszi az adaptív tesztelés alkalmazását. Ennek hatására a legjelentősebb nemzetközi oktatási vonatkozású projektekben, kutatásokban is fokozatosan előtérbe kerül. Mind az európai OECD PISA-, mint az amerikai *No Child Left Behind*- (NCLB) kutatások kapcsán is felmerült az igény alkalmazására.

Frey, Seitz és Kröhne (2011) szimulációs kísérletben vizsgálták az adaptív tesztelés bevezetésének hatását, lehetőségeit a PISA-felmérések vonatkozásában. A kutatás adat-

bázisát 14 624 15 éves tanuló korábbi (2000-es, 2003-as és 2006-os) PISA-méréseken elért válasza adta. Az elemzést összesen 348 PISA-ítemen végezték el. A szimulációban az itemszintű elemzések helyett részteszt szintű elemzéseket végeztek. Eredményeik szerint a mérés hatékonysága (mérési precizitás/prezentált itemek száma) 74%-kal nőtt. Ugyanazon pontosság biztosítása esetén a szükséges itemszám a korábbi rögzített formátumú tesztek alkalmazása során szükséges 55-ről 26-ra csökkent, és a tesztelés időtartama 120 percről 57 percre csökkent. Ezen eredmények alapján a PISA szakértői csoportja a többszakaszos adaptív tesztek részleges bevezetését javasolta a 2015-ös méréstől kezdődően (OECD, 2012).

Kingsbury (2004) az NCLB-törvény előírásainak megfelelő teszt kidolgozásával kapcsolatban végzett vizsgálatot adaptív tesztekre vonatkozóan. A szimulációs kísérlet 4. és 8. évfolyamos tanulók számára készült matematikai és szövegértési rögzített formátumú tesztekhez hasonlított össze adaptív verziójukkal. A kutatás eredményei szerint az adaptív teszt mindkét szinten több információt, vagyis pontosabb mérési eredményt szolgáltatott, mint a hagyományos verzió.

Azonban egy jól működő feladat- vagy részteszt szintű adaptív rendszer kidolgozása bonyolult és összetett feladat, miközben a rögzített formátumú tesztelésről az adaptív tesztelésre való átállás számos kérdést is felvet. A tanulmányban bemutatott kutatásnak négy célja volt: (1) egy 5–8. évfolyamos tanulók induktív gondolkodás-fejlettségi szintjének meghatározására – többszakaszos adaptív tesztelés használatával – alkalmas itembank összeállítása, (2) a diákok hagyományos, rögzített formátumú teszten és részteszt szintű, adaptív teszten elért teljesítményeinek összehasonlítása, (3) az adaptív tesztelés során kiosztott itemek, illetve résztesztek nehézségi szintjének, ennek változásmintázatának jellemzése, valamint (4) a rögzített és az adaptív tesztelés során kinyert információ és a mérési hiba nagyságának összehasonlítása képességszint szerinti bontásban.

Módszerek

Minta

Az adatfelvétel 2012 őszén 158 5–8. évfolyamos diák részvételével zajlott. A diákok 45%-a volt fiú. A minta évfolyam és nem szerinti eloszlását az 1. táblázat tartalmazza.

1. táblázat. A minta évfolyam és nem szerinti eloszlása

Évfolyam	N (fő)	Nemek aránya
5.	22	1,64
6.	44	1,52
7.	51	1,59
8.	41	1,49

Megjegyzés: fiú: 1, lány: 2.

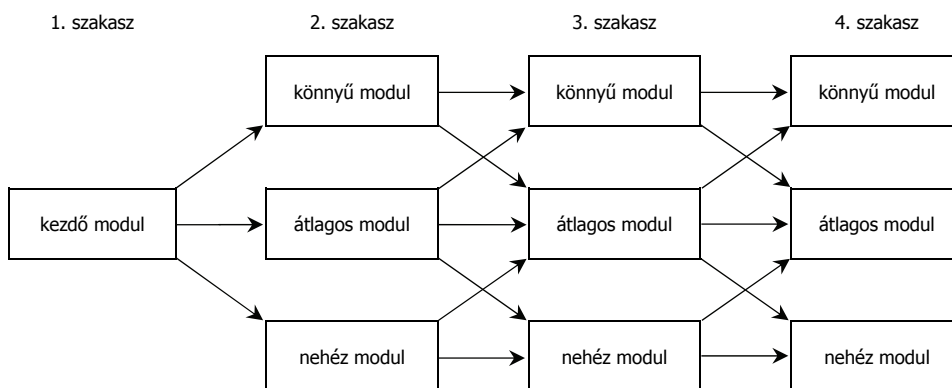
Mérőeszköz

Az induktív gondolkodás fejlettségét korábbi kutatásokban már gyakran alkalmazott itemek (l. pl. *Molnár és Csapó, 2012*) felhasználásával mértük. A több teszt feladatait tartalmazó, a korábbi kutatási eredmények alapján kalibrált itembank 96 verbális és non-verbális itemet tartalmaz.

Adatfelvétel és eljárások

Az adatfelvétel első fázisában minden osztályt véletlenszerűen kettéosztottunk. A diákok egyik része egy 28 itemből álló ($n=79$), rögzített formátumú induktív gondolkodás-tesztet oldott meg, a másik fele egy szintén 28 itemes, négyszakaszos, 1-3-3-3 szerkezetű, többszintű adaptív tesztet. Két hét elteltével, az adatfelvétel második fázisában, a korábban rögzített formátumú tesztet megoldó tanulók adaptív tesztet, az adaptív tesztet megoldó tanulók rögzített formátumú tesztet kaptak. A diákok számára rendelkezésre álló idő mindkét esetben 45 perc volt. Az itemek nehézségi paraméterei $-4,3$ és $+4,3$ logitegység között mozogtak. A rögzített formátumú teszt feladatai széles képességtartomány vizsgálatát célozták meg, azaz a teszt különböző nehézségű feladatokból állt. A tanulók képességszintjének megállapítása a Rasch-modell segítségével történt, majd a logitegységben adott értékeket egy 500 pontos átlagú és 100 pontos szórású skálára transzformáltuk (az eljárásról részletesebben l. *Molnár, 2013*).

A többszakaszos adaptív teszt kezdő résztesztje (1. szakasz) 10, szélesebb itemnehézségi skálán mozgó itemet tartalmazott. A 2–4. szakasz résztesztjei 6-6 itemből álltak, és minden szakaszon belül három különböző nehézségi szintű résztesztet tartalmaztak. Összesen 10 különböző részteszt (modul) kialakítása történt, melyből 17 különböző tesztváltozat összeállítására volt lehetőség. A részteszt (modulok) egymáshoz való viszonyát és az egész tesztelés során elfoglalt helyét az 1. ábra szemlélteti.



1. ábra
A többszakaszos adaptív teszt szerkezete

A modulok közötti elágazási szabály meghatározása az NC-módszer (*Number Correct*) segítségével történt (*Zenisky és Hambleton, 2004*). Minden feladatban a helyes válaszáért 1, a helytelen vagy hiányzó válaszáért 0 pontot adott a rendszer. Az első és a második modul közötti elágazásnál az osztópontot a kezdő modul tesztkarakterisztikus görbéje (erről részletesebben l. *Molnár, 2013*) alapján számítottuk ki a korábbi adatfelvételek eredményeire alapozva, így a 4-nél kevesebb pontot elért tanulók a könnyű modult kapták, a 7-nél több pontot elérik a nehezet (457 és 543 képességpontok által meghatározott értékek). A második és a harmadik elágazás esetén is hasonló módon határoztuk meg az osztópontokat: a könnyű modulnál 0–4 pont elérésekor a könnyű modul felé, 5–6 pont esetén a közepesen nehéz modul felé ágazott el a teszt. A legnehezebb részteszten 0–2 pontot elérő diákok a tesztelés következő szakaszában a közepes nehézségű modult kapták, míg a 3–6 pontot teljesítők maradtak a legnehezebb itemeket tartalmazó tartományban. A közepes (átlagos) nehézségű résztesztet megoldók három irányban léphetek tovább. A 0–2 pontot elérik a legkönnyebb, a 3–4 pontot teljesítők a közepes nehézségű, az 5–6 pontot kapott diákok a legnehezebb feladatokat tartalmazó részteszt felé léptek tovább. Ezzel a módszerrel, előzetes hipotézisünk szerint, a negyedik szakasz végére három egyenlő részre osztottuk a diákokat képességszintjük szerint.

A rögzített és az adaptív teszten elért teljesítmények összehasonlíthatóságát horgonyitemek biztosították. A bevezető tesztben 10, a többi modulban 2-2 horgonyitem szerepelt. Ezáltal bármelyik útvonalon ment végig a tanuló, 16 horgonyitem biztosította az eredmények közös képességskálán való megjeleníthetőségét. A tesztekben kinyerhető információ nagyságát a tesztinformációs görbékkel jellemeztük (erről részletesebben l. *Molnár, 2013*), ami a tesztből kinyert információ nagyságát a tesztet megoldó tanulók átlagos képességszintje és az itemek nehézségi szintje közötti különbségek segítségével jellemzi. A kinyert információ nagyságát akkor tekintettük maximálisnak, ha a feladatok nehézségi szintje és az azokat megoldó diákok képességszintje azonos. Minél távolabb van egymástól ez a két érték, annál kisebb a tesztelés során kinyert információ nagysága.

Az empirikus vizsgálat eredményei

A tesztek reliabilitása

Az eredmények kiterjeszhetőségének, általánosíthatóságának körét első szinten jól jellemzi a teszt reliabilitásmutatójának értéke, aminek meghatározására rögzített formátumú teszt esetén a Cronbach- α -t, adaptív teszt esetén ennek kiterjesztését, a személyszeparációs reliabilitásmutatót használtuk. Az adaptív teszt WLE (*Weighted Likelihood Estimate*) személyszeparációs reliabilitásmutatója 0,85, ami magasabb, mint a rögzített formátumú teszt megbízhatósági mutatója (Cronbach- $\alpha=0,83$). A reliabilitásmutatók alapján megállapítható, hogy a kidolgozott itembank megbízhatóságát tekintve alkalmas 5–8. évfolyamos diákok induktív gondolkodásának, e gondolkodás fejlettségi szintjének meghatározására.

A diákok rögzített formátumú teszten mutatott teljesítménye (átlag=500, szórás=100) és az adaptív teszt alapján számolt képességszintje (átlag=489, szórás=100) erősen korrelált egymással ($r=0,82$, $p<0,01$). A két tesztkörnyezetben egymástól függetlenül meghatározott, azonos diákra vonatkozó képességszintek átlagosan azonosnak tekinthetők, ugyanakkor a korrelációs együttható nagysága eltérésekre is utal.

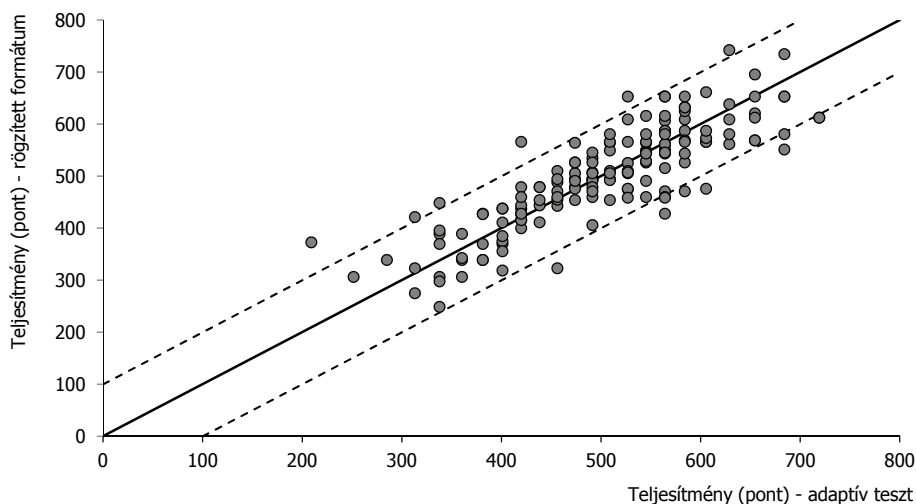
A becsült képességszintek összehasonlítása évfolyamonként és személyenként

A diákok rögzített, illetve adaptív tesztkörnyezetben mutatott teljesítményének alapstatisztikai mutatóit évfolyamonkénti bontásban a 2. táblázat mutatja. Az átlagosan legmagasabb képességszintű, 8. évfolyamos diákok esetén volt szignifikáns különbség a rögzített, illetve adaptív tesztkörnyezetben becsült képességszint között. Az 500 pont körül, azaz átlagosan teljesítő diákok között nem, miután a rögzített formátumú teszt is, hagyományos módon, elsősorban az ő képességszint-mérésüket célozza.

2. táblázat. Rögzített, illetve adaptív tesztkörnyezetben mutatott teljesítmények átlaga és szórása évfolyamonkénti bontásban

Évfolyam	N	Rögzített formátumú teszt		Adaptív teszt		t	p
		átlag	szórás	átlag	szórás		
5.	22	439	90	432	95	-0,55	0,59
6.	44	510	94	500	96	-1,35	0,18
7.	51	498	106	491	99	-0,76	0,45
8.	41	524	93	547	86	2,60	0,01

A 2. ábra a két tesztkörnyezetben nyújtott teljesítmények diákszintű összehasonlítását ábrázolja. Ha a diák képességszintje tesztkörnyezettől függetlenül számszerűen ugyanannak bizonyult, akkor a diákot reprezentáló alakzat a folytonos vonalon helyezkedik el. Amennyiben megállapított képességszintje nem különbözik egymástól szignifikánsan rögzített és adaptív környezetben, az őt reprezentáló jel a szaggatott vonalakon belül található. A szaggatott vonalak által képzett sávon kívül elhelyezkedő diákok számára a rögzített vagy az adaptív tesztkörnyezet bizonyult kedvezőbbnek. Előfordulásuk elenyésző a mintában, azaz különböző mérési hiba alkalmazása mellett, de szignifikanciaszinten belüli, közel azonos képességszint-beclést végeztünk rögzített, illetve adaptív teszt-kiosztással.



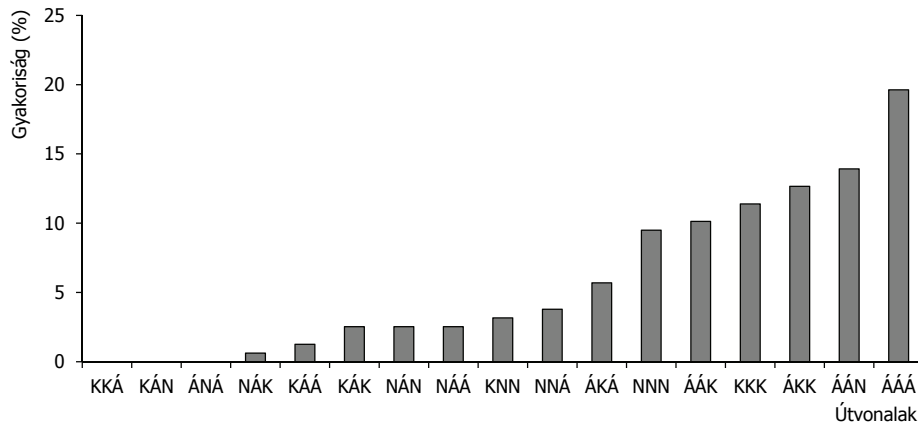
2. ábra

A rögzített formátumú és az adaptív teszten nyújtott teljesítmények összehasonlítása diákonkénti bontásban

A rögzített és az adaptív tesztelés során kiosztott résztesztek nehézségi szintjének változásmintázata

Az adatfelvétel során a többszakaszos adaptív teszt esetében a négy szakaszból összeállítható 17 különböző teszt közül 14-et osztottunk ki (3. ábra). Az esetek ötödében a részteszteken nyújtott teljesítmények alapján kizárólagosan az átlagos nehézségi szintű feladatokból álló tesztet közvetítettük ki. 11%-ban a kizárólagosan könnyű és közel 10%-ban a kizárólagosan nehéz résztesztekből álló, nehézségi szint tekintetében homogen teszteket vettük fel. Mindezek alapján megállapítható, hogy a diákok 40%-a a kezdő részteszten nyújtott teljesítménye alapján egyértelműen besorolható volt a három képességsáv egyikébe. Egyetlen diák esetében fordult elő, hogy két nehézségi szintet is ugrott a tesztelés folyamán. A kezdő modul után megállapított képességszintje a legmagasabb képességtartományba sorolta őt, ugyanakkor a tesztelés végére átkerült az átlagnál alacsonyabban teljesítő diákok csoportjába.

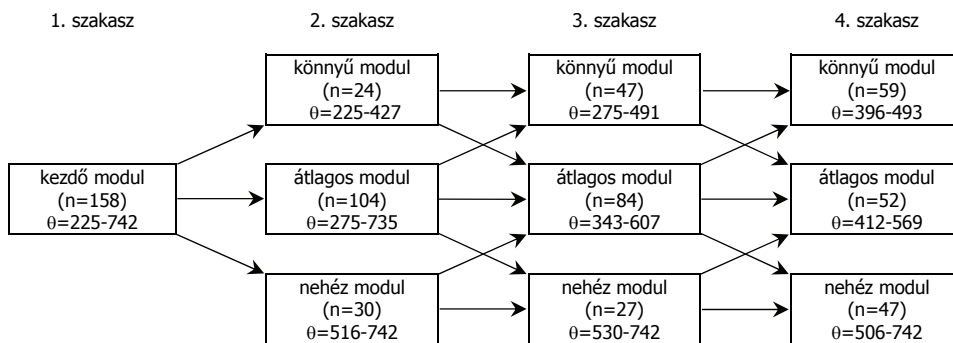
A diákok kétharmadának viselkedését jól jellemzi a hat leggyakoribb útvonal, melyek között a három azonos nehézségi szintű modulból álló tesztek mellett szerepel két olyan útvonal, ahol a két átlagos nehézségi szintű részteszt után a teljesítmények alapján a záró modulon a nehéz, illetve a könnyű részteszt irányába ágazott el a rendszer. A négy szakaszból álló adaptív tesztelési modell előnye a háromszakaszos modellhez képest az átlagos képességszinthez közeli, ugyanakkor azt vagy nem elérő, vagy kicsit túlteljesítő tanulók pontosabb képességszint-meghatározásában mutatkozott meg.



3. ábra

Az adaptív tesztsziszteren belül a második, harmadik és negyedik szakaszban kiosztott útvonalak gyakorisága (K: könnyű, Á: átlagos, N: nehéz modul)

A szakaszokon belül a tanulók modulonkénti eloszlását képességi szint és gyakoriság szerinti bontásban a 4. ábra szemlélteti. A kezdő modulon mutatott teljesítmény alapján a tanulók háromötöd része a teszt második szakaszában közepes nehézségű résztesztet kapott, majd a teszt harmadik szakaszában mutatott teljesítmények alapján a teszt negyedik szakaszában közel azonos módon oszlottak el az átlagos (34%), az átlagnál alacsonyabb (37%) és az átlagnál magasabb (29%) képességi szintű diákok. A rögzített formátumú teszttel ellentétben, ahol állandó volt a teszten belüli könnyebb, átlagos és nehezebb feladatok aránya, az adaptív feladat kiosztás során a magasabb képességi szintű diákok nagyobb arányban kaptak nehezebb, míg az alacsonyabb képességi szintű diákok könnyebb feladatokat. A teszt utolsó szakaszában, a diákokat közel harmadolva azonos átfedéssel, egyértelműen kialakult a három képességi sáv.



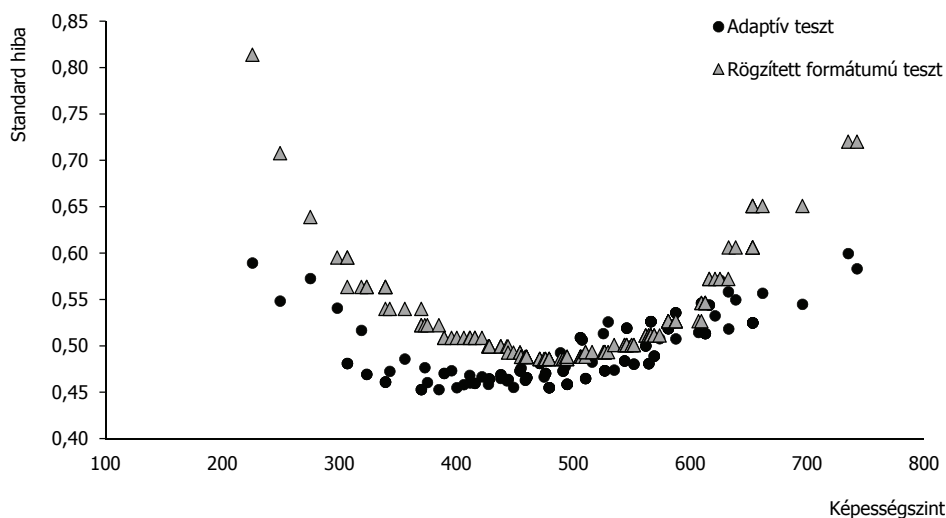
4. ábra

A tanulók gyakorisági és képességi szint szerinti eloszlása a szakaszokon és a modulokon belül

A rögzített és az adaptív tesztelés során kinyert információ és a mérési hiba nagyságának összehasonlítása

A kutatás felépítése diákszinten lehetővé teszi a becült képességszintek pontosságának összehasonlítását. Diákonkénti bontásban a rögzített, illetve az adaptív tesztkörnyezetben történt képességszint-becslés során elkövetett hiba nagyságát összehasonlítva (Wang, 2001, 2010) megállapítható, hogy a rögzített formátumú teszt alapján történt képességszint-becslés hibáinak nagysága diákszinten átlagosan nagyobb ($t=-7,54$, $p<0,01$; $se_{\text{átlag}}=0,53$), mint ugyanazon diákok adaptív tesztkörnyezetben történt képességszint-becslésének hibája ($se_{\text{átlag}}=0,49$). A teljes minta vonatkozásában pontosabban, kisebb mérési hibával történt adaptív tesztkörnyezetben a diákok képességszintjének becslése. Hipotézisünk alapján azonban a hiba nagysága nem egyenletesen oszlik el a teljes képességskálán: különböző mintázat várható az alacsonyabb, az átlagos és a magasabb képességtartományban.

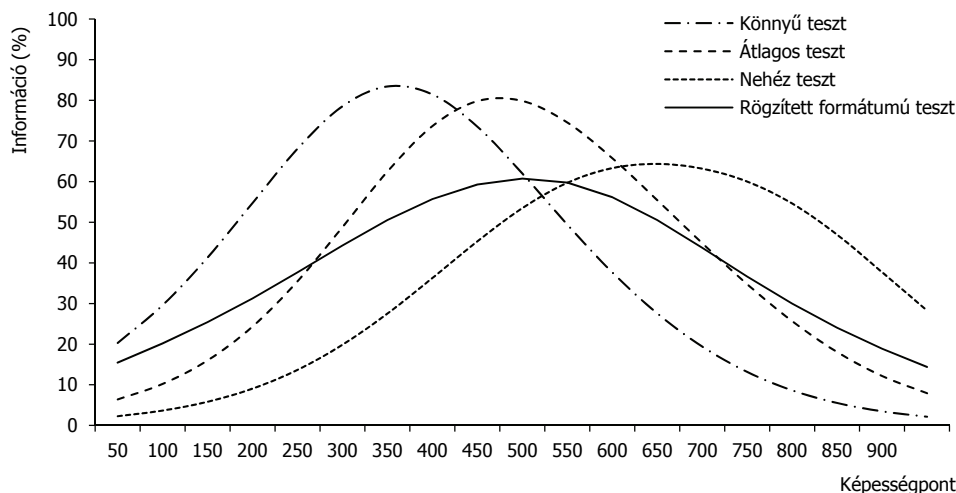
Összehasonlítva a rögzített formátumú és az adaptív teszten elért eredmények sztenderd hibáinak változását (5. ábra), hipotéziseinknek megfelelően, az alacsonyabb és a magasabb képességtartományban nagyobb hibával mér a rögzített formátumú teszt, mint az átlagos képességszintű diákok körében. Ez azt jelenti, hogy az adaptívteszt-algoritmus alkalmazásának előnye ezen képességtartományok esetén a legjelentősebb, átlagos képességszintű diákok mérése során közel azonosnak bizonyult a két tesztkörnyezetben becült képességszintek sztenderd hibáinak alakulása.



5. ábra

Az adaptív és a rögzített formátumú lineáris teszt standard hibáinak alakulása a tanulók képességszintjének függvényében

A pontosság egy másik mutatója a tesztelés során kinyert információ nagysága, amit jelen esetben a (rész)teszteken nyújtott teljesítmények alapján felrajzolt tesztinformációs görbék segítségével jellemzünk. A 6. ábra görbéi grafikusán szemléltetik, hogy már akár a kizárólagosan könnyű, átlagos, illetve nehéz modul résztesztjeiből összeállított tesztek (a 17 tesztváltozatból csak 3) is több információt szolgáltatottak a tesztelés során, mint az egyetlen, sokféle nehézségű feladatot tartalmazó rögzített formátumú teszt. A görbéket a 4. ábrán ismertetett képességszintekkel egybevetve megállapítható, hogy minden képességszinten több információt tudunk kinyerni adaptív tesztek alkalmazásával (a görbék minden esetben a rögzített formátumú teszt által adott információs függvény felett futnak az érintett képességtartományokban). A többletinformáció már abban az esetben is kimutatható, ha a diákokat az első részteszten nyújtott teljesítményük alapján három csoportba soroljuk, majd az adott képességtartományhoz közeli nehézségi szintű feladatokból állítjuk össze a tesztet.



6. ábra

Az adaptív technikával összeállított és a rögzített formátumú, azonos nehézségű részteszteket tartalmazó tesztek információs függvényei

Összegzés

A vizsgálat során ugyanazon mintán ugyanazon konstruktum mérése kapcsán összehasonlítottuk a számítógépes, rögzített formátumú tesztelés és a többszakaszos adaptív tesztelés során becsült képességszinteket, a becslés pontosságát, valamint a kinyert információ nagyságát. A kutatás célja annak vizsgálata volt, hogy a hagyományos rögzített formátumú tesztekkel az adaptív tesztelésre való átállás biztosítja-e és milyen mértékben a nagyobb mérési precizitás elérését. A kutatásban egy – ugyanazon a mintán fel-

vett – 28 ítemes rögzített formátumú és egy összességében 28 ítemes, ám 1-3-3-3 szerkezetű négyszakaszos adaptív teszten elért teljesítményeket, a becült képességszinteket, azok sztenderd hibáit, a tesztelés során kinyert információ nagyságát, valamint a tesztek jóságmutatóit hasonlítottunk össze.

Az eredmények alapján megállapítható, hogy a teljes minta szintjén több információt nyerünk ki a tesztelés során, szignifikánsan pontosabb képességszint-meghatározást végzünk adaptívteszt-algoritmus alkalmazásával, mint hagyományos, rögzített formátumú teszteléskor. A kinyert információ százalékos nagyságát összehasonlítva, míg a rögzített teszt átlagosan 60%-os információt szolgáltatott, addig az adaptív tesztelés során kinyert átlagos információ nagysága 76% volt. Az eltérés elsősorban az alacsony és a magas képességszintű tanulók esetében volt jelentős, előbbi esetén közel 34%, utóbbi során közel 24%-kal több volt az adaptív tesztből kinyert információ mennyisége. A képességszintek becslése során elkövetett hiba nagysága is ezzel párhuzamosan csökkent az adaptívteszt-algoritmus alkalmazása során. A diákok képességszintjéhez igazodó tesztelési eljárás hatékonysága már abban az esetben is jelentős, ha egy rövid, megfelelő jóságmutatókkal rendelkező kezdő részteszten elért eredmény alapján három képességszintbe soroljuk a diákokat, majd mindenki a saját képességszintjének képességtartományában legtöbb információt szolgáltató, rögzített formátumú tesztet kap.

A kutatás eredményei több szempontból korlátozottak – a minta és az itembank mérete, struktúrája, az alkalmazott itemek tartalmi és pszichometrikus jellemzői – mely korlátok nagyban befolyásolhatják a képességszintek becslését, ezért további kutatások szükségesek különböző méretű és tartalmi lefedésű itembankok felhasználásával a kinyert információ mértékének pontosabb meghatározására. A kutatás egyedisége, hogy az adaptív tesztelés hatékonyságát vizsgáló legtöbb kutatással szemben nem szimulált adatbázison, hanem empirikus adatok segítségével hasonlítottuk össze a rögzített és az adaptív tesztkörnyezetben becült képességszintek alakulását, továbbá az azonos minta alkalmazása lehetővé tette a diákszintű összehasonlítást is. Az eredmények alátámasztották a szimulációs kísérletekben is tapasztaltakat, miszerint jelentős mértékű mérési precizitás érhető el akár már három nehézségi szintet megkülönböztető adaptívteszt-algoritmus alkalmazásával a hagyományos lineáris tesztekhez képest.

A kutatást a TAMOP 3.1.9/11 kutatási program, az Oktatásméleti Kutatócsoport és az MTA-SZTE Képességfejlődés Kutatócsoport támogatta.

Irodalom

- Al-A'ali, M. (2007): Implementation of an improved adaptive testing theory. *Educational Technology & Society*, **10**. 4. sz. 80–94.
- Armstrong, R. D. (2002): *Routing rules for Multiple-Form Structures*. (Computerized Testing Report 02-08). Law School Admission Council. <http://www.lsac.org/lisacresources/research/ct/pdf/ct-02-08.pdf>. Utolsó letöltés: 2013. április 07.

- Armstrong, R. D., Jones, D. H., Koppel, N. B. és Pashley, P. J. (2004): Computerized adaptive testing with multiple-form structures. *Applied Psychological Measurement*, **28**. 147–164.
- Baker, F. B. (2001): *The basics of item response theory*. ERIC Clearinghouse on Assessment and Evaluation, University of Maryland, College Park, MD.
- Csapó Benő, Molnár Gyöngyvér és R. Tóth Krisztina (2008): A papír alapú tesztekől a számítógépes adaptív tesztekig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 3–4. sz. 3–16.
- Frey, A. és Seitz, N. N. (2009): Multidimensional adaptive testing in educational and psychological measurement: Current state and future challenges. *Studies in Educational Evaluation*, **35**. 2–3. sz. 89–94.
- Frey, A., Seitz, N. N. és Kröhne, U. (2011): Reporting differentiated literacy results in PISA by using multidimensional adaptive testing. In: Prenzel, M., Kobarg, M., Schöps, K. és Rönnebeck, S. (szerk.): *Research in the context of the Programme for International Student Assessment*. Springer, Berlin. 1–33.
- Jodoin, M., Zenisky, A. és Hambleton, R. K. (2006): Comparison of the psychometric properties of several computer-based test designs for credentialing exams with multiple purposes. *Applied Measurement in Education*, **19**. 3. sz. 203–220.
- Keng, L. (2008): *A comparison of the performance of testlet-based computer adaptive tests and multistage tests*. The University of Texas, Austin.
- Kingsbury, G. G. és Hauser, C. (2004): Computerized adaptive testing and no child left behind. Előadás, Annual Meeting of the American Educational Research Association, San Diego, CA.
- Linacre, J. M. (2000): *Computer-adaptive testing: A methodology whose time has come*. MESA Psychometric Laboratory, University of Chicago.
- Magyar Andrea (2012): Számítógépes adaptív tesztelés. *Iskolakultúra*, **22**. 6. sz. 52–60.
- Molnár Gyöngyvér (2011): Az információs-kommunikációs technológiák hatása a tanulásra és oktatásra. *Magyar Tudomány*, 9. sz. 1038–1047.
- Molnár Gyöngyvér (2013): *A Rasch modell alkalmazási lehetőségei az empirikus kutatások gyakorlatában*. Gondolat Kiadó, Budapest.
- Thompson, T. és Way, D. (2007): Investigating CAT designs to achieve comparability with a paper test. In: Weiss, D. J. (szerk.): *Proceedings of the 2007 GMAC Conference on Computerized Adaptive Testing*. www.psych.umn.edu/psylabs/CATCentral/. Utolsó letöltés: 2013. április 04.
- Tian J., Miao, D. és Zhu Xia, G. J. (2007): An introduction to the computerized adaptive testing. *US-China Education Review*, **4**. 1. sz. 72–81.
- Tóth Zoltán (2005): A tudásszerkezet és a tudás szerveződésének vizsgálata a tudástér-elmélet alapján. *Magyar Pedagógia*, **105**. 1. sz. 59–82.
- Wainer, H. (2000): *Computerized adaptive testing: A primer* (2nd Edition). NJ: Erlbaum, Hillsdale.
- Wang, T. és Kolen, M. J. (2001): Evaluating comparability in computerized adaptive testing: Issues, criteria and an example. *Journal of Educational Measurement*, **38**. 1. sz. 19–49.
- Wang, H. (2010): Comparability of computerized adaptive and paper-pencil tests. *Test, measurements and research services bulletin*. http://www.pearsonassessments.com/NR/rdonlyres/057A4A04-9DCB-4B68-9CB0-3F32DDF396F6/0/Bulletin_13.pdf. Utolsó letöltés: 2013. április 07.
- Zenisky, A., Hambleton, R. K. és Luecht, R. M. (2010): Multistage testing: Issues, designs and research. In: der Linden, W. J. és Glas, C. A. W. (szerk.): *Elements of adaptive testing*. Springer, New York. 355–372.

ABSTRACT

ANDREA MAGYAR AND GYÖNGYVÉR MOLNÁR: COMPARING THE EFFICACY OF
COMPUTERIZED ADAPTIVE AND FIXED-ITEM TESTING

With the rapid spread of computers in the past two decades, linear paper-and-pencil tests are gradually being replaced by computer-based assessment. The most advanced form is computerized adaptive testing, in which the test is adapted to examinees' ability level by only administering items of appropriate difficulty. The aim of this paper is to compare the effectiveness of fixed-item and adaptive tests from an assessment perspective by: (1) relating differences in student level achievement; (2) outlining item difficulties of delivered tests; and, finally, (3) comparing measurement error and test information functions in linear and adaptive test environments. The samples from the pilot study were drawn from children in Years 5 and 8 at Hungarian primary schools (N=158). A fixed-item test was administered to half of the participants; the other part took four-stage adaptive tests (1-3-3-3 structure). Two weeks later, the types of test were switched. Both tests measured inductive reasoning. A one-parameter Rasch model was used for the analyses. The reliability of the adaptive tests proved to be higher (Cronbach- α =.85) than that of the fixed test form (Cronbach- α =.83). The adaptive test provided consistently higher information at every skill level than that of the fixed-form test. The standard error of the four-stage test was significantly lower, especially in upper and lower ability levels. The study provided a promising step towards more precise educational assessment in using multistage testing with even three stages besides traditional linear test forms.

Magyar Pedagógia, **113**. Number 3. 181–193. (2013)

Levelezési cím / Address for correspondence: Magyar Andrea, SZTE Neveléstudományi Doktori Iskola, Molnár Gyöngyvér, SZTE Neveléstudományi Intézet, MTA-SZTE Képességfejlesztés Kutatócsoport H-6722 Szeged, Petőfi S. sgt. 30–34.

A folyóirat megjelenését a Magyar Tudományos Akadémia és az Oktatásért Közalapítvány támogatta.



Terjeszti a Magyar Posta Rt.

Magyar Posta Zrt. Értékesítési Ágazati Igazgatóság (1008 Budapest, Orczy tér 1.)

faxon: 06/1-303-3440 e-mailen: hirlapelofizetes@posta.hu

További információ: 06/80/444-444

Előfizetési díj egy évre 3200,- Ft. Ára példányonként 800,- Ft.

Külföldön terjeszti a KULTURA Külkereskedelmi Rt (H-1035 Budapest, Kerék u. 80.).

Az MTA Könyv- és Folyóiratkiadó Bizottsága megbízásából kiadja az SZTE BTK,
a kiadásért felel a BTK dékánja.

A szedés a Szegedi Tudományegyetem Neveléstudományi Intézetében készült.

Tördelőszerkesztő: Börcsökne Soós Edit.

Nyomták a GENERAL Nyomda Kft-ben. Felelős vezető: Hunya Ágnes.

Megjelent 4,2 (A/5) ív terjedelemben.

HU ISSN 0025-0260