

MAGYAR PEDAGÓGIA

A MAGYAR TUDOMÁNYOS AKADÉMIA
PEDAGÓGIAI TUDOMÁNYOS BIZOTTSÁGÁNAK FOLYÓIRATA

SZÁZHUSZONHATODIK ÉVFOLYAM

2. SZÁM



2026

MAGYAR PEDAGÓGIA

A MAGYAR TUDOMÁNYOS AKADÉMIA
PEDAGÓGIAI TUDOMÁNYOS BIZOTTSÁGÁNAK FOLYÓIRATA

Alapítás éve: 1892

A megjelenés szünetelt 1948-ban és 1951–60 között
A folyóirat megjelenését a Magyar Tudományos Akadémia Könyv- és Folyóiratkiadó
Bizottsága támogatta

SZÁZHUSZONHATODIK ÉVFOLYAM

Főszerkesztő:

CSÍKOS CSABA

Szerkesztők:

Habók Anita, Tóth Edit

Szerkesztőbizottság:

FALUS IVÁN, FÜLÖP MÁRTA, HALÁSZ GÁBOR, JÓZSA KRISZTIÁN,
KÉRI KATALIN, KÖLLŐ JÁNOS, MOLNÁR GYÖNGYVÉR, NÉMETH ANDRÁS,
NIKOLOV MARIANNE, PUSZTAI GABRIELLA, ZSOLNAI ANIKÓ

Nemzetközi tanácsadó testület (International Advisory Board):

LAVICZA ZSOLT (Linz), SUZANNE HIDI (Toronto)
LÁZÁR SÁNDOR (Kolozsvár), MARTON FERENC (Göteborg) SZŰCS DÉNES (Cambridge)

Szerkesztőség:

Szegedi Tudományegyetem, Neveléstudományi Intézet

6722 Szeged, Petőfi sgt. 32–34.

Tel./FAX: (62) 544–354

Technikai szerkesztő: Biró Fanni és Varga Andrea

Szerkesztőségi titkár: B. Németh Mária

Journal of the Educational Committee of the Hungarian Academy of Sciences

Editor-in-chief: Csaba Csikos, University of Szeged, H–6722 Szeged, Petőfi sgt. 32–34.

Tel./FAX: 36–62–544354 E-mail: szerk@magyarpedagogia.hu / Web: <https://www.magyarpedagogia.hu>

TARTALOM

TANULMÁNYOK

Majoros Erika: Trendskalázási módszerek: A korai és közelmúltbeli IEA matematikai és természettudományos felmérések összekapcsolása	83
Magyar Andrea, Habók Anita és Molnár Gyöngyvér: Technológiaalapú szókincsfejlesztés és olvasásértés: egy nagymintás kontrollcsoportos kutatás eredményei	99
Malhazjan Szergej és Pikó Bettina: Narratív irodalmi áttekintés a pedagógusok munkahelyi stresszéről, munkaelégedettségéről és pszichoszomatikus stressz-indikátorairól	123



TRENDSKÁLÁZÁSI MÓDSZEREK: A KORAI ÉS KÖZELMÚLTBELI IEA MATEMATIKAI ÉS TERMÉSZETTUDOMÁNYOS FELMÉRÉSEK ÖSSZEKAPCSOLÁSA

Majoros Erika

Umeå University, Svédország

A Tanulói Teljesítmények Vizsgálatának Nemzetközi Társasága (International Association for the Evaluation of Educational Achievement, IEA) által szervezett nemzetközi nagymintás felmérések egyike a Trends in International Mathematics and Science Study (TIMSS), amely 1995-től mér matematikai és természettudományos tudást negyedik és nyolcadik osztályban. A TIMSS-t négy IEA felmérés előzte meg, amelyek eredményit nem kapcsolták össze a TIMSS skálával. Jelen tanulmány lehetséges trendskálázási módszereket ismertet, és a megvalósult összekapcsolás illusztrálására a magyar diákok újra skálázott matematikai teljesítmény és motiváció, valamint a természettudományos teljesítmény eredményeit mutatja be.

Elméleti keret

Az IEA 1995 óta végez rendszerszintű felméréseket a matematikai és természettudományi tanulói teljesítmények vizsgálatára, mely felmérések eredményeit trendskálákon közli. Ezt megelőzően az IEA négy nemzetközi nagymintás felmérést végzett ezekben a tantárgyakban, azonban ezeknek a korai felméréseknek eredményeit nem kapcsolta össze az 1995-ban alapított trendskálával. Magyarország az összes korai felmérésben részt vett az IEA-hoz történt 1968-as csatlakozása után (B. Németh et al., 2012; Csíkos & Vidákovich, 2012).

Az 1995 előtt végzett IEA felméréseket első szakaszbéli, az azt követőket pedig második szakaszbéli felméréseknek nevezzük Gustafsson (2008) nyomán. Az első szakaszbéli rendszerszintű felméréseket az 1. táblázat tartalmazza. A két szakasz közötti összekapcsolás hiányát a korai felmérések közötti, a mérési eszközökben, a vizsgált populációkban és a mérési eljárásokban bekövetkezett változtatások indokolták (Martin & Kelly, 1996a). Az akkori technológiai és módszertani lehetőségek is korlátozhatták az összekapcsolás megvalósíthatóságát. Az első felmérés óta eltelt évtizedek során új technikai lehetőségek jelentek meg, így változtatások történtek például a tanulók és a tesztfeladatok mintavételével vagy az attitűdöket vizsgáló kérdések megfogalmazásával kapcsolatban. Ezek a döntések kihívások elé állítják az összehasonlíthatóságot, következésképpen az eredmények összekapcsolását. A modern technikai és módszertani fejlesztések azonban lehetővé teszik az ilyen kihívások kezelését.

1. táblázat

Az IEA által az első szakaszban szervezett nemzetközi matematikai és természettudományos felmérések

Felmérés	Adatgyűjtés ideje	Részt vevő oktatási rendszerek száma
Első Nemzetközi Matematikai Vizsgálat (First International Mathematics Study; FIMS)	1964	12
Első Nemzetközi Természettudományi Vizsgálat (First International Science Study; FISS)	1970–71	17
Második Nemzetközi Matematikai Vizsgálat (Second International Mathematics Study; SIMS)	1980–82	20
Második Nemzetközi Természettudományi Vizsgálat (Second International Science Study; SISS)	1983–84	24

Korábbi kutatások igazolják, hogy a két szakaszbeli IEA felmérések eredményeit össze lehet kapcsolni, különféle módszertani megközelítésekkel. Az egyik megközelítés az, hogy olyan tesztek kapcsolnak össze, amelyek megismételt feladatokat tartalmaznak, modern tesztelméleti (item response theory, IRT) modellek alkalmazásával. Afrassa (2005), Majoros és munkatársai (2021), valamint Strietholt és Rosén (2016) ezzel a megközelítéssel kapcsoltak össze matematika és olvasási teljesítményt vizsgáló felméréseket. Azonban ezek a matematikai összekapcsolási tanulmányok (Afrassa, 2005; Majoros et al., 2021) korlátozottak a TIMSS-skálával való összehasonlíthatóság és a bevont országok számának szempontjából. A nyolcadik évfolyamos matematika és természettudományos trendskála minden résztvevő országra kiterjedő kibővítése megvalósult Majoros (2023) munkájában. A negyedik évfolyamos természettudományos trendskálázás megvalósult Sosa Paredes és Andersson (2025) tanulmányában.

Egy másik lehetséges megközelítés különböző regionális, országos vagy nemzetközi teszt-eredmények összekapcsolására az, amelyben a tesztek nem tartalmaznak megismételt feladatokat, így a klasszikus tesztelméletre (classical test theory, CTT) támaszkodik. Ebben a megközelítésben az összekapcsolás szigorúbb feltételekre épül a képességek eloszlására vonatkozóan (ld. pl. Chmielewski, 2019; Hanushek & Woessmann, 2012).

Az matematika- és természettudományos tanulás affektív aspektusainak trendmérése az IEA által a TIMSS 2011-es adminisztrációjával kezdődött. A 2011-es, a 2015-ös és a 2019-es háttérkérdőívek tartalmaznak megismételt kérdéseket, és eredményeiket trendskálákon jelentik, a kognitív eredményekhez hasonlóan (Martin et al., 2012, 2016; Yin & Fishbein, 2020). Ezen affektív skálák longitudinális kiterjesztését vizsgálja Majoros és munkatársai (2022) tanulmánya.

Összegezve megállapítható, hogy a tesztelméleti módszertani fejlődés és a nemzetközi nagymintás felmérések jelentőségének növekedése miatt érdemes feltárni a korai adatokban rejlő lehetőségeket. Jelen tanulmány kétféle módon járul hozzá a szakirodalomhoz. Egyrészt, a tárgyalt összekapcsolási technikák alkalmazhatók más olyan felmérésekre is, amelyekben az adminisztrációk között változások történtek. Másrészt, a bemutatott skálák felhasználhatók jövőbeni longitudinális tanulmányokhoz.

Rendszerszintű oktatási eredmények mérése

A mérés számok hozzárendelése tárgyak vagy események aspektusaihoz egy vagy másik szabály vagy egyezmény szerint (Stevens, 1968). Az oktatási mérés olyan mérést jelent, amely oktatási környezetben történik. Az oktatási rendszerek összetett és többszintű struktúrák, melyek tervezéséhez és irányításához az oktatási mérések szolgálnak visszajelzésül (Csíkos et al., 2020). Rendszerszinten a visszajelzést nemzetközi és/vagy nemzeti értékelések biztosítják. A nemzetközi és nagymintás felmérések a viszonylag hosszú visszajelzési idő miatt előnyösek lehetnek a közép- vagy hosszú távú tervezéshez. Ezen felmérésekben való részvétel különösen előnyös lehet azoknak az országoknak, amelyek nem tartanak fenn nemzeti értékeléseket.

Az IEA felmérésekben a méréseket teljesítménytesztek és háttérkérdőívek segítségével végzik. Ezen mérések a pszichometriában, azaz a pszichológiai tesztek tulajdonságainak értékelésével foglalkozó tudományában kifejlesztett elméleteket és módszertanokat alkalmazzák (Furr & Bacharach, 2014). Cronbach (1960) a pszichológiai tesztet két vagy több ember viselkedésének összehasonlítására szolgáló szisztematikus eljárásként határozta meg (p. 21). A pszichológiai tesztek fejlesztésének alapelvei az elmúlt évtizedekben drámaian megváltoztak (Embretson & Reise, 2000).

A valószínűséginek is nevezett modern tesztelmélet a pszichológiai értékelések fő elméleti alapjává vált, felváltva a CTT-ként emlegetett elveket. A CTT-keretrendszerben a módszereknek és a mérési eljárásoknak számos hiányosságuk van (Hambleton et al., 1991). A legfontosabb korlát az, hogy a tesztmegoldó jellemzői, mint például a képesség, valamint a teszttulajdonságok, mint például a nehézség, csak egymással összefüggésben értelmezhetők. Ezért a CTT keretében az eredmények tesztfüggőek, ami negatív következményekkel jár a tesztteredmények összehasonlíthatóságára nézve. Ez és a CTT egyéb korlátai arra késztették a pszichometrikusokat, hogy alternatív elméleteket és modelleket dolgozzanak ki a pszichológiai mérésre. A következőkben a látens, azaz nem közvetlenül megfigyelhető pszichológiai konstrukciók, mint például a matematikai képesség mérésére szolgáló két fő keretrendszert vázoljuk: a valószínűségi tesztelméletek és az ellenőrző faktorelemzés (confirmatory factor analysis, CFA) keretrendszereit.

A modern tesztelmélet alkalmazása

Az IRT egy modellalapú mérési keretrendszer (Embretson & Reise, 2000). Az IRT két alapvető gondolata, hogy a tesztmegoldó teljesítménye egy adott itemen látens tulajdonságok vagy képességek alapján előre jelezhető, és hogy az itemteljesítmény és a látens tulajdonság közötti kapcsolat kifejezhető az item karakterisztikus görbéje segítségével (ld. pl. Embretson & Reise, 2000; Hambleton et al., 1991; Molnár, 2006).

Hambleton és Jones (1993) rámutattak az IRT tesztelésre való alkalmazásának négy előnyére a klasszikus tesztelméleti megközelítésekkel szemben. Először is, az itemstatisztikák függetlenek attól a mintától, amelyből becsülték őket. Másodszor, a jártassági pontszámok függetlenek a teszt nehézségétől. Harmadszor, az IRT-modellek alapot biztosítanak a tesztítemek képességszintekhez való illesztéséhez. Végül, az IRT-modellek nem igényelnek szigorúan párhuzamos teszteket a megbízhatóság értékeléséhez. Az IRT további előnye, hogy lehetővé teszi az ítemek és a tesztmegoldók mátrix mintavételét, amely mintavételi módszert gyakran használnak nagymintás felmérésekben.

A FIMS és a FISS kivételével az itt tárgyalt teljesítménytesztekben ún. mátrix mintavételi megközelítést alkalmaztak. Ez azt jelenti, hogy a tesztek összesen több itemet tartalmaznak, mint amennyit egyes tesztmegoldók kézhez kapnak. Lényegében a mátrix mintavétel során az

itemek mintáit diákok mintáin adminisztrálják (Mazzeo et al., 2006). Ezt a kettős mátrix mintavételt két ok motiválja. Először, a mért képességek és ismeretek terjedelme annak megfelelően nagy itemkészletet igényel. Másodsor, a diákok természettudományos és matematikai ismereteinek, valamint alkalmazásának tesztelése időigényes itemtípusokat igényel. A mátrix mintavétel előnye, hogy csökkenti az egyes diákok tesztelési terheit. Továbbá, a mátrix mintavételezéssel növelhető a teszt érvényessége, mivel a diákok nagyobb valószínűséggel töltik ki a tesztet a megadott időn belül. Összekapcsolási módszerek alkalmazásával összehasonlítható eredmények érhetők el különböző tesztformák használatával.

Ellenőrző faktorelemzés alkalmazása

A látens változók modellezésének egy másik megközelítése az ellenőrző faktoranalízis (Jöreskog, 1969). A CFA egy statisztikai technika a strukturális egyenletek modellezése (structural equation modeling, SEM) keretrendszerében. A CFA az indikátorként szolgáló megfigyelt változók és a látens változók, azaz faktorok közötti kapcsolatot vizsgálja. A faktor egy mögöttes konstrukció, amelyről feltételezik, hogy befolyásolja az indikátorokat, és figyelembe veszi a közöttük lévő korrelációkat. A kapcsolatok mintázatát faktorsúlyok fejezik ki. A faktorpontszámok faktorsúlyok segítségével számíthatók ki. A faktorpontszámok a tesztelők manifest pontszámát jelentik a látens dimenzió (Brown, 2015).

A CFA gyakran használt eszköz a skálafejlesztés során. Ezt a módszert gyakran alkalmazzák például a mérőeszközök pszichometriai tulajdonságainak értékelésére, konstrukciók validálására, a mérési invariancia csoportok közötti tesztelésére vagy a módszerhatások vizsgálatára (Brown, 2015).

A látens konstrukciókat felmérő kérdőíves adatok CFA módszerekkel skálázhatók. Az affektív itemek skálázására szolgáló CFA-modellezés alkalmazásának egyik példája az OECD által 2008 óta végzett Teaching and Learning International Survey (TALIS). Ezenkívül mérési invariancia tesztelést alkalmaztak többcsoportos ellenőrző faktoranalízissel (multiple-group confirmatory factor analysis, MGCFA) a résztvevő oktatási rendszerek összehasonlíthatóságának értékelésére (OECD, 2019).

Skálázás

Stevens (1968) négy típusba sorolja a skálákat: nominális (pl. sorrend nélküli kategóriák, mint a futballisták számozása), ordinális (pl. korlátozott számú kategória a Likert-skála szerinti sorrenddel, mint például az ásványok keménységi skálája), intervallum (folytonos és egyenlő távolságra lévő számok valódi nulla nélküli sora, mint pl. a Celsius- és Fahrenheit-fokok, naptári idő), és arányskálák (folytonos, egyenlő távolságra lévő számok valódi nullát tartalmazó sora, mint pl. a hosszúság). Kolen (2006) az oktatási mérési keretrendszerben a skálázást úgy határozta meg, mint számoknak vagy más rendezett mutatóknak egy oktatási teszten elért teljesítményhez történő hozzárendelése. A tesztitemeket jellemzően helyesnek, részben helyesnek vagy helytelennek pontozzák, és a nyers pontszám a helyes válaszok számán alapul.

Különböző típusú transzformációk léteznek a skálapontszámok nyers pontszámokból történő előállítására. A nemzetközi nagymintás felmérésekben a képességskálák előállítására az IRT-alapú skálázás a jellemző. Az IRT képességskálázási konvenciói olyan skálapontszámokat eredményeznek, amelyek átlaga 0 és szórása 1, és amelyek lineáris transzformációt igényelnek az értelmezéshez (Kolen, 2006). Például a TIMSS nemzetközi teljesítményskáláit az első adminisztrációkor 500 átlaggal és 100 szórással állították be (Martin & Kelly, 1996b).

A skálázás során a valószínűsíthető érték (plausible value, PV) elnevezésű módszertant (ld. pl. Beaton & Johnson, 1992; Mislevy, 1991; Mislevy et al., 1992; Rutkowski et al., 2010) alkalmazzák a populáció szintű képességbecslések generálására egy mátrix mintavételezésű tesztervből. A PV módszertan szerint a tanulói teljesítményt hiányzó értéként kezelik. Minden tanuló válaszol a tesztelemelek egy részhalmazára, és ezeket a válaszokat használják fel a tanulói képességeloszlás generálására a populációban. A populáció képességeloszlásának becsléséhez Rubin (1987) többszörös imputációs megközelítésének mérési modellen alapuló kiterjesztését alkalmazzák. A becsült képességeloszlásból véletlenszerű PV értékeket rendelnek minden tesztmegoldóhoz. A modern nemzetközi nagymintás felmérésekben alkalmazott PV-módszertan eltér az egyéni jártassági pontszámokra összpontosító más megközelítésektől (L. Rutkowski et al., 2010; Mazzeo et al., 2006). A PV-megközelítés lehetővé teszi a nagy itembankok használatát a pontosabb jártassági becslés érdekében, de a mátrixmintavétel miatt a jelentett pontszámokat csoportszinten kell kezelni. Az ilyen típusú értékeléseket csoportos pontszámértékeléseknek nevezik (Mazzeo et al., 2006). A tanulmányi eredmények mérése nemzetközi kontextusban sajátos kihívásokkal jár. A következő szakaszban ezeket a kihívásokat vázoljuk fel.

Mérési invariancia és itemtorzítás

A tanulói eredmények több oktatási rendszert magában foglaló és összehasonlító értékelésénél figyelembe kell venni a kulturális különbségeket, például a mérendő konstrukciók jelentését illetően. A mérési invariancia vagy ekvivalencia egy mérés azon tulajdonságára utal, amely arra vonatkozik, hogy a pontszámok ugyanazt a jelentést hordozzák-e különböző körülmények között, pl. az értékelés időpontja, az adminisztrációs módszerek, országok vagy populációk között (Kline, 2016; Meade & Lautenschlager, 2004). Ahogy Kline (2016) rámutat, az irodalomban számos címke létezik ugyanarra az invarianciatípusra. A legtöbb terminológia a Meredith (1993) által megállapított invarianciaszintek hierarchiájából származik. Ezek az invarianciaszintek a paraméterekre vonatkozó csoportok közötti korlátozások hierarchiáját képviselik (Wu & Estabrook, 2016).

A faktoriális invariancia hierarchiája négy szintből áll: konfigurációs (mintázati), gyenge (metrikus), erős (skaláris) és szigorú (reziduális) invariancia (Kline, 2016; Meredith, 1993; Meredith & Teresi, 2006; Putnick & Bornstein, 2016). A konfigurációs invariancia a legkevésbé korlátozó szint, és azt jelenti, hogy a konstrukciók alapvető szerveződése, azaz a faktortöltések mintázata azonos a csoportok, pl. országok között. A gyenge faktoriális invariancia a konfigurációs invarianciára épít, és azt jelenti, hogy a faktortöltések egyenértékűek a csoportok között. Az erős faktoriális invariancia azt jelenti, hogy a faktorátlagok egyenértékűek a csoportok között. A hierarchia tetején található szigorú faktoriális invariancia azt jelenti, hogy a reziduális értékek varianciái egyenértékűek a csoportok között.

A mérési invariancia, vagyis az itemtorzítás hiányának fogalmát eredetileg az IRT kontextusában fogalmazták meg (Meredith & Teresi, 2006). A mérési invariancia – a szigorú terminológia szerint a faktoriális invariancia – és az itemtorzítás hiánya közötti kapcsolatról elmondható, hogy a két fogalom különálló, de szorosan összefügg egymással (Meredith & Teresi, 2006). Az itemtorzítás hiányát Meredith és Teresi (2006) az itemparaméterek csoportok közötti invarianciájaként definiálja. Amikor ez az invariancia nem teljesül, különböző itemműködés (differential item functioning, DIF) következik be.

Skálák összekapcsolása

Az előző szakaszok a nemzetközi nagymintás értékelések keresztmetszeti aspektusával foglalkoztak. A második fázisú IEA felmérések egyik fő célja a kognitív eredmények rendszerszintű trendmérésének fenntartása, és újabban bizonyos háttérváltozók trendskalázása is. Ebből a célból a felméréseket összekapcsolják annak érdekében, hogy az eredmények ugyanazon a trendskálán legyenek értelmezhetőek. Ebben a szakaszban a keresztmetszeti mérési pontok összekapcsolásának keretrendszerét vázoljuk fel.

Az nemzetközi nagymintás felmérésekben a trendskálák felépítése két vagy több teszt összekapcsolásán alapul. Az összekapcsolás az a folyamat, amelynek során a tartalmukban és/vagy nehézségükben eltérő tesztek eredményeit statisztikai módszerekkel összemérhetővé teszik (Holland & Dorans, 2006; Linn, 1993; Mislevy, 1992). Az összekapcsolás legerősebb változata a kiegyenlítés (equating). Mazzeo és von Davier (2013) a skálák összekapcsolásának folyamatát úgy definiálják, mint felmérések sorozata által előállított eredmények olyan skálájának elérését, amely hosszútávon stabil és összehasonlítható jelentést tart fenn.

Két vagy több értékelés összekapcsolása az IRT-keretrendszerben magában foglalja a különböző tesztadminisztrációkból származó itemparaméter-bebecslések egy közös skálára helyezését. A Hambleton és munkatársai (1991) által meghatározott négy összekapcsolási design az egycsoportos, az ekvivalens csoportokon alapuló, a közös személyeken alapuló és a horgonyteszt módszertan.

Egycsoportos design esetén ugyanaz a hallgatói csoport tölti ki az összekapcsolandó teszteket. Ez a design nem feltétlenül praktikus, mert a tesztelési idő túl hosszú az egyének számára, ezért a fáradtság befolyásolhatja az eredményeket. Ezenkívül a több teszt kitöltése által okozott gyakorlási hatás akadályozhatja a paraméterbecslés pontosságát.

Az ekvivalens csoportokon alapuló design alkalmazása során véletlenszerűen kiválasztott és egyenértékű hallgatói csoportokkal veszik fel az összekapcsolandó teszteket. Ez a design praktikusabb, mint az első, mert elkerüli a fáradtság és a gyakorlás hatásait.

A közös személyeken alapuló design azt jelenti, hogy az összekapcsolandó teszteket megoldó csoportokban van egy közös tesztmegoldó halmaz. Ez a design hasonló kihívásokat jelent a közös csoport számára, mint az egycsoportos terv, ezért kevésbé tekinthető praktikusnak.

A horgonyteszt design megoldja a fent említett hiányosságokat. A hallgatók különböző csoportjai töltik ki az összekapcsolandó teszteket. Minden teszt tartalmaz egy csoport közös itemet, amelyek összességét horgonytesztnak nevezzük. Ez a design nagyon népszerű a megvalósíthatósága miatt.

Az IEA a teljesítményteszteket különböző időpontokban (a TIMSS esetén négyévente) veszi fel, és az egymást követő adminisztrációk tartalmaznak megismételt tesztitemeket. Ezen horgonytesztet alkalmazó designon belül számos lehetséges paraméterbecslési módszer alkalmazható az értékelések összekapcsolására.

A különböző tesztekben származó válaszadatokon alapuló paraméterbecslés elvégezhető együtt vagy külön-külön. Az az eljárás, amelynek során a különböző tesztformákból származó válaszadatokat összevonják a paraméterbecsléshez, egyidejű kalibrálásnak nevezzük. Ez a módszer kisebb standard hibákat biztosít, kevesebb feltételezést tartalmaz, mint más IRT-eljárások, és jó összekapcsolás érhető el kevés közös itemmel (Wingersky & Lord, 1984).

Külön paraméterbecslés esetén a tesztek összekapcsolásának egyik módja, hogy a horgonytétel tételeparamétereit az egyik teszten becsült értékekhez rögzítik, amikor a többi teszt tételét kalibrálják. Ezt a folyamatot Kolen és Brennan (2014) fixparaméteres kalibrációnak nevezi. Alternatív megoldásként a paraméterek külön-külön is becsülhetők a horgonytesztre, és skálatranszformáció alkalmazható úgy, hogy a közös tételparaméterek eloszlása megegyezzen.

A Dorans és munkatársai (2011) által leírt legjobb kiegyenlítési gyakorlatok néhány aspektusa a skálák összekapcsolására is vonatkozik. Először is, mivel a gyűjtött adatok mennyisége jelentős hatással van az eredmények hasznosságára, a nagy reprezentatív minták kedvezőek. Az ilyen adatok biztosítják, hogy a statisztikai bizonytalanság (standard hiba) sokkal kisebb legyen, mint az eredmények egyéb variációs forrásai. Másodszor, a horgonyitemek viselkedését a különböző adminisztrációkban értékelni kell. A feltételezés az, hogy az itemek viselkedése hasonló az azonos évfolyamú diákcsoportok között, és ezt DIF-módszerekkel tesztelik. Végül, a horgonyteszten elért teljesítménynek szorosan korrelálnia kell a teljes teszteredménnyel.

Skálák összekapcsolása a TIMSS felmérésekben

A TIMSS-felmérésekben alkalmazott skálázási és összekapcsolási eljárásokat az IEA által készített különféle technikai jelentések részletesen ismertetik (ld. pl. Martin et al., 2020; Martin & Kelly, 1996b). Jelen tanulmány kétféle oktatási eredményre összpontosít. Először a kognitív eredményekre, azaz a tanulók nyolcadik évfolyamos matematika és természettudományos teljesítményére. Másodszor affektív eredményekre, pontosabban, hogy a diákok mennyire motiváltak a matematika tanulására. Az alábbiakban a mért konstrukció típusa alapján foglaljuk össze az eljárásokat.

Kognitív teljesítmény: matematika és természettudomány

A skálázási eljárás három fő lépésből áll: itemparaméter-bebecslés, személypontosítás és populációmodellezés. Az első lépést IRT egyidejű kalibrációval végzik, azaz az egyes TIMSS-ciklusok itemparamétereit az aktuális és az előző értékelés adatai alapján becsülik meg. Az első lépésben a háromparaméteres logisztikus modellt (3PL) alkalmazzák a dichotóm feleletválasztós itemekhez, a kétparaméteres logisztikus modellt (2PL) a dichotóm szabad választású itemekhez, és az általánosított parciális kreditmodellt (generalized partial credit model, GPCM) a politóm szabad választású itemekhez, azaz azokhoz, amelyek maximális pontszáma egynél több. A korai ciklusokban a skálázást olyan mintákon végezték, amelyek azonos méretű, véletlenszerűen kiválasztott almintákból álltak a részt vevő országok mindegyikéből (ld. pl. Adams et al., 1996). Újabban a teljes mintákat használják és úgy súlyozzák, hogy minden ország egyenlően járuljon hozzá az itemkalibrációhoz (ld. pl. Foy et al., 2020). Ez utóbbi megközelítés azt jelenti, hogy több információt használnak fel az itemkalibrációhoz, ami pontosabb paraméterbecsléseket biztosít, de számítási szempontból is igényesebb.

A személypontosítási eljárás során öt PV-t rendelnek minden tesztmegoldóhoz az IRT modellalapú skálapontszám-eloszlásból. A tanulói pontszámok becslésének megbízhatóságának növelése érdekében a TIMSS skálázás harmadik lépése egy populációmodellezésnek vagy kondicionálásnak nevezett folyamat. Ez azt jelenti, hogy az összes rendelkezésre álló háttér-adatot bevonják egy főkomponens-elemzésbe, és az adatok varianciájának 90%-át képviselő komponenseket minden oktatási rendszer regressziós modelljébe foglalják.

Miután a skálázás befejeződött, az újonnan létrehozott skálát összekapcsolják a TIMSS trendskálával. Ez két lépésben valósul meg. Az első lépés az új értékelési adatok látens képességeloszlásának lineáris transzformációja, hogy megfeleljen az előző ciklus eloszlásának. Az eloszlások egyeztetéséhez a közös országok, azaz a két ciklus trendországainak adatait használják fel. Ezek a trendországok egyenlően járulnak hozzá a transzformációs állandók kiszámításához. A második lépés ennek a lineáris transzformációnak az alkalmazása az újonnan

gyűjtött felmérési adatok egészére. Ezen lépések befejezése után az átalakított PV-k kerülnek a trendskálára (ld. pl. Foy et al., 2020).

Tanulási motiváció és más attitűdskálák

Az 1995 és 2007 között végzett TIMSS felmérésekben ún. indexeket képeztek bizonyos háttérkérdőív-skálához. A legújabb TIMSS háttérkérdőíves skálák (2011-től kezdődően) IRT-skálázással készülnek, a Rasch parciális kreditmodell (partial credit model, PCM) alkalmazásával. Az affektív konstrukciók longitudinális skálázása először a TIMSS 2011 és a TIMSS 2015 összekapcsolásával jelent meg (Martin et al., 2016). Jelenleg bizonyos háttérkérdőíves skálák, amelyek a 2011-es, 2015-ös és 2019-es TIMSS-ciklusokban számos horgonyitemet tartalmaztak, egy kétlépéses transzformációs folyamaton keresztül kapcsolódnak össze. Az első transzformáció a TIMSS 2019 logit skála pontszámait a TIMSS 2015 skálájára helyezi az átlag/szigma módszer (Kolen & Brennan, 2014; Marco, 1977) alkalmazásával. Ezeket a skálákat a TIMSS 2019 és a TIMSS 2015 adatok külön kalibrálásával becsülik meg. A második lépés a TIMSS 2015 logit pontszámok transzformálása a trendskálára (átlag 10, szórás 2).

Módszer

Jelen tanulmány a trendskálázási módszerek illusztrálására a magyar diákok újra skálázott matematikai teljesítmény és motiváció, valamint a természettudományos teljesítmény eredményeit mutatja be. Az empirikus munka a 13 éveseket (FIMS és SIMS), a 14 éveseket (FISS és SISS) és a nyolcadikos diákokat (TIMSS ciklusok) reprezentáló populációk adatain alapul.

Az első fázisú vizsgálatok adatait a későbbi felmérések adataihoz képest eltérően dolgozták fel. A korai felmérések adatait és dokumentációját frissített formátumokban a Center for Comparative Analysis of Educational Achievement (COMPEAT) projekt¹ tette elérhetővé. A TIMSS ciklusokhoz tartozó adatokat és dokumentációt az IEA Study Data Repository-ről² töltöttük le.

Eljárások

Az összekapcsolásról részletesen lásd Majoros (2023) és Majoros és munkatársai (2022). Jelen tanulmányban csak az értelmezéshez szükséges részletekre térünk ki. Az eljárások során alkalmazott eszközök SPSS 25 (IBM Corp., 2017) és különböző R (R Core Team, 2025) csomagok voltak.

Kognitív trendskálák

Az itemparaméterek egyidejű kalibrálását végeztük el a FIMS és SIMS itemválaszokon, melyen belül a horgonyteszt itemparamétereit az IEA által jelentett TIMSS 1995-ös értékekhez rögzítettük. Ezeket az itemparamétereket az 1999-es felmérésben végzett újra skálázási eljárás után publikálták (Martin et al., 2020). Ezután a tanulók képességeit külön becsültük meg a FIMS és a SIMS esetében, öt PV-t rendelve a diákokhoz. A tanulók képességbecsléseinek a

¹ <https://www.gu.se/en/center-for-comparative-analysis-of-educational-achievement-compeat>

² <https://www.iea.nl/data-tools/repository/timss>

TIMSS trendskálán való elhelyezéséhez a jelentett TIMSS 1995-ös skálázáshoz használt eredeti transzformációs állandókat alkalmaztuk. A természettudományos teljesítményskála felállításakor ugyanezt az eljárást alkalmaztuk.

Az országos szintű átlagok kiszámításakor az egyénenként becsült öt PV-t rendszerszinten aggregáltuk Rubin (1987) szabályait követve az adathalmazok közötti összevonásra vonatkozóan, amelyeket a Gonzalez (2014) által felvázolt többszörös imputációval értünk el. Ez az eljárás biztosítja, hogy mind a mintavételi zajt, mind a hiányzó adatokból adódó bizonytalanságot figyelembe vegyük az átlagok standard hibáinak kiszámításakor. Az átlagokat szenátusi súlyokkal (senate weight) súlyoztuk, amelyek összege minden ország adatai esetében 500. A TIMSS adminisztrációk átlageredményeit ugyanezzel az eljárással számoltuk ki.

Motivációs trendskálák

A motivációs skálák megalkotásakor egy ún. piaci kosár megközelítést alkalmaztunk. A piaci kosár megközelítés feltételezi, hogy az értékelésbe vagy felmérésbe bevont itemek meghatározzák a mért konstrukciót. Ebben az esetben a feltételezés az volt, hogy a különböző adminisztrációkban szereplő kérdőív itemek meghatározzák az egyes mérendő konstrukciókat, és reprezentatív itemek piaci kosarának tekinthetők. Az összes adminisztrációt egyidejűleg tekintve a hiányzó válaszok a kérdőívek ciklusok közötti változásainak következményei.

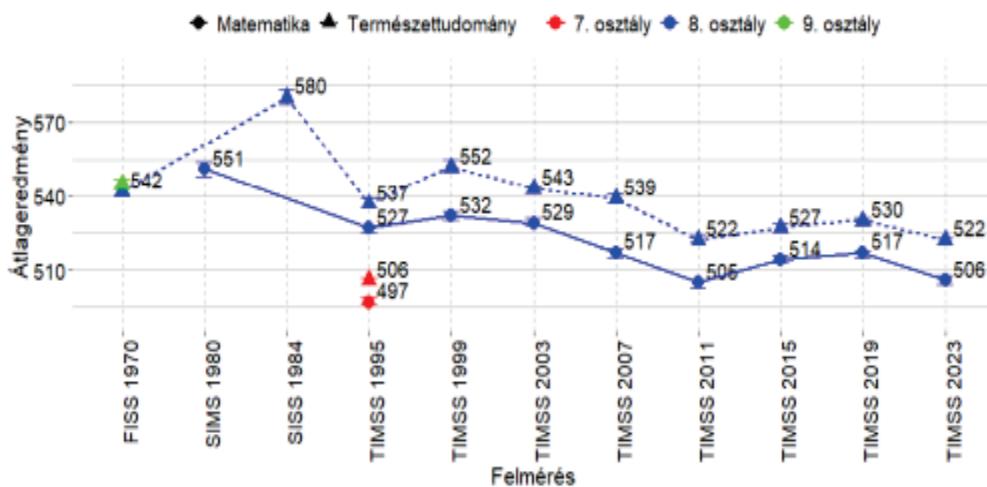
Országokonként egy mérési modellt alkalmaztunk a hiányzó válaszokat pótló, valószínűsíthető válaszok generálására, a Zwitser és munkatársai (2017) által javasolt eljárást követve. A mérési modellt minden országra külön illesztettük, ezzel figyelembe véve az országok közötti lehetséges különbségeket. A mérési modell a GPCM modell volt a TIMSS eljárással való összhang érdekében. A mérési modellek illesztésével becsült itemparaméterek felhasználásával a hiányzó válaszokat válaszdónként ötször imputáltuk. Ezután kiszámítottuk az egyéni összpontszámokat, így diákonként öt PV-t becsülve. Végül egy 5 átlagú és 1 szórású skálává transzformáltuk ezeket a pontszámokat, és a fent leírt aggregálást alkalmaztuk az országos átlagok kiszámításához.

Eredmények és értelmezésük

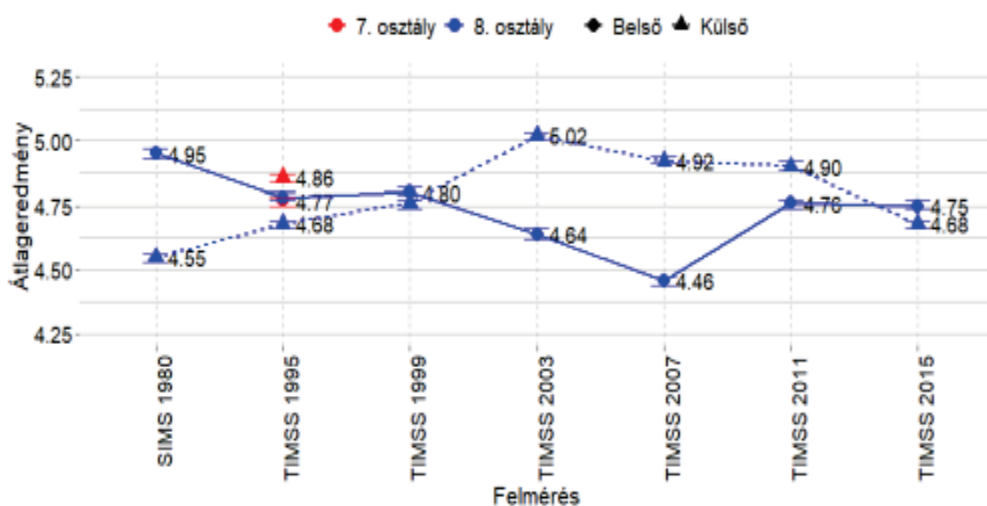
Kognitív trendek

A nyolcadik osztályos matematikai és természettudományos átlagok alakulását 1970 és 2023 között az 1. ábra mutatja be, iskolai évfolyam szerint. A matematikai korai eredményeket, tehát a SIMS mérésen elért eredményeket részletesebben tárgyalta Csikos és Vidákovich (2012). Az összekapcsolt skálára helyezett eredmény világosan kiemelkedik a későbbi értékektől (551). A korai természettudományos eredményeket (FISS, SISS) részletesebben lásd B. Németh és munkatársai (2012) munkájában. Összességében elmondható, hogy a magyar diákok kiemelkedően jól teljesítettek a korai kognitív teszteken. A természettudományos eredmények a TIMSS felmérések kezdete óta magasabbak, mint a matematikai eredmények.

A hosszútávú trendek alakulásának elemzése összetett feladat, mely során figyelembe kell venni többek között a rendszerszintű reformokat, a tantervekben bekövetkezett változásokat, illetve a mérőeszközök illeszkedésének mértékét a magyar tantervhez viszonyítva. Az IEA-vizsgálatok elméleti háttere a részt vevő országok tantervei, valamint a helyi szakértők által megjelölt kézikönyvek elemzésén alapul (B. Németh et al., 2012), azonban a részt vevő országok számának növekedése értelemszerűen hatással van erre az illeszkedésre.

1. ábra*A magyar kognitív trendek 1970 és 2023 között***A matematikatanulási motivációk trendjei**

A nyolcadikos matematikatanulási motivációs átlagok alakulását 1980 és 2015 között a 2. ábra mutatja be, iskolai évfolyam szerint. Eccles és Wigfield (2002) modelljét követve a motiváció forrása alapján különböztettük meg és soroltuk be a tanulói kérdőívek itemeit. Az instrumentális okokat (pl. jutalom, szülői elismerés) érintő kérdéseket a külső (extrinsic) motiváció, a matematikatanulás élvezetét érintő kérdéseket pedig a belső (intrinsic) motiváció fogalmához társítottuk.

2. ábra*A magyar matematikatanulási motivációs trendek 1980 és 2015 között*

Ami különösen érdekes az átlageredmények áttekintésekor, hogy a belső motivációs trend hasonló a kognitív eredmények alakulásának trendjéhez, míg a külső motivációs trend közel ellenkezőleg alakult 2015-ig. A fent említett okokból itt nem vállalkozunk a trendek elemzésére.

Korlátok

A skálák használatakor figyelembe kell venni a mintavételi különbségeket. Például Strietholt és munkatársai (2013) kidolgoztak egy korrekciós modellt az országok és az IEA olvasási felmérések közötti összehasonlíthatóság javítására életkor és iskolázottság tekintetében. Az időbeli és országonkénti különbségek figyelembevételének egy másik lehetséges megközelítése az életkor és az évfolyam magyarázó változóként való kezelése.

Számos tényező befolyásolja az összekapcsolás minőségét. Ezek a tényezők magukban foglalják az értékelések közötti hasonlóság mértékét, a konstrukciók stabilitását tartalmi, jelentésbeli és kontextusbeli szempontból, valamint a horgonyitemek számát és viselkedését. Bizonyos fokú bizonytalanság lehet ezen szempontok értékelése során, mivel az összekapcsolás régi adatokat tartalmaz.

A longitudinális teljesítményskálák egyik lehetséges korlátja az országon belüli összehasonlíthatóságban rejlik a felmérések célpopulációi miatt. Az IEA mintavételi stratégiájának változását a homogén minták lehető legjobb közelítésével kezelték. Az új skálapontszámokat használó további elemzésekben az életkor és az évfolyam kontrollváltozóként kezelhető.

A TIMSS 1995-höz kapcsolódó horgonytesztekben található közös elemek száma további aggodalomra ad okot. Az affektív skálákban a horgonyitemek aránya jóval kisebb, mint a kognitív skálák esetében. Azonban az alkalmazott egyidejű kalibrációs eljárás biztosítja a legjobb megközelítést erre az esetre, ahogy arra Wingersky és Lord (1984) rámutat.

Egy másik korlátozás a hiányzó válaszok különböző típusainak kódolására és kezelésére vonatkozik a teljesítménytesztekben. Az első szakaszbeli tesztekben a hiányzó válaszok „el nem ért” típusát nem különböztették meg, ellentétben a TIMSS skálázási eljárással. Azonban lehetséges lenne ezt a megkülönböztetést elvégezni, és feltárni az eredményekre gyakorolt hatást.

A motivációs felméréseket a közelmúltig nem trendmérésre tervezték. Idővel módosítások történtek, pl. a válaszlehetőségek száma 1995-től megváltozott ötről négyre. A középső, neutrális opció kezelésének módja a SIMS-ben korlátozza a skálákat. Ezenkívül az alkalmazott összekapcsolási eljárások nem vették figyelembe a motivációs pontszám eloszlásának időbeli különbségeit, mivel longitudinálisan egycsoportos modellezést alkalmaztak.

Az affektív trendskálák egy másik korlátja, hogy az átlagok standard hibái alábecsültek. Ennek az az oka, hogy a TIMSS-ben alkalmazott rétegzett, többlépcsős mintavételi terv miatt a becslések standard hibáinak kiszámításához használt eljárásban feltételezett egyszerű véletlenszerű mintavétel nem alkalmazható (Rutkowski et al., 2010).

Végül az egyik legnagyobb kihívást jelentő fennmaradó kérdés az, hogy a kérdések megfogalmazásának változásai befolyásolják-e a motivációs indikátorok közötti belső kapcsolatot (pl. faktorszerkezet). A kérdőív itemek száma szinte minden időpontban változott, ami megnehezíti ennek a kérdésnek a vizsgálatát.

A skálák lehetséges alkalmazásai

Az itt ismertetett skálák lehetőséget teremtenek hosszú távú változások összehasonlító vagy országokon belüli vizsgálatára. Az első szakaszbeli felmérések kognitív skálái nyilvánosan elérhetők a COMPEAT weboldalon³, a skálák összekapcsolásának dokumentációjával együtt. A rendszerszintű változások időt vesznek igénybe, ezért az iskolareformok értékeléséhez hosszú távú adatokra van szükség. Továbbá, az ok-okozati kutatási kérdések megválaszolására szolgáló hatékony statisztikai megközelítések alkalmazhatók a rendszerszintű longitudinális adatokra.

Mindazonáltal megfigyelési adatokból érvényes ok-okozati következtetések levonása kihívást jelent (ld. pl. Allardt, 1990; Rutkowski & Delandshere, 2016). A megfigyelési adatok ebben az összefüggésben olyan adatokat jelentenek, amelyeket a tanulók kezelésének vagy csoportosításának ellenőrzése nélkül gyűjtöttek, mint az IEA felmérések. Számos kutató tett javaslatokat az ilyen adatokat alkalmazó eljárásokra (ld. pl. Gustafsson, 2008; Gustafsson & Nilsen, 2022; Robinson, 2013; Schlotter et al., 2014).

Ilyen kutatási terület például az oktatáspolitikák társadalmi-gazdasági teljesítményszakadéokra gyakorolt hatása. Strietholt és munkatársai (2019) áttekintették a nemzetközi összehasonlító szakirodalmat, és megállapították, hogy a meglévő kutatások nagy része leíró jellegű, és csupán keresztmetszeti adatokon alapuló korrelációkat becsl. Az első szakaszbeli IEA-felmérések társadalmi-gazdasági háttérmutatóinak feltérképezésére irányuló további kutatások a teljesítményskálákkal kombinálva potenciálisan hozzájárulhatnak ehhez a vizsgálati irányhoz. A longitudinális skálák kihasználásának egy másik lehetséges területe a globális oktatási reformmozgalommal kapcsolatos kérdések (Fuller & Stevenson, 2019; Sahlberg, 2016). Ilyen kapcsolódó jelenségek például a privatizáció, a szabad iskolaválasztás, az iskolai verseny vagy a tanárképzés (Tóth & Sipos, 2014).

Irodalom

- Adams, R. J., Wu, M., & Macaskill, G. (1996). Scaling methodology and procedures for the mathematics and science scales. In M. O. Martin & D. L. Kelly (Eds.), *Third international mathematics and science study technical report* (Vol. 2, pp. 111–146). TIMSS & PIRLS International Study Center, Boston College.
- Afrassa, T. M. (2005). Monitoring mathematics achievement over time: A secondary analysis of FIMS, SIMS and TIMS: A Rasch analysis. In C. Alagumalai, D. David, & N. Hungi (Eds.), *Applied Rasch measurement: A book of exemplars* (pp. 61–77). Springer. https://doi.org/10.1007/1-4020-3076-2_4
- Allardt, E. (1990). Challenges for comparative social research. *Acta Sociologica*, 33(3), 183–193. <https://doi.org/10.1177/000169939003300302>
- B. Németh, M., Korom, E., & Nagy, L. (2012). A természettudományos tudás nemzetközi és hazai vizsgálata. In B. Csapó (Ed.), *Mérlegen a magyar iskola* (pp. 131–190). Nemzeti Tankönyvkiadó.
- Beaton, A. E., & Johnson, E. G. (1992). Overview of the scaling methodology used in the national assessment. *Journal of Educational Measurement*, 29(2), 163–175. <https://doi.org/10.1111/j.1745-3984.1992.tb00372.x>
- Brown, T. A. (2015). *Confirmatory factor analysis for applied research* (Second edition). The Guilford Press.
- Chmielewski, A. K. (2019). The global increase in the socioeconomic achievement gap, 1964 to 2015. *American Sociological Review*, 84(3), 517–544. <https://doi.org/10.1177/0003122419847165>

³ <https://www.gu.se/en/center-for-comparative-analysis-of-educational-achievement-compeat/mathematics-and-science>

- Cronbach, L. J. (1960). *Essentials of psychological testing* (2nd ed.). Harper & Row.
- Csikós, Cs., Pásztor, A., Rausch, A., & Sztányi, J. (2020). A matematikai nevelés kutatásának aktuális irányzatai. *Magyar Tudomány*, 181(1), 24–33. <https://doi.org/10.1556/2065.181.2020.1.3>
- Csikós, Cs., & Vidákovich, T. (2012). A matematikatudás alakulása az empirikus vizsgálatok tükrében. In B. Csapó (Ed.), *Mérlegen a magyar iskola* (pp. 83–130). Nemzeti Tankönyvkiadó.
- Dorans, N. J., Moses, T. P., & Eignor, D. R. (2011). Equating test scores: Toward best practices. In A. A. von Davier (Ed.), *Statistical models for test equating, scaling, and linking* (pp. 21–42). Springer. https://doi.org/10.1007/978-0-387-98138-3_2
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53, 109–132. <https://doi.org/10.1146/annurev.psych.53.100901.135153>
- Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. Psychology Press.
- Foy, P., Fishbein, B., von Davier, M., & Yin, L. (2020). Implementing the TIMSS 2019 scaling methodology. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and procedures: TIMSS 2019 technical report* (pp. 12.1–12.146). TIMSS & PIRLS International Study Center, Lynch School of Education, Boston College.
- Fuller, K., & Stevenson, H. (2019). Global education reform: Understanding the movement. *Educational Review*, 71(1), 1–4. <https://doi.org/10.1080/00131911.2019.1532718>
- Furr, R. M., & Bacharach, V. R. (2014). *Psychometrics: An introduction* (2nd ed.). SAGE Publications.
- Gonzalez, E. J. (2014). Calculating standard errors of sample statistics when using international large-scale assessment data. In R. Strietholt, W. Bos, J.-E. Gustafsson, & M. Rosén (Eds.), *Educational policy evaluation through international comparative assessments* (pp. 59–73). Waxmann Verlag.
- Gustafsson, J.-E. (2008). Effects of international comparative studies on educational quality on the quality of educational research. *European Educational Research Journal*, 7(1), 1–17. <https://doi.org/10.2304/eeerj.2008.7.1.1>
- Gustafsson, J.-E., & Nilsen, T. (2022). Methods of causal analysis with ILSA data. In T. Nilsen, A. Stancel-Piątak, & J.-E. Gustafsson (Eds.), *International handbook of comparative large-scale studies in education* (pp. 1–28). Springer International Publishing. https://doi.org/10.1007/978-3-030-38298-8_56-1
- Hambleton, R. K., & Jones, R. W. (1993). Comparison of classical test theory and item response theory and their applications to test development. *Educational Measurement: Issues and Practice*, 12(3), 38–47. <https://doi.org/10.1111/j.1745-3992.1993.tb00543.x>
- Hambleton, R. K., Swaminathan, H., & Rogers, H. J. (1991). *Fundamentals of item response theory*. SAGE Publications.
- Hanushek, E. A., & Woessmann, L. (2012). Do better schools lead to more growth? Cognitive skills, economic outcomes, and causation. *Journal of Economic Growth*, 17(4), 267–321. <https://doi.org/10.1007/s10887-012-9081-x>
- Holland, P. W., & Dorans, N. J. (2006). Linking and equating. In R. L. Brennan (Ed.), *Educational measurement* (pp. 187–220). Praeger Publishers.
- IBM Corp. (2017). *IBM SPSS Statistics for Windows*. IBM Corp.
- Jöreskog, K. G. (1969). A general approach to confirmatory maximum likelihood factor analysis. *Psychometrika*, 34(2), 183–202. <https://doi.org/10.1007/bf02289343>
- Kline, R. B. (2016). *Principles and practice of structural equation modeling*. Guilford Press.
- Kolen, M. J. (2006). Scaling and norming. In R. L. Brennan (Ed.), *Educational measurement* (pp. 155–186). Praeger Publishers.
- Kolen, M. J., & Brennan, R. L. (2014). *Test equating, scaling, and linking: Methods and practices* (3. ed.). Springer.
- Linn, R. L. (1993). Linking results of distinct assessments. *Applied Measurement in Education*, 6(1), 83–102. https://doi.org/10.1207/s15324818ame0601_5
- Majoros, E. (2023). Linking the first- and second-phase IEA studies on mathematics and science. *Large-Scale Assessments in Education*, 11(1), 14. <https://doi.org/10.1186/s40536-023-00162-y>
- Majoros, E., Christiansen, A., & Cuellar, E. (2022). Motivation towards mathematics from 1980 to 2015: Exploring the feasibility of trend scaling. *Studies in Educational Evaluation*, 74, 101174. <https://doi.org/10.1016/j.stueduc.2022.101174>

- Majoros, E., Rosén, M., Johansson, S., & Gustafsson, J.-E. (2021). Measures of long-term trends in mathematics: Linking large-scale assessments over 50 years. *Educational Assessment, Evaluation and Accountability*, 33(1), 71–103. <https://doi.org/10.1007/s11092-021-09353-z>
- Marco, G. L. (1977). Item characteristic curve solutions to three intractable testing problems. *Journal of Educational Measurement*, 14(2), 139–160. <https://doi.org/10.1111/j.1745-3984.1977.tb00033.x>
- Martin, M. O., & Kelly, D. L. (Eds.). (1996a). *Third international mathematics and science study technical report: Design and development* (Vol. 1). TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., & Kelly, D. L. (Eds.). (1996b). *Third international mathematics and science study technical report: Implementation and analysis* (Vol. 2). TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., Mullis, I. V. S., Foy, P., Arora, & Alka (2012). Creating and interpreting the TIMSS and PIRLS 2011 context questionnaire scales. In M. O. Martin & I. V. S. Mullis (Eds.), *Methods and procedures in TIMSS and PIRLS 2011*. TIMSS & PIRLS International Study Center, Boston College. <https://timssandpirls.bc.edu/methods/t-context-q-scales.html>
- Martin, M. O., Mullis, I. V. S., Hooper, M., Yin, L., Foy, P., & Palazzo, L. (2016). Creating and interpreting the TIMSS 2015 context questionnaire scales. In M. O. Martin, I. V. S. Mullis, & M. Hooper (Eds.), *Methods and procedures in TIMSS 2015* (pp. 15.1–15.312). TIMSS & PIRLS International Study Center, Boston College.
- Martin, M. O., von Davier, M., & Mullis, I. V. S. (Eds.). (2020). *Methods and procedures: TIMSS 2019 technical report*. TIMSS & PIRLS International Study Center, Boston College.
- Mazzeo, J., Lazer, S., & Zieky, M. J. (2006). Monitoring educational progress with group-score assessments. In R. L. Brennan (Ed.), *Educational measurement* (4th ed., pp. 681–699). Praeger Publishers.
- Mazzeo, J., & von Davier, M. (2013). Linking scales in international large-scale assessments. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment: Background, technical issues, and methods of data analysis* (pp. 230–258). CRC Press. <https://doi.org/10.1201/b16061-15>
- Meade, A. W., & Lautenschlager, G. J. (2004). A comparison of item response theory and confirmatory factor analytic methodologies for establishing measurement equivalence/invariance. *Organizational Research Methods*, 7(4), 361–388. <https://doi.org/10.1177/1094428104268027>
- Meredith, W. (1993). Measurement invariance, factor analysis and factorial invariance. *Psychometrika*, 58(4), 525–543. <https://doi.org/10.1007/BF02294825>
- Meredith, W., & Teresi, J. A. (2006). An essay on measurement and factorial invariance. *Medical Care*, 44(11), S69–S77. <https://doi.org/10.1097/01.mlr.0000245438.73837.89>
- Mislevy, R. J. (1991). Randomization-based inference about latent variables from complex samples. *Psychometrika*, 56(2), 177–196. <https://doi.org/10.1007/BF02294457>
- Mislevy, R. J. (1992). *Linking educational assessments: Concepts, issues, methods, and prospects*. Educational Testing Service, Policy Information Center. <https://eric.ed.gov/?id=ED353302>
- Mislevy, R. J., Beaton, A. E., Kaplan, B., & Sheehan, K. M. (1992). Estimating population characteristics from sparse matrix samples of item responses. *Journal of Educational Measurement*, 29(2), 133–161. <https://doi.org/10.1111/j.1745-3984.1992.tb00371.x>
- Molnár, Gy. (2006). A Rasch-modell alkalmazása a társadalomtudományi kutatásokban. *Iskolakultúra*, 16(12), 99–113.
- OECD. (2019). *TALIS 2018 technical report*. OECD Publishing. http://www.oecd.org/education/talis/TALIS_2018_Technical_Report.pdf
- Putnick, D. L., & Bornstein, M. H. (2016). Measurement invariance conventions and reporting: The state of the art and future directions for psychological research. *Developmental Review*, 41, 71–90. <https://doi.org/10.1016/j.dr.2016.06.004>
- R Core Team. (2025). *R: A language and environment for statistical computing*. <https://www.R-project.org/>

- Robinson, J. P. (2013). Causal inference and comparative analysis with large-scale assessment data. In L. Rutkowski, M. von Davier, & D. Rutkowski (Eds.), *Handbook of international large-scale assessment* (pp. 521–545). CRC Press.
- Rubin, D. B. (1987). *Multiple imputation for nonresponse in surveys*. John Wiley & Sons, Inc. <https://doi.org/10.1002/9780470316696>
- Rutkowski, D., & Delandshere, G. (2016). Causal inferences with large scale assessment data: Using a validity framework. *Large-Scale Assessments in Education*, 4(1), Article 6. <https://doi.org/10.1186/s40536-016-0019-1>
- Rutkowski, L., Gonzalez, E., Joncas, M., & von Davier, M. (2010). International large-scale assessment data: Issues in secondary analysis and reporting. *Educational Researcher*, 39(2), 142–151. <https://doi.org/10.3102/0013189x10363170>
- Sahlberg, P. (2016). The global educational reform movement and its impact on schooling. In K. E. Mundy, A. Green, B. Lingard, & A. Verger (Eds.), *The handbook of global education policy* (pp. 128–144). John Wiley & Sons, Ltd. <https://doi.org/10.1002/9781118468005.ch7>
- Schlotter, M., Schwerdt, G., & Woessmann, L. (2014). Econometric methods for causal evaluation of educational policies and practices: A non-technical guide. In R. Strietholt, W. Bos, J.-E. Gustafsson, & M. Rosén (Eds.), *Educational policy evaluation through international comparative assessments* (pp. 95–126). Waxmann.
- Sosa Paredes, Y. K., & Andersson, B. (2025). Linking the first, second and third IEA science studies. *Studies in Educational Evaluation*, 87, 101529. <https://doi.org/10.1016/j.stueduc.2025.101529>
- Stevens, S. S. (1968). Measurement, statistics, and the schemapiric view. *Science*, 161(3844), 849–856. <https://doi.org/10.1126/science.161.3844.849>
- Strietholt, R., Gustafsson, J.-E., Hogrebe, N., Rolfe, V., Rosén, M., Steinmann, I., & Hansen, K. Y. (2019). The impact of education policies on socioeconomic inequality in student achievement: A review of comparative studies. In L. Volante, S. Schnepf, J. Jerrim, & D. Klinger, *Socioeconomic inequality and student outcomes. Education Policy & Social Inequality* (Vol. 4, pp. 17–38). Springer Singapore. https://doi.org/10.1007/978-981-13-9863-6_2
- Strietholt, R., & Rosén, M. (2016). Linking large-scale reading assessments: Measuring international trends over 40 years. *Measurement: Interdisciplinary Research and Perspectives*, 14(1), 1–26. <https://doi.org/10.1080/15366367.2015.1112711>
- Strietholt, R., Rosén, M., & Bos, W. (2013). A correction model for differences in the sample compositions: The degree of comparability as a function of age and schooling. *Large-Scale Assessments in Education*, 1(1), Article 1. <https://doi.org/10.1186/2196-0739-1-1>
- Tóth, E., & Sipos, J. (2014). A rendszerszintű mérések hatása a pedagógusok tanítási gyakorlatára: Egy tanári kérdőíves vizsgálat tanulságai. *Iskolakultúra*, 24(10), 38–53.
- Wingersky, M. S., & Lord, F. M. (1984). An investigation of methods for reducing sampling error in certain IRT procedures. *Applied Psychological Measurement*, 8(3), 347–364. <https://doi.org/10.1177/014662168400800312>
- Wu, H., & Estabrook, R. (2016). Identification of confirmatory factor analysis models of different levels of invariance for ordered categorical outcomes. *Psychometrika*, 81(4), 1014–1045. <https://doi.org/10.1007/s11336-016-9506-0>
- Yin, L., & Fishbein, B. (2020). Creating and interpreting the TIMSS 2019 context questionnaire scales. In M. O. Martin, M. von Davier, & I. V. S. Mullis (Eds.), *Methods and procedures* (pp. 16.1–16.331). TIMSS & PIRLS International Study Center, Boston College.
- Zwitser, R. J., Glaser, S. S. F., & Maris, G. (2017). Monitoring countries in a changing world: A new look at DIF in international surveys. *Psychometrika*, 82(1), 210–232. <https://doi.org/10.1007/s11336-016-9543-8>

ABSTRACT

TREND SCALING METHODS: LINKING RECENT AND OLDER IEA STUDIES ON MATHEMATICS AND SCIENCE

Erika Majoros

Keywords: linking scales, international large-scale assessments, mathematics achievement, science achievement, mathematics motivation, TIMSS

One of the international large-scale assessments administered by the International Association for the Evaluation of Educational Achievement (IEA) is the Trends in International Mathematics and Science Study (TIMSS), which has been measuring mathematics and science achievement in grades four and eight since 1995. TIMSS was preceded by four IEA surveys, the results of which were not linked to the TIMSS scale. This study describes possible trend scaling methods and, to illustrate the achieved linkage, presents the rescaled mathematics and science performance and motivation for mathematics of Hungarian students.

Magyar Pedagógia, 126(2). 83–98. (2026)
<https://doi.org/10.14232/mped.2026.2.83>

Majoros Erika:  <https://orcid.org/0000-0003-4417-5016>
Umeå University, Svédország
SE-901 87 Umeå, Svédország
erika.majoros@umu.se



TECHNOLÓGIAALAPÚ SZÓKINCSEFEJLESZTÉS ÉS OLVASÁSERTÉS: EGY NAGYMINTÁS KONTROLLCSOPORTOS KUTATÁS EREDMÉNYEI

Magyar Andrea¹, Habók Anita^{1,2}, Molnár Gyöngyvér^{1,2}

¹ MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport

² Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola

Bevezetés

Az információszerzés és a társadalomban való eligazodás egyik meghatározó tényezője az értő olvasás, azaz az olvasási képesség legalább olyan minimum szintje, amely lehetővé teszi az olvasó számára az olvasott szöveg alapján egyszerű összefüggések felismerését és rövid következtetések levonását. A PISA 2022-es eredményei szerint az OECD-országok 15 éves tanulóinak 26%-a nem éri el ezt az olvasási képességszintet, azaz ők funkcionális analfabétának tekinthetők. A tanulók mintegy 26%-a mutatott kettes szint alatti teljesítményt, amely azt jelenti, hogy az adott szintű feladatokkal sikerült ugyan megbirkóznia a tanulónak, de a nagyobb kihívást jelentő feladatok megoldásánál már nehézségekbe ütközött. A magyar 15 éves diákok esetében ez az arány hasonlóan alakult, 25,9% volt 2022-ben (OECD, 2023). A gyenge olvasási képesség gyökerei gyakran a korai fejlődési szakaszokra vezethetők vissza, amely indokoltá teszi a korai diagnosztikai értékelést és a szókincs fejlesztésének szükségességét a sikeres olvasási készségek fejlesztése érdekében (Bosch et al., 2020). Számos kutatás hangsúlyozza a különböző tantárgyakhoz kötődő szókincsismeret lényeges szerepét, mint a felső tagozaton a szakszövegek olvasási készségeinek egyik legfontosabb nyelvi összetevőjét (Dujardin et al., 2023; Oakhill et al., 2019; Ouellette, 2006; Ouellette & Shaw, 2014; Sparapani et al., 2018).

Ahhoz, hogy a tanulók tantárgyi szakszókincisének fejlesztésére fejlesztőprogramot lehessen kidolgozni, előzetesen olyan korpusz létrehozása szükséges, amely a tankönyvekben előforduló szavakat gyakoriságukkal és különböző egyéb paramétereikkel együtt tartalmazza. Tudásunk szerint hazánkban jelenleg nem létezik olyan nyilvános korpusz, amely tartalmazná az általános iskola felső tagozatában az oktatásban alkalmazott tankönyvekben szereplő szakszavak listáját, és nem érhető el olyan szókincsfejlesztő program sem, amellyel ezt a szókincset osztálytermi keretek között lehetne fejleszteni.

Kutatásunk előzménye volt a fent említett korpusz létrehozása, jelen kutatásunk fókuszja pedig egy arra építő fejlesztőprogram kidolgozása és hatásvizsgálatának megvalósítása. Mind az általános iskolás diákok számára kidolgozott fejlesztőprogram, mind a fejlesztőprogram hatásvizsgálatát lehetővé tevő mérőeszköz megvalósítása számítógépes környezetben történt. A számítógépes környezet oka, hogy a mérőeszköz tekintetében azonnali diagnosztika és visszacsatolás megvalósítására törekedtünk, míg a fejlesztőprogram feladatai egyrészt több ezer szavas korpuszon alapultak, amelyek integrálását a mesterséges intelligencia alkalmazásával

algoritmusok és automatikus itemgenerálás segítségével valósítottunk meg, másrésről a számítógép-alapú rendszerrel lehetővé vált a több ezer item személyre szabott kiköszvetítése, azaz az egyes tanulók szintjén az adaptív fejlesztés megvalósítása. A fejlesztőprogram alkalmazásával tantárgyspecifikus környezetben fejleszthetik a tanulók tantárgyi szókincsüket.

Szakirodalmi háttér

A mentális lexikon és a szókincs szerepe a szövegértésben

Amikor szókincsről beszélünk, fontos különbséget tenni a szókincessel kapcsolatos különböző fogalmak között. Ezek közül az egyik legfontosabb a mentális lexikon fogalma, amelyet a legfrissebb elméletek szerint olyan összetett tudatos és tudattalan lexikai folyamatokat szabályozó rendszernek tekinthetünk, amelyben a szókincs különböző elemei összetett és folyamatosan változó hálózatot alkotnak (Dujardin et al., 2023; Jarema & Libben, 2007). A hálózatban a szavak és különböző szemantikai elemek jelentésbeli asszociációik segítségével kapcsolódnak egymáshoz (Aitchison, 2012). A szakirodalom ezeknek az elemeknek a kapcsolódását reprezentáló modelleket lényegében két alapkategóriába sorolja; az atomgömb- és a pókháló-elméletekre. Az atomgömb-hipotézis szerint a mentális lexikon különböző méretű szemantikai mezőket tartalmaz, amelyekben az egységek egynél több mezőbe is tartozhatnak. A pókháló elméletekben ezzel szemben az egyes szemantikai egységek pókhálószerűen kapcsolódnak egymáshoz, és ezáltal az egymással közvetlenül nem kapcsolódó egységek is kapcsolatba kerülhetnek egymással a pókháló összeköttetései révén (Gósy, 2005).

A szókincs ebben az összetett, dinamikusan változó hálózatos rendszerben helyezkedik el, amely Gósy (2005) szerint három egymással rugalmasan érintkező, egymásba átalakuló részből áll. A rendszer aktív része az adott személy által gyakran használt szavakat és nyelvi egységeket tartalmazza, a passzív részben tárolódnak a ritkán használt nyelvi fogalmak és jelek, míg a harmadik terület azt a szókincset fedi le, amelyet az adott kommunikáció helyzetben a beszélő éppen használ. A három terület átjárható, időben is folyton változó rendszert alkot (Gósy, 2005).

A szókincs bővülése egy életen át tartó folyamat, amely során az életkor előrehaladtával folyamatosan nő az aktivált szavak száma, és a szótudás minősége is változik. A szóismeret pontos meghatározása ezért kihívást jelent, mivel gyakran nem csupán arról van szó, hogy egy szót tudunk-e vagy sem, a szavak ismerete inkább egy spektrumon képzelhető el. Daller et al. (2007) a lexikai tér kifejezését vezették be, amely három alapvető dimenzióban határozza meg a szótudást. Az első dimenzió a szókincs terjedelme (breadth of vocabulary) vagy szókincs nagysága (vocabulary size), amely azt jelenti, hogy hány szót ismerünk. A második dimenzió a szótudás mélység (vocabulary depth), amely arra vonatkozik, hogy mennyire részletes ismerettel rendelkezünk egy-egy szóval kapcsolatban. A harmadik a lexikai folyékonyosság (lexical fluency), amely az emlékezetből a szó gyors előhívásának képességét jelöli.

A szókincs másik megközelítése arra fókuszál, hogy mit tud az egyén kezdeni egy-egy lexikai elemmel. E megközelítés alapján gyakran tesznek különbséget a receptív és produktív jártasság között (Nation, 2001). A receptív jártasság azt jelenti, hogy megértjük a szavakat, amikor halljuk vagy olvassuk őket, míg a produktív jártasság arra utal, hogy képesek vagyunk helyesen használni ezeket a szavakat beszédben vagy írásban (Nation, 2001; Nation, K. 2017; Schmitt, 2014). Pearson et al. (2007) felhívták a figyelmet arra, hogy a gyermekek receptív szókincse lényegesen nagyobb, mint produktív szókincsük. Ez azt jelenti, hogy lényegesen több szót értenek meg, mint amennyit aktívan használni tudnak beszédben.

A szókincs olvasott szövegértésben betöltött szerepét emeli ki Nagy József (2004), és a szóolvasó készség fontosságára hívja fel a figyelmet. Értelmezésében a szóolvasó készség az olvasáskészség egyik alapvető összetevője, amelynek lényege, hogy az olvasó a leírt szavakat gyorsan, pontosan és automatikusan felismeri, és ezzel azonnal aktiválja azok jelentését a mentális lexikonban. Az olvasás folyamatában ez a készség hidat képez a betűolvasás és a mondatolvasás között: a betűk és betűkapcsolatok felismerése után lehetővé teszi a szavak rutinszerű azonosítását, amely nélkül a mondatok és a szövegek megértése nem valósulhat meg. Nagy hangsúlyozza, hogy a szóolvasó készség optimális működéséhez nem elegendő a betűző olvasás, hanem szükség van a leggyakoribb magyar szavak gyors, automatikus felismerésére is. Kutatásai szerint a folyékony szóolvasás elengedhetetlen feltétele annak, hogy a tanulók elkerüljék a funkcionális analfabetizmust, és eredményesen használják az olvasást tudásszerzésre és élményszerzésre (Nagy, 2004).

Coxhead és Nation (2001) a szókincsben szereplő szavakat előfordulási gyakoriságuk alapján kategorizálták. A négy kategória közül az első a magas gyakoriságú, mindennapi életben előforduló szavakat tartalmazza, amelyek a folyószövegek közel 70%-át alkotják. A következő kategória az akadémiai és tudományos szavakat, kifejezéseket foglalja magában, amelyek a különböző tudományágakhoz tartozó szövegekben találhatók meg, és ezek a tudományos szövegek körülbelül 5%-át teszik ki. A harmadik kategória olyan szavakat tartalmaz, amelyek specifikusan egy adott tudományterülethez kapcsolódnak, és ezek a tudományos szövegek körülbelül 10%-ában fordulnak elő. Az utolsó kategóriába az alacsony gyakoriságú szavak tartoznak, amelyek ritkán jelennek meg az olvasott szövegekben. Mivel több kutatás is alátámasztotta, hogy a szövegek megértéséhez a benne szereplő szavak legalább 98%-ának ismerete szükséges, a tudományos szövegek megértése szoros összefüggésben áll az adott tudományterület szakszókincsének megfelelő szintű elsajátításával (Dujardin et al., 2023; Hu & Nation, 2000; Oakhill et al., 2019; Ouellette, 2006; Sparapani et al., 2018). Carver (1994) arra is rávilágított, ha egy szövegben a kulcsszavak 100%-a ismert, akkor a szöveg könnyűnek tekinthető, míg 99%-os ismert szó esetén átlagos nehézségűnek, 98%-os ismert szó esetén pedig nehéznek számít.

Összefoglalva, a mentális lexikon működéséről és a szókincs szerkezetéről szóló elméletek egyértelműen alátámasztják, hogy a szókincs nagysága, mélysége és az egyes szavak gyors előhívásának képessége meghatározó szerepet játszik az olvasott szövegek megértésében. Nagy József szóolvasó készségről alkotott modellje hangsúlyozza, hogy a szavak gyors és automatikus felismerése teremti meg a kapcsolatot a mentális lexikon és a szövegértés között, míg Nation és Coxhead munkái azt mutatják meg, hogy a szógyakoriság és a szókincs szervezettsége alapvetően meghatározza a megértés sikerességét.

A szakszókincs szerepe a diszciplináris oktatásban

Az iskola felső tagozatában az olvasás mint önálló tantárgy megszűnik, azonban a szövegértés továbbra is meghatározó szerepet tölt be az oktatásban, mivel integrálódik a különböző tantárgyi kontextusokba (például történelem, földrajz, természettudományok). Ennek következtében a tanulók számára az adott tantárgyi követelmények elsajátításának előfeltétele a megfelelő szövegértési kompetencia, valamint az adott diszciplína szakszókincsének ismerete (Nagy & Townsend, 2012; Józsa & Steklács, 2012). A szakirodalom hangsúlyozza, hogy a diszciplináris nyelvhasználat sajátos lexikai és szövegszervezési mintázatokkal rendelkezik, amelyek nélkül a tantárgyi tudás megszerzése és alkalmazása jelentős akadályokba ütközik (Beck et al., 2002; Moje, 2008).

A szakirodalom Beck et al. (2002) nyomán három szinthez sorolja az iskolai szakszókincset: az alapszókincs (Tier 1), az általános akadémiai szókincs (Tier 2), valamint a diszciplináris szakszókincs (Tier 3) szintjéhez. Az alapszókincs a mindennapi életben gyakran használt, kontextusfüggetlen szavakat foglalja magában. Ezek a kifejezések általában a gyermekek első nyelvi szocializációja során elsajátításra kerülnek, és nem igényelnek külön pedagógiai fókuszot a tantárgyi környezetben, az alsós olvasástanítás során főként ez a szókincs van fókuszban. A második szint, az általános akadémiai szókincs (Tier 2), amely már komplexebb kognitív műveletekhez kapcsolódik. Ide tartoznak azok a fogalmak, amelyek több tudományterületen is előfordulnak, ugyanakkor magasabb absztrakciós szintet igényelnek. Például az „elemez”, „összehasonlít” vagy „kategorizál” szavak nem kötődnek kizárólag egy diszciplinához, hanem az iskolai tudás számos területén használatosak (Nagy & Townsend, 2012). Az ilyen szókincs célzott fejlesztése kulcsfontosságú, mivel lehetővé teszi a tanulók számára, hogy értelmezni tudják a tananyag szerkezetét, logikai összefüggéseit, valamint el tudják sajátítani a tananyagot (Snow, 2010). A harmadik szint, a diszciplináris szakszókincs (Tier 3) olyan terminusokat foglal magában, amelyek kizárólag egy-egy szaktárgyhoz kötődnek, és elsősorban a tudományos ismeretek precíz kifejtését szolgálják. Ilyenek például a „fotoszintézis” a biológiában vagy a „reneszánsz” a történelemben. E fogalmak elsajátítása nem csupán lexikai tudást jelent, hanem egyben a tudományos gondolkodásmódhoz való hozzáférést is biztosítja (Moje, 2008). A pedagógiai gyakorlatban ez jelenti a legnagyobb kihívást, mivel a tanulóknak nemcsak a szóalakot és jelentést kell megtanulniuk, hanem el kell sajátítaniuk azt is, hogy az adott tudományág szakszókincsét az adott diszciplína sajátos nyelvi-kommunikációs keretein belül miként kell használniuk (Shanahan & Shanahan, 2008).

Összességében a három szint megkülönböztetése nem pusztán taxonómiai kérdés, hanem a tanulásszervezés alapvető kiindulópontja kellene, hogy legyen. Az oktatás során figyelembe kellene venni e szintek eltérő sajátosságait, és a tanulók számára biztosítani a fokozatos bevezetést a mindennapi kommunikációtól az általános akadémiai nyelvhasználaton át a szaktudományos terminológiáig. Az utóbbi évek kutatásai (pl. Colwell et al., 2023; Van Orman et al., 2021) kiemelik, hogy a diszciplináris nyelv tudatos fejlesztése nélkülözhetetlen a tantárgyi sikerességhez és az élethosszig tartó tanuláshoz. A jövőbeli kutatások és pedagógiai fejlesztések szempontjából fontos kérdés, hogy miként lehet a különböző tantárgyakban tanuló diákokat fokozatosan bevezetni a diszciplináris nyelvhasználatba, és hogyan lehet kialakítani olyan tanulásszervezési stratégiákat, amelyek támogatják a szakszókincs elsajátítását és aktív használatát.

A szókincs mérése és fejlesztése

A szókincsre vonatkozó elméletekkel összhangban a szókincs mérésének és értékelésének folyamata is összetett és többdimenziós jelleget ölt. A fent említett megkülönböztetésekkel összhangban, mint például a terjedelem vagy mélység, receptív vagy produktív, valamint diszciplináris vagy általános szakszókincs, a kutatások azt javasolják, hogy mindegyiket egyaránt mérni és értékelni szükséges (Hoffman et al., 2014; Kieffer & Lesaux, 2012). Kremmel és Schmitt (2016) is felhívják a figyelmet arra, hogy a célszavakat mind izolált formájukban, mind kontextuális környezetben mérni és értékelni kell, annak érdekében, hogy feltárjuk, olvasás közben az adott szavak valóban érthetők-e.

A szókincs fejlesztésének alapját olyan szó-korpuszok (szólisták) jelentik, amelyek az egyes szavak előfordulásait, frekvenciáját, jelentését, valamint a hozzájuk kapcsolódó morfológiai és szemantikai információkat gyűjtik össze. Ilyen lista a Coxhead Academic Word List (AWL), amely egy olyan szószedet, amely az angol nyelvben gyakran előforduló tudományos

szavakat tartalmazza. Az 570 szócsaládot tartalmazó lista Paul Coxhead (2000) kutatása alapján készült, és elsődleges célja az volt, hogy segítse az angolt mint második nyelvet tanulókat a tudományos szókincsük fejlesztésében. A listában szereplő szavak a különböző tudományos szövegekben és a tudományos írásban különösen gyakoriak, függetlenül attól, hogy milyen tudományterületen használják őket. Az AWL tehát nem tartalmaz olyan szavakat, amelyek a mindennapi kommunikációban gyakoriak, hanem kizárólag a tudományos kontextusban előforduló kulcsszavakat gyűjti össze (Coxhead, 2000). Az ilyen korpusz alapú szókincsfejlesztés segíti a tanulókat abban, hogy autentikus nyelvi környezetben fejlesszék tudásukat, így nemcsak a szavak listáját sajátítják el, hanem azok mélyebb megértését és használatát is.

Technológialapú programok a szókincsfejlesztésben

A számítógép és a különféle számítógépes alkalmazások számos módon segíthetnek a szókincs mérésében és fejlesztésében. Számítógépes rendszerek egy megadott korpusz alapján képesek online szókincstesztet generálni, amelyek különböző nehézségi szintű szavakat tartalmaznak. Az ilyen tesztek kifejezetten a receptív szókincset mérik, például a szavak jelentéseinek felismerésére irányuló kérdésekkel. A szókincs terjedelmét mérő tesztek a szavak felismerésére (például szókincs kvíz) összpontosítanak, míg a mélység mérésére olyan feladatok is alkalmazhatók, amelyek a szóhoz kapcsolódó információk részletesebb feldolgozását igénylik, például a szó többes jelentéseinek vagy szinonimáinak felismerése (Schmitt, 2014).

A szókincs mérésére kidolgozott tesztek között széles körben alkalmazott eljárás az álszavak (pseudowords) használata, amely lehetővé teszi a szóalakok gyors felismerésének és a lexikai döntésnek a vizsgálatát. Ennek egyik legismertebb formája Paul Meara YES–NO tesztje, amelyben a vizsgálati személyeknek egy szólista elemeiről kell eldönteniük, hogy valós szavak-e vagy sem (Meara, 1980). A módszer előnye, hogy rövid idő alatt megbízható becslést ad a receptív szókincs nagyságáról, és számítógépes környezetben is könnyen automatizálható.

A lexikai döntésen alapuló tesztek továbbfejlesztett változata a LexTALE (Lexical Test for Advanced Learners of English), amelyet Lemhöfer és Broersma (2012) dolgoztak ki. A teszt 60 valódi és álszó alapján méri a szókincsismeretet, online is elérhető, kitöltése mindössze néhány percet vesz igénybe, és eredményei szoros összefüggést mutatnak más, hosszabb nyelvi teljesítménymérő eljárásokkal. A LexTALE jelentősége abban áll, hogy egyszerűsége ellenére valid és standardizált eszközt kínál a szókincs gyors becslésére, ezért a pszicholingvisztikai kutatásokban széles körben alkalmazzák.

A receptív szókincs mérésének másik elterjedt típusa a képalapú szóértés-vizsgálat. Ennek legismertebb példája a Peabody Picture Vocabulary Test (PPVT), amelyet Dunn és Dunn (2007) fejlesztettek ki. A tesztben a vizsgálati személynek négy kép közül kell kiválasztania azt, amely megfelel a hallott szónak. A PPVT számítógépes változata (Q-interactive, Q-global) automatizált adminisztrációt és azonnali kiértékelést tesz lehetővé, így a digitális szókincsmérés egyik legismertebb standardizált eszközének tekinthető.

Magyarországon technológialapú programok kidolgozása elsősorban a Szegedi Tudományegyetem Neveléstudományi Műhelyéhez kötődik, ahol Nagy József kutatásait továbbfejlesztve Molnár, Csapó és Magyar diagnosztikus és számítógép-alapú mérőeszközöket dolgoztak ki a szóolvasás fejlődésének vizsgálatára (Csapó et al., 2014; Molnár & Csapó, 2013; 2019a; Molnár & Magyar, 2018). Az SZTE Oktatásméleti Kutatócsoport és MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport (MTA-SZTE DTTK) által létrehozott és folyamatosan fejlesztett eDia-rendszer új távlatokat nyitott a szóolvasó készség mérés-értékelésében. Nagy (2004) az ezredfordulón, a Magyar Értelmező Kéziszótár (Pusztai & Csábi, 2003)

alapján összeállított a köznyelvi szövegekben leggyakrabban előforduló 5000 szó szóolvasási készségének papíralapú vizsgálatára alkalmas feladatokat. A feladatok technológiai környezetbe történő adaptációját és empirikus alkalmazását követően Molnár és Magyar (2018) kidolgozták a szóolvasási készség adaptív online mérőrendszerét. Az új mérőeszköz lehetővé tette a tanulók egyéni fejlődési szintjének személyre szabott és differenciált feltérképezését (Magyar & Molnár, 2014; 2015; Molnár & Magyar, 2018).

A mesterséges intelligencia (Artificial Intelligence – AI) és a természetes nyelvek feldolgozása (Natural Language Processing – NLP) technológiák különböző szinteken képesek segíteni a szókincs fejlesztésében. Az NLP rendszerek például képesek felismerni a szavak különböző jelentéseit különböző kontextusokban, és értékelní, hogy a tanuló hogyan használja a szót adott szöveggörnyezetben (Porkoláb & Fekete, 2023). Az AI segíthet a szavak különböző szókincsdimenziókba való besorolásában (pl. gyakoriság, szinonimák, szóalakok), és képes lehet szemantikai elemzést végezni annak megértésére, hogy a szó jelentése hogyan változik a különböző kontextusokban (Jurafsky & Martin, 2013).

Egyes alkalmazások, mint a Duolingo vagy a Memrise, automatikusan személyre szabják a szókincsfejlesztést, és biztosítanak olyan szókincsgyakorlatokat, amelyek figyelembe veszik az egyes tanulók nyelvi szintjét és fejlődését. Az alkalmazások képesek a tanulót azokra a szavakra fókuszáltatni, amelyeket még nem sajátított el teljesen, és a tanulási folyamatot fokozatosan, a szókincs mélységét is figyelembe vevő módon alakítják (Hatcher & Yu, 2018). Az alkalmazások gyakran alkalmaznak *spaced repetition* (időzített ismétlés) algoritmusokat, amelyek folyamatosan fejlesztik a tanuló szókincsét, figyelembe véve az egyes szavak elsajátításában mutatott erősségeit és gyengeségeit (Behzad et al., 2017). A számítógép figyelemmel kísérheti, hogy a tanuló milyen gyakran és hogyan használ bizonyos szavakat a szövegekben, és értékelheti, hogy a szó megfelelő kontextusban és helyes formában szerepel-e. Ezzel lehetőség nyílik a produktív szókincs pontosabb mérésére (Ding et al., 2023). A gépi rendszerek továbbá hatékonyan segíthetnek a tanulónak egyénre szabott visszajelzést adni, például, hogy mely szavakban kell javulnia, melyek azok a szavak, amelyeket még nem használ aktívan, és melyek azok, amelyeket mélyebben kellene megértenie (Molnár, 2024; Porkoláb & Fekete, 2023; Szűts et al., 2022).

Napjainkban egyre inkább teret nyernek a digitális játékalapú programok, amelyek nagyban hozzájárulnak a tanulás hatékonyságához és élvezetességéhez (Csapó et al., 2012a). Ezek a programok egy adott szabályrendszeren alapján valamilyen cél elérésére irányulnak, ezáltal kihívást jelentve a játékos számára, továbbá különböző játékosított elemek, például pontok, szintek és kihívások alkalmazásával teszik vonzóbbá a tanulást. A játékosított elemek a motivációt fokozzák, ezáltal a tanulás szórakoztatóbbá válik (Csapó et al., 2012b; Pásztor, 2013). A játékosított programok további előnye, hogy azonnali visszajelzést biztosítanak a tanulónak, lehetővé téve számukra, hogy gyorsan felismerjék és kijavítsák hibáikat. Ez elősegíti a tanulás mélyebb megértését és hosszú távú memóriában történő rögzítését (Chen et al., 2018).

Összefoglalva, a számítógép és a digitális technológia segítségével tehát a szókincs mérése és fejlesztése sokkal gyorsabbá, pontosabbá és személyre szabottabbá válhat. Az automatikus szókincsfelmérők, a mesterséges intelligencia alapú elemzések, a gépi tanulás és a különböző fejlesztő alkalmazások lehetőséget adnak a tanulók szókincsének minden dimenziójában történő részletes mérésére és finomhangolására. A fent ismertetett mérőeszközök a leggyakoribb köznyelvi szavak ismeretét mérik fel, azonban a diákok sikeres iskolai előrehaladásához nélkülözhetetlen az egyes tantárgyakhoz kötődő szakszókincs ismerete, ami alapját képezi a tankönyvi szövegek megértésének. Hiányterületnek számít annak feltérképezése, hogy a különböző tantárgyak keretében hogyan, milyen sikerességgel valósítják meg a tanárok az adott tantárgy szakszókincsének fejlesztését, miközben az kulcsfontosságú a tanulók szövegértési és

kommunikációs készségeinek erősítésében. Az adott tudományterület évfolyamnak megfelelő szintű szókincsének elsajátítása elősegítheti a tanulók számára a tantárgyi tartalmak mélyebb megértését és a tudományos gondolkodás fejlesztését. A megfelelő szakszavak használata lehetővé teszi a tanulók számára, hogy pontosan és hatékonyan kommunikáljanak a különböző tantárgyi területeken (Terbe, 2019).

Kutatási előzmények

A kutatás célja egy olyan online diagnosztikus mérő- és fejlesztőprogram kidolgozása, amely az 5–8. évfolyamos tanulók tantárgyi szókincsének mérésére és célzott fejlesztésére szolgál, és megbízhatóan alkalmazható a mindennapi pedagógiai gyakorlatban. A fejlesztés elméleti és empirikus alapját egy, az MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport kutatói által az általános iskolai tankönyvekben előforduló szövegek feldolgozásával előzetesen létrehozott szókincs-adatbázis képezi, amely a magyarországi általános iskolák felső tagozatán használt hivatalos tankönyvek szógyakorisági adatait tartalmazza. Az adatbázis összesen 53 tankönyv lexikai állományát foglalja magában a matematika, irodalom, történelem, magyar nyelv és természettudomány tantárgyak köréből. A korpuszban szereplő több ezer lexéma gyakorisági elemzését követően a mérő- és fejlesztőfeladatok hierarchikus struktúráját úgy alakítottuk ki, hogy a tanulók a legnagyobb előfordulási gyakoriságú szavaktól fokozatosan haladjanak a ritkábban előforduló, ezáltal nagyobb nyelvi kihívást jelentő elemek felé.

A korpuszelemzés eredményei azt mutatják, hogy az 5. évfolyamos tankönyvekben 5619 különböző szó fordul elő, összesen 314 770 előfordulással. Ez a lexikai mennyiség az évfolyamok előrehaladtával fokozatosan növekszik, 8. évfolyamra több mint 533 ezer szóelődulással. Bár az egyes évfolyamok szókincse között jelentős átfedés figyelhető meg, az adatok jól tükrözik a tanulókra nehezedő fokozódó lexikai terhelést és a tantárgyi nyelvhasználat egyre nagyobb komplexitását (1. táblázat).

1. táblázat

A korpuszban szereplő szavak évfolyamok szerint

Évfolyam	Különböző szavak száma	Összes előfordulás
5	5 619	314 770
6	5 797	327 675
7	7 004	528 412
8	7 134	533 089

Kutatási kérdések

A kidolgozott fejlesztőprogram hatékonyságát eltérő tanulói csoportok körében vizsgáltuk, különös figyelmet fordítva a különböző képességszintű tanulókra. A következő kutatási kér-

dések megválaszolását tűztük ki célul: (1) Mennyire hatékonyan fejleszthető az 5–8. évfolyamos tanulók szókincse az online fejlesztőprogrammal? (2) Van-e különbség évfolyamonként a fejlesztés hatékonyságában? (3) Milyen hatékonysággal alkalmazható a fejlesztőprogram a különböző szókincsszinttel rendelkező tanulók körében?

Módszerek

Minta

Az általános iskolák önkéntesen regisztrálhattak a programba, ennek eredményeként összesen 123 általános iskola 5025 tanulója kapcsolódott be a vizsgálatba. A fejlesztésben való részvétel önkéntes volt. Az érintett osztályokból kontroll- és kísérleti csoportokat alakítottunk ki. A kutatás kvázi-kísérleti elrendezésben zajlott: mind a kísérleti, mind a kontrollcsoportot két időpontban mértük fel, a beavatkozás előtt és azt követően. A vizsgálatba bevont tanulói minta kialakítása a következő kritériumok szerint történt: (1) azok a tanulók, akik nem vettek részt mind az előmérésben, mind az utómérésben, kikerültek a végső adatbázisból; (2) a kísérleti csoportból kizártuk azokat a tanulókat, akik a fejlesztőmodulok legalább 85%-át nem teljesítették; (3) eltávolítottuk azokat a tanulókat is, akik ugyan kitöltötték a teszteket, de nem töltöttek elegendő időt az instrukciók elolvasására és a feladatok elvégzésére; (4) végül, a pedagógusok egyéni jellemzőinek és tanítási módszereinek torzító hatását minimalizálандó, a tanulókat *propensity score matching* módszerrel párosítottuk a bemeneti mérési eredmények, valamint az iskola, osztály és évfolyam adatai alapján. Ezt követően a párokat véletlenszerűen osztottuk be a kísérleti, illetve a kontrollcsoportba.

Az adatok tisztítása után összesen 60 iskola 1562 tanulója került be az elemzésekbe, fele a kísérleti, fele pedig a kontrollcsoportban. A diákok átlagéletkora 13,33 év volt ($SD = 0,62$ év; ld. 2. táblázat).

2. táblázat

A minta mérete, kor szerinti eloszlása

Évfolyam	Fejlesztőcsoport (n)	Kontrollcsoport (n)	Fejlesztőcsoport életkora Átlag (M) Szórás (SD)	Kontrollcsoport életkora Átlag (M) Szórás (SD)
5	242	242	11,90 (0,74)	11,84 (0,66)
6	193	193	12,89 (0,65)	12,97 (0,72)
7	243	243	13,84 (0,62)	13,86 (0,62)
8	103	103	14,68 (0,47)	14,67 (0,47)

Mérőeszköz

A fejlesztőprogram hatásvizsgálatára a tanulók szókincsének a terjedelmét és a mélységét mérő számítógépes tesztet dolgoztunk ki. A teszt összesen 75 itemből állt. A tesztben szereplő feladatok kétharmada a tanulók szókincsét szövegtől és kontextustól függetlenül mérte, míg egyharmaduk a tanulók szókincsét konkrét szöveggörnyezetbe ágyazva vizsgálta. A szövegtől és kontextustól függetlenül mérő feladattípusnál négyféle feladat került alkalmazásra: (1) különálló szavak, amelyek egy adott képhez kapcsolódnak, és a tanulónak meg kellett állapítania, hogy a szó jelentése illik-e a képhez; (2) összetett szavak, amelyek esetében a feladat annak eldöntése volt, hogy a szó létezik-e vagy sem; (3) a szóasszociációs tesztek alapján kialakított feladat, ahol a tanulónak hat szó közül három olyan szót kellett kiválasztania, amely a megadott kulcsszó jelentéséhez a legközelebb áll; (4) az utolsó feladattípusban öt szó szerepelt, amelyről azt kellett eldönteni, hogy kapcsolódik-e valamelyik megadott definícióhoz. A második részben a kontextus segítette a tanulókat a szavak jelentésének értelmezésében. A feladatok között szerepeltek szinonimák és ellentétek keresése, valamint igaz-hamis és feleletválasztós feladatok a szöveg tartalmával kapcsolatban (1. ábra). A teszt működését előzetesen teszteltük. Az elő- és utótesztként alkalmazott mérőeszközök megbízhatóságát korábbi kutatásokban monitoroztuk (Magyar et al., 2023), ahol megállapítottuk, hogy az elő- és utótesztek megbízhatóan alkalmazhatók általános iskolás, valamint középiskolás tanulók körében. Mind az elő-, mind az utótesztek magas megbízhatósági értékeket mutattak minden évfolyamon (Cronbach $\alpha = 0,871$ – $0,921$). A teljes teszt megbízhatósága szintén magas volt mind az elő-, mind az utótesztek esetén (Cronbach $\alpha = 0,918$ és $0,932$). A rendszerbe épített automatikus értékelés biztosította az objektív pontozást. Minden diák ugyanazt a pontszámot kapta az azonos válaszáért, mivel a rendszer a korrekciós tényező alapján rendelte a pontszámokat, elősegítve ezzel a teszt objektivitását és megbízhatóságát (Csapó et al., 2014).

1. ábra

Példafeladatok a mérőtesztből

Olvasd el figyelmesen a következő szót. Keresd meg a szövegben az ellentétes jelentésű párját és kattints rá!

segítő

Valós ember - virtuális személyiség

Emberként lassan hajlamosak leszünk túlzott szerepet tulajdonítani annak, ami ebben a kis világban történik. Azt gondolhatjuk, az a tévképzet alakulhat ki bennünk, hogy mindenki hozzánk hasonlóan gondolkodik - pedig ez a legritkább esetben van így. Ezt különösen akkor fontos szem előtt tartanunk, ha egyszer csak azon kapjuk magunkat, hogy valamilyen rossz érzésünket, félelmünket, szorongásunkat, esetleg a bennünk bujkáló ártó szándékot felerősítő tartalmakkal van tele a napunk, mert újra meg újra ilyen cikkek, képek és videók kerülnek elénk.

Vissza Mehets

Fejlesztőprogram

A fejlesztőprogram négy szintje a szakirodalomban ismertetett szókinckategóriák figyelembevételével került kialakításra. A besorolás alapját Beck et al. (2002) háromszintű modellje adta, amely megkülönbözteti az alapszókinccet (Tier 1), az általános akadémiai szókinccet (Tier 2), valamint a diszciplináris szakszókinccet (Tier 3). A program szintjei a tanulók életkori sajátosságaihoz, nyelvi fejlettségéhez és az elsajátítandó fogalmak absztrakciós szintjéhez igazodtak.

Az első szinten elsősorban az alapszókinccshez (Tier 1) tartozó szavak jelentek meg. A program ezen szakasza az alsó tagozatos olvasástanítás gyakorlatához igazodva a biztos szófelismerést, a jelentésazonosítást és a szókinccs aktiválását támogatta. A második és harmadik szint már fokozatos átmenetet képezett az általános akadémiai szókinccs (Tier 2) irányába. Ide olyan szavak kerültek, amelyek több tantárgyban is használatosak, és magasabb szintű gondolkodási műveletekhez kapcsolódnak. Az olyan kifejezések, mint például az „elemez”, „összehasonlít”, „következtet” vagy „kategorizál”, a tanulók számára nem csupán nyelvi elemekként, hanem a tananyag feldolgozását segítő gondolkodási eszközökként is megjelentek. A negyedik szint a diszciplináris szakszókinccsre (Tier 3) épült, amely már egy-egy tudományterülethez szorosan kapcsolódó terminusokat foglalt magában. Ezek a természettudományos, történelmi vagy nyelvtani szakkifejezések nemcsak speciális lexikai tudást igényelnek, hanem az adott tudományterület gondolkodásmódjának és kommunikációs sajátosságainak megértését is. A program legmagasabb szintjén ezért hangsúlyt kapott a fogalmak kontextusba ágyazott használata, a definíciók értelmezése, valamint a szakszókinccs aktív alkalmazása különböző feladathelyzetekben.

A fejlesztőprogram szintjeinek kialakítása így nem pusztán nehézségi fokozatok mentén történt, hanem a szókinccs funkcionális és kognitív sajátosságainak figyelembevételével is, támogatva a tanulók fokozatos átmenetét a hétköznapi nyelvhasználattól az akadémiai és diszciplináris nyelvhasználat felé.

2. ábra

Példafeladat a fejlesztőprogramból



A program bemeneti teszttel indul, majd a szókincsfejlesztő része négy nehézségi szinten kínál feladatokat. A tanulók szintenként hat-nyolc feladattípus közül választhatnak, és a feladatokat tízes egységekben pontozza a rendszer. A pontszámok számítása a számítógépes játékokhoz hasonlóan történik: 200 pont elérésekor a tanuló továbbléphet a következő szintre. A záróteszt megoldására 800 pont megszerzése után van lehetőség. A feladatok megoldása többször is lehetséges, mivel a rendszer dinamikusan változtatja az itemeket. Egy feladatsor játékos formában 10–15 perc alatt teljesíthető, ami egyaránt alkalmassá teszi az osztálytermi, tanórán kívüli és otthoni gyakorlásra.

A fél-automatikus itemgenerálás (AI) lehetőségeinek köszönhetően bármely feladattípus tetszőlegesen bármennyiszer megoldható volt, ugyanakkor a feladatok előre meghatározott szerkezetben alapultak, amely biztosította azok megfelelő megbízhatóságát, objektivitását és validitását. A 2. ábra egy példafeladatot mutat be, a rendszer egy előre definiált adatbázis és algoritmus alapján képes különböző válaszlehetőségeket létrehozni ugyanazon szerkezetű feladathoz.

Eljárások

A kutatás módszertani kerete egy kvázi-kísérleti designra épült, amelyben a fejlesztő- és kontrollesoportok teljesítményét két mérési ponton, a fejlesztés előtt és után mértük. Az elő- és utóteszt közötti intervallum öt hónap volt. Az előteszt egy iskolai tanórát, azaz 45 percet vett igénybe, ezt követően a kísérleti csoport tagjait célzott, személyre szabott online fejlesztésben részesítettük, amelyet az iskolai órakeretben vagy tanórán kívüli tevékenységként biztosítottak a résztvevő iskolák tanárai. A képzési feladatok elvégzésére és a minijátékok gyakorlására nem volt meghatározott minimum vagy maximum időkorlát. A programban részt vevő diákok heti egy- három alkalommal foglalkoztak a fejlesztőprogrammal. Egy-egy foglalkozás időtartama 10–15 perc volt (3. ábra). Mind az elő- mind az utótesztelés, mind a fejlesztések az iskola számítógéptermeiben zajlottak online módon. Az online teszteket az eDia-rendszer, a fejlesztést az eLea-rendszer biztosította (Csapó & Molnár, 2019; Molnár & Csapó, 2019b; Molnár et al., 2019).

A kutatás célját előzetesen ismertettük a pedagógusokkal, továbbá tájékoztatást nyújtottunk a fejlesztőprogram részletes leírásáról, valamint az eDia- és eLea-rendszerek használatáról. A programhoz való hozzáféréshez egyedi azonosítókat kaptak a résztvevő iskolák, amelyek kizárólag a saját tanulóik beazonosítására szolgáltak. Az online rendszerekhez való hozzáférés csak a meghatározott időszakban, az iskoláknak küldött webes hivatkozáson keresztül volt elérhető. A tanulók és szüleik írásbeli beleegyezése az iskola közreműködésével történt. A kutatást a Szegedi Tudományegyetem Neveléstudományi Doktori Iskola Etikai Bizottsága hagyta jóvá.

Alkalmazott statisztikai eljárások

A mérőteszt belső konzisztenciájának vizsgálatához a Cronbach-alfa mutatót alkalmaztunk. Az elő- és utóteszt, valamint a vizsgálati csoportok közötti teljesítménykülönbségek feltárására t próbát, valamint kétutas ismételt méréseket ANOVA-t (*Mixed-design ANOVA*) használtunk. Először Levene-próbával vizsgáltuk a csoportok varianciáinak homogenitását, vagyis azt, hogy a kísérleti és kontrollesoport szórása statisztikailag különbözik-e egymástól. A két csoport közötti teljesítményeltérések mértékének számszerűsítésére a hatásméretet (effect size) számítottuk, amelyet a részleges eta-négyzet (partial η^2) mutatóval határoztuk meg. Ez az érték azt jelzi, hogy az adott tényező (pl. az idő, a csoport vagy azok interakciója)

a teljes variancia mekkora hányadát magyarázza. A szakirodalomban elfogadott irányelvek szerint a 0,01 alatti érték kismértékű, a 0,06 körüli közepes, míg a 0,14 feletti nagymértékű hatást jelez (Cohen, 1988). A csoportok közötti különbségek nagyságát Cohen's d -vel is jellemeztük. Cohen (1988) iránymutatása szerint $d \approx 0,20$ kis hatás, $d \approx 0,50$ közepes hatás, és $d \approx 0,80$ nagy hatást jelez. A páros (elő–utó) összehasonlításoknál a páros hatásmértéket (paired Cohen's d) tekintettük. Minden hatásmértéket 95% CI -vel jeleztünk.

Eredmények

(1) A fejlesztőprogram hatékonyságának vizsgálata

A program hatékonyságát kontrollcsoportos vizsgálattal mértük, az elő- és utótesztek teljesítményeredményeinek összehasonlítása alapján. Az előtesztre vonatkozó eredmények alapján a Levene-statisztika értéke 0,060 ($p = 0,806$), ami nem szignifikáns. Ez azt jelzi, hogy a két csoport varianciája az előteszten nem különbözött egymástól, tehát a homogenitás feltétele teljesült. Az utótesztre kapott érték (Levene-statisztika = 3,335; $p = 0,068$) szintén nem szignifikáns $p = 0,05$ szinten, így itt sem mutatható ki lényeges különbség a csoportok szórásai között, így elvégezhetőek voltak a varianciaanalízis és a t -próbák, és eredményei megbízhatóan értelmezhetők.

Az előteszten a fejlesztő- ($M = 43,74$, $SD = 16,69$) és a kontrollcsoport ($M = 43,90$, $SD = 16,55$) teljesítménye között nem mutatkozott szignifikáns eltérés ($t = 0,192$; $p = 0,848$). Az ismételt méréses kétutas varianciaanalízis eredményei alapján szignifikáns fejlesztőhatás volt kimutatható, ($F(1,1560) = 136,47$, $p < 0,001$, $\eta^2 = 0,080$) és a fejlesztési idő és a csoport közötti interakció is szignifikánsnak bizonyult ($F(1,1560) = 10,93$; $p = 0,001$; $\eta^2 = 0,007$). A fejlesztőhatás gyakorlati jelentőségének becslésére hatásméret-mutatókat számítottunk, amely a beavatkozás tényleges erejét, azaz a változás gyakorlati jelentőségét fejezi ki. Az idő és a csoport interakciójára vonatkozó részleges eta-négyzet értéke $\eta^2 = 0,007$ volt, ami kismértékű, de statisztikailag szignifikáns hatást jelez. A fejlesztő- és kontrollcsoport utóteszt-eredményeinek különbségét a Cohen-féle hatásmértékkel is becsültük. A számítás szerint Cohen's $d = 0,10$ (95% CI [-0,00; 0,20]), amely a szakirodalmi értelmezés szerint nagyon kis mértékű hatásnak tekinthető. Az eredmény ugyanakkor arra enged következtetni, hogy a fejlesztőprogram alkalmazása minimális, de pozitív irányú elmozdulást eredményezett a tanulók teljesítményében a kontrollcsoportéhoz képest.

Az ismételt méréses varianciaanalízis csoportokon belüli hatásokat mutató táblázatának eredményei alapján a csoportfőhatás nem érte el a statisztikai szignifikancia szintjét ($F(1,1560) = 0,860$; $p = 0,354$; $\eta^2 = 0,001$). Ez arra utal, hogy a kísérleti és a kontrollcsoport teljesítménye között az egész mérési időszakra vetítve nem volt kimutatható különbség. A rendkívül alacsony parciális eta-négyzet érték ($\eta^2 = 0,001$) azt jelzi, hogy a csoporttagság a teljes variancia mindössze 0,1%-át magyarázza, tehát a csoporthatás gyakorlati szempontból elhanyagolható.

Az intervenciót követően a fejlesztő- ($M = 47,95$; $SD = 17,28$) és a kontrollcsoport ($M = 46,25$; $SD = 18,42$) teljesítménye hasonló szórást jelzett. A Levene-teszt az utótesztre vonatkozóan a szignifikanciaszinthez közeli eredményt mutatott ($F(1,1560) = 3,335$; $p = 0,068$), ami a csoportok közötti varianciakülönbség tendenciájára utal.

(2) A fejlesztés évfolyamonkénti különbségeinek vizsgálata

A második kutatási kérdésünk az évfolyamonkénti eltérések vizsgálatára vonatkozott. A kétutas ANOVA elemzés eredményei szerint szignifikáns főhatás mutatkozott az idő tényezőre vonatkozólag ($F(1, 1554) = 120,51, p < 0,001, \eta^2 = 0,072$), ami azt jelzi, hogy az utóteszthez viszonyítva szignifikánsan jobb eredményeket értek el a tanulók, a hatásméret közepes. A csoport és az idő interakciója is szignifikáns volt ($F(1, 1554) = 15,26, p < 0,001, \eta^2 = 0,010$), tehát a fejlesztőcsoport javulása szignifikánsan nagyobb mértékű volt, mint a kontrollcsoporté. Összességében az évfolyam és a fejlesztés interakciója nem volt szignifikáns ($p = 0,816$), vagyis az elő- és utóteszt közti javulás nem különbözött évfolyamonként. A háromutas interakció (fejlesztés $\times$ csoport $\times$ évfolyam) azonban szignifikáns eredményt mutatott ($F(3, 1554) = 2,79, p = 0,039, \eta^2 = 0,005$), ami arra utal, hogy a fejlesztés hatása a fejlesztő- és kontrollcsoportokban az évfolyam függvényében változott. Volt olyan évfolyam, ahol a fejlesztés nagyobb hatású volt, és olyan is volt, ahol kevésbé; a fejlesztőprogram hatékonysága a kontrollcsoportéhoz képest az 5. évfolyamon szignifikánsan alacsonyabb volt a 7. és 8. évfolyamhoz képest ($p = 0,001$ és $p = 0,002$), míg a 6., 7. és 8. évfolyam között nem volt szignifikáns különbség.

A 3. táblázat összegzi a négy évfolyam előteszten és utóteszten elért eredményeinek leíró statisztikai mutatóit és a Cohen's d hatásmértéket. A Cohen's d értékei minden évfolyamon 0,5 alatt maradtak, vagyis a hatások kicsik a hagyományos értelmezés szerint (Cohen, 1988). Az 5. évfolyamon gyakorlatilag nem mutatkozott különbség a két csoport teljesítménye között ($d = -0,01$), míg a 6–8. évfolyam esetében a fejlesztőcsoport enyhén jobb eredményeket ért el, különösen a legidősebb korcsoportban ($d = 0,22$).

3. táblázat

Leíró statisztikai mutatók

Évfolyam	Fejlesztőcsoport			Kontrollcsoport			Hatásmérték
	N	Előteszt $M(SD)\%$	Utóteszt $M(SD)\%$	N	Előteszt $M(SD)\%$	Utóteszt $M(SD)\%$	
5	242	37,89 (12,84)	40,77 (14,38)	24 2	37,99 (12,66)	40,99 (16,01)	-0,01 [-0,19, 0,16]
6	193	43,88 (16,48)	49,00 (15,54)	19 3	43,99 (16,35)	46,12 (17,00)	0,18 [-0,02, 0,38]
7	243	47,13 (17,79)	51,39 (18,40)	24 3	47,28 (17,75)	49,85 (19,66)	0,08 [-0,10, 0,26]
8	103	49,25 (18,41)	54,72 (18,56)	10 3	49,66 (17,91)	50,39 (20,38)	0,22 [-0,05, 0,50]

(3) A különböző szókincsű tanulók körében végzett hatékonyságvizsgálat

A második kutatási kérdés esetében megvizsgáltuk az eredményeket az egyes évfolyamokon belül és az évfolyamok között. Mivel nem mutatkoztak nagy különbségek az évfolyamok között, a továbbiakban az évfolyamokon belüli különbségeket kívántuk megvizsgálni. Ehhez a vizsgálatban részt vevő tanulókat az előteszten nyújtott kezdeti teljesítményük alapján három

csoportra osztottuk. Az első csoportba (A alcsoport) 373 tanuló került, akik gyenge teljesítményt értek el, vagyis az átlagnál több mint egy szórással alacsonyabb pontszámot szereztek (0–33%). A második csoportot (B alcsoport) 1562 tanuló alkotta, akiknek az eredménye az átlag körüli tartományba esett (33,001–63%). A harmadik csoportba (C alcsoport) 281 tanuló került, akik az átlagnál több mint egy szórással jobb eredményt értek el (63,001–100%). A 4. táblázat összefoglalja a tanulók eredményeit a három képességcsoportban.

4. táblázat

Eredmények különböző szintű tanulócsoporthoz szerinti bontásban

Csoport	N	Előteszt		Utóteszt		Hatásmérték
		M%	SD%	M%	SD%	Cohen's <i>d</i> [95% CI]
Fejlesztő A	189	23,32	4,85	30,31	8,05	1,05 [0,83; 1,26]
Kontroll A	184	23,54	4,37	30,71	12,18	0,78 [0,57; 0,99]
Fejlesztő B	453	44,53	9,70	48,34	13,33	0,33 [0,20; 0,46]
Kontroll B	455	44,36	9,66	45,74	15,18	0,11 [-0,02; 0,24]
Fejlesztő C	139	68,94	4,66	70,68	8,47	0,25 [0,01; 0,49]
Kontroll C	142	68,83	4,70	68,02	12,25	-0,09 [-0,32; 0,15]

Megjegyzés. A alcsoport = az előteszten gyengén teljesítők; B alcsoport = az előteszten átlagosan teljesítők; C alcsoport = az előteszten jól teljesítők.

Az elő- és utóteszt eredmények azt mutatják, hogy a fejlesztőprogram hatása a tanulók szókincsétől függően eltérő mértékben érvényesült. Az alacsony képességű tanulók esetében az előteszten nem mutatkozott szignifikáns különbség a kontrollcsoport és a fejlesztőcsoport teljesítménye között ($t = 0,455$, $p = 0,65$). Az utóteszten szintén nem volt kimutatható eltérés a két csoport között ($t = 0,383$, $p = 0,70$). A fejlődés mértékét szórásegységben kifejező hatás-méret alapján mindkét csoport szignifikáns fejlődést mutatott, és a két csoport átlagos fejlődése hasonló mértékűnek bizonyult. Az utóteszten azonban eltérően alakult a teljesítmények szórása: a kontrollcsoportban a szórással növekedett, vagyis a tanulók közötti különbségek nagyobbá váltak, míg a kísérleti csoportban ez a növekedés jóval kisebb mértékű volt. Ez arra utalhat, hogy a fejlesztőprogram nemcsak általános fejlődést eredményezett, hanem a gyengébb képességű tanulók fejlődését is hatékonyabban támogatta. Ennek következtében a kísérleti csoport teljesítménye homogénebbé vált, míg a kontrollcsoportban a különbségek inkább felerősödtek.

A közepes képességű tanulók esetében az előteszten szintén nem mutatkozott szignifikáns különbség a kontroll- és a fejlesztőcsoport teljesítménye között ($t = -0,268$, $p = 0,789$), ami jelzi, hogy a két csoport kezdeti szintje azonos volt. Az utóteszt eredményei azonban már szignifikáns különbséget mutattak ($t = -2,73$, $p = 0,006$) a fejlesztőcsoport javára. A fejlesztőcsoportban kismértékű, de egyértelmű pozitív hatás volt kimutatható ($d = 0,33$), míg a kontrollcsoport fejlődése minimálisnak bizonyult ($d = 0,11$). Ezek az eredmények arra utalnak, hogy a program a közepes szókincsszinttel rendelkező tanulók fejlődését is hatékonyan támogatta. A fejlesztőcsoport kisebb szórása azt jelzi, hogy a fejlődés ebben a csoportban egyenletesebben valósult meg.

A magas képességszintű tanulók esetében sem mutattak az előteszt eredményei szignifikáns különbséget a kontroll- és a fejlesztőcsoport között ($t = -0,191$, $p = 0,849$), vagyis a két csoport kiinduló teljesítménye hasonló volt. Az utóteszt eredményei azonban már itt is szignifikáns eltérést jeleztek ($t = -2,114$, $p = 0,036$), ahol a fejlesztőcsoport teljesítménye tovább javult, míg a kontrollcsoportban enyhe visszaesés volt tapasztalható. A fejlesztőcsoportban kisebb, de pozitív hatás jelentkezett ($d = 0,25$), míg a kontrollcsoportban enyhe teljesítménycsökkenés volt megfigyelhető ($d = -0,09$). Ez azt jelzi, hogy a program a magasabb kezdeti szókincsszinttel rendelkező tanulók esetében is kedvező hatást gyakorolt a teljesítményre, jól lehet a fejlődés mértéke mérsékeltebb volt.

Diszkusszió

A kutatás fő célja egy olyan technológia alapú fejlesztőprogram kidolgozása volt, amely hatékonyan alkalmazható a felsős tanulók körében a tantárgyi szókincs mérésére és fejlesztésére osztálytermi környezetben. A program hatékonyságvizsgálatát 123 általános iskola 5025 5–8. évfolyamos tanulója részvételével végeztük, akik különböző képességszintűek voltak. A program hatékonyságának igazolására a következő kutatási kérdéseket tűztük ki: (1) Mennyire hatékonyan fejleszthető az 5–8. évfolyamos tanulók szókincse az online fejlesztőprogrammal? (2) Van-e különbség évfolyamonként a fejlesztés hatékonyságában? (3) Mekkora hatékonysággal alkalmazható a fejlesztőprogram a különböző szókincsszinttel rendelkező tanulók körében?

A program hatékonyságát kontrollcsoportos vizsgálattal mértük. A fejlesztés megkezdése előtt a kísérleti és a kontrollcsoport teljesítménye nem különbözött szignifikánsan, ami biztosította az összehasonlíthatóságot, és a beavatkozás hatásának tisztább értelmezését. Az előteszten mért hasonló átlagok arra utalnak, hogy a tanulók kiinduló szintje kiegyenlített volt, így az utóteszten kimutatott különbségek valóban a program hatásának tulajdoníthatók. Az ismételt mérés kétutas varianciaanalízis eredményei alapján egyértelműen kimutatható volt a fejlesztőhatás, ami általános teljesítményjavulást jelezett a tanulók körében. A fejlesztés és a csoport interakciója szignifikánsnak bizonyult, ami azt jelzi, hogy a fejlesztőprogramban részt vevő tanulók fejlődése eltérő mintázatot követett a kontrollcsoportéhoz képest. A fejlesztőcsoport tagjai nagyobb mértékű javulást mutattak az utóteszten, ami a program pozitív és célzott hatására utal. A hatás nagyságát jellemző részleges eta-négyzet ugyan kismértékű, de statisztikailag megbízható beavatkozási hatást jelez. Ez azt sugallja, hogy a fejlesztés gyakorlati jelentősége mérsékelt, ugyanakkor az eredmény szisztematikus és nem véletlenszerű. A csoportfőhatás nem bizonyult szignifikánsnak, ami azt mutatja, hogy a két csoport összteljesítménye között nem volt lényegi eltérés a teljes mérési időszak során. Ez tovább erősíti azt az értelmezést, hogy a beavatkozás hatása nem a kiinduló különbségekből fakadt, hanem a fejlesztés

során bekövetkezett változások eredménye volt. A szórások elemzése szintén érdekes mintázatot mutatott. A kontrollcsoportban az utóteszten a teljesítmények szórása enyhén növekedett, ami a tanulók közötti különbségek növekedésére utal. Ezzel szemben a fejlesztőcsoportban a variancia stabil maradt, ami a teljesítmények kiegyenlítettebb javulására enged következtetni. Az eredmények megerősítik azt az elméleti megközelítést is, amely szerint a mentális lexikon hálózatos szerveződése lehetővé teszi az új lexikai elemek fokozatos integrálását a meglévő tudásrendszerbe (Aitchison, 2012; Gósy, 2005).

A második kutatási kérdés az évfolyamonkénti különbségek vizsgálatára irányult, vagyis arra, hogy a fejlesztőprogram hatékonysága az egyes évfolyamokon eltérően érvényesült-e. Az ismételt méréses kétutas varianciaanalízis eredményei alapján a fejlesztőhatás szignifikánsnak bizonyult, ami azt mutatja, hogy a tanulók teljesítménye az utóteszten minden évfolyamon javult. A hatásméret közepesnek mutatkozott, ami a beavatkozás általános pozitív hatását támasztja alá. A mérések szerint a fejlesztőprogramban részt vevő tanulók javulása nagyobb mértékű volt, mint a kontrollcsoporté. Ez az eredmény összhangban áll a fejlesztés hatékonyságát alátámasztó korábbi elemzésekkel, és megerősíti, hogy a program képes volt szignifikánsan előmozdítani a tanulók teljesítményét.

Az évfolyam és a fejlesztés interakciója ugyanakkor nem bizonyult szignifikánsnak, ami azt jelenti, hogy ha minden tanulót (fejlesztő- és kontrollcsoport) együtt vizsgálunk, az évfolyamok között nem volt lényegi eltérés a fejlődés mértékében. Más szóval, a javulás általános mintázata hasonló volt az 5–8. évfolyamok között, így az évfolyam önmagában nem befolyásolta a fejlődés ütemét. Ezzel szemben a háromutas interakció (fejlesztés $\times$ csoport $\times$ évfolyam) szignifikáns eredményt mutatott, ami azt jelzi, hogy a fejlesztőprogram hatékonysága nem volt egyforma az évfolyamok között. A fejlesztés hatása az évfolyam függvényében változott: A páronkénti összehasonlítások alapján az 5. évfolyam tanulóinak fejlődése kisebb mértékű volt, mint a 7. és 8. évfolyamé, míg a 6., 7. és 8. évfolyam között nem mutatkozott szignifikáns eltérés. A leíró statisztikák szintén ezt a tendenciát támasztották alá. A fejlesztőcsoportban minden évfolyamon javulás volt tapasztalható, de a növekedés mértéke a magasabb évfolyamokon nagyobb volt. Különösen a 6–8. évfolyamon látható erőteljesebb teljesítménycsökkenés, ami arra utalhat, hogy a program hatékonysága a tanulók életkori és kognitív fejlettségével párhuzamosan nőtt. Az évfolyamok közötti különbségek értelmezésében fontos szempont lehet a tantárgyi nyelvhasználat fokozódó komplexitása. Az évfolyamok előrehaladtával jelentősen nő a tankönyvek lexikai terhelése, valamint a tantárgyi szövegek absztrakciós szintje. Ez összhangban áll Moje (2008), valamint Shanahan és Shanahan (2008) megállapításaival, amelyek szerint a felsőbb évfolyamokon egyre hangsúlyosabbá válik a diszciplináris nyelvhasználat szerepe. A 7–8. évfolyamon tapasztalt kedvezőbb fejlesztőhatás arra utalhat, hogy ezek a tanulók már stabilabb általános akadémiai szókinccsel rendelkeztek, amely megfelelő alapot biztosított a szaktárgyi terminológia hatékonyabb elsajátításához.

A harmadik kutatási kérdés célja az volt, hogy feltárjuk, mennyire hatékonyan alkalmazható a fejlesztőprogram a különböző szókinccszinttel rendelkező tanulók körében, illetve mely képességszint esetében érvényesült legerőteljesebben a program hatása. Ennek érdekében a tanulókat az előteszten nyújtott teljesítményük alapján három csoportba soroltuk: gyengén teljesítők, átlagos szintűek és jól teljesítők. Az eredmények alapján az online tantárgyi szókinccs-fejlesztő program hatékonysága eltérő módon érvényesült a különböző kezdeti képességszinttel rendelkező tanulók körében. Az előtesztetek egyik képességcsoportban sem jeleztek szignifikáns különbséget a fejlesztő- és a kontrollcsoportok között, ami arra utal, hogy a csoportok kiinduló teljesítménye összehasonlítható volt. Ez megerősíti, hogy az utóteszt során kimutatott eltérések nagy valószínűséggel a fejlesztőprogram hatásának tulajdoníthatók, és nem a kezdeti

különbségekből fakadnak. Az alacsony szókincsszintű csoportban a kontroll- és a fejlesztőcsoport teljesítménye között sem az elő-, sem az utóteszten nem mutatkozott szignifikáns különbség. Ez azt jelenti, hogy a program nem váltott ki mérhető többlethatást ebben a csoportban, bár mindkét csoportban általános fejlődés volt megfigyelhető. Az utóteszten tapasztalt szórásnövekedés viszont arra utal, hogy a fejlődés mértéke egyénenként eltérő volt, egyes tanulók jelentősen javultak, mások viszont kevésbé. Ez a heterogenitás különösen a kontrollcsoportban volt erősebb, ami arra enged következtetni, hogy a gyengébb tanulók esetében is volt a programnak felzárkóztató hatása. Az eredmények összhangban állnak Beck et al. (2002) háromszintű szókincsmodelljével, amely szerint a diszciplináris szókincs fejlesztése fokozatos, explicit támogatást igényel.

A közepes képességszintű tanulóknál már egyértelműen kirajzolódott a fejlesztőprogram pozitív hatása. Míg az előteszten a két csoport teljesítménye nem különbözött szignifikánsan, az utóteszten a fejlesztőcsoport szignifikánsan jobb eredményt ért el. Ez a tendencia arra utal, hogy a program különösen ebben a tanulói rétegben volt a leghatékonyabb. A szórásértékek alapján a fejlesztőcsoportban egyenletesebb fejlődés volt megfigyelhető, míg a kontrollcsoport eredményei nagyobb szórást mutattak, ami kevésbé stabil teljesítménynövekedést jelez. A magas képességszintű tanulók esetében a fejlesztőprogram szintén statisztikailag szignifikáns javulást eredményezett. A fejlesztőcsoport teljesítménye tovább emelkedett, míg a kontrollcsoportban enyhe visszaesés volt tapasztalható. Ez arra enged következtetni, hogy a program nemcsak a közepes, hanem a jó képességű tanulók számára is előnyös lehet, mivel számukra is biztosított további fejlődési lehetőséget. A fejlesztőcsoport kisebb szórása pedig arra enged következtetni, hogy a jobb képességű tanulók viszonylag egységesen reagáltak a programra, míg a kontrollcsoportban nagyobb teljesítménybeli eltérések alakultak ki. Ez összhangban áll Nation (2001), Nation, K. (2017) és Schmitt (2014) megállapításaival, amelyek szerint a szókincsfejlesztés hatékonyságának egyik feltétele a már meglévő receptív és produktív lexikai tudás megfelelő szintje. A jobb kezdeti szókincssel rendelkező tanulók könnyebben képesek új fogalmakat integrálni a mentális lexikonba, valamint gyorsabban alakítanak ki szemantikai kapcsolatokat az új és meglévő lexikai elemek között. Hasonló eredményekre utalnak Perfetti (2007) és Perfetti és Stafura (2014) kutatásai is, amelyek szerint az olvasási megértés sikerességét jelentősen meghatározza a lexikai reprezentációk minősége és automatizáltsága.

A kutatás eredményei egyúttal alátámasztják Nagy és Townsend (2012), valamint Snow (2010) megállapításait is, amelyek szerint az általános akadémiai szókincs célzott fejlesztése alapvető feltétele a tantárgyi tudás elsajátításának. A programban alkalmazott fokozatos szintstruktúra – az alapszókincstől az akadémiai, majd diszciplináris szókincsig – lehetővé tette a tanulók számára, hogy fokozatosan kapcsolódjanak be a tudományos nyelvhasználatba. Ez különösen fontos lehet a felső tagozatos oktatásban, ahol a szövegértés már elsősorban a tantárgyi tanulás eszközeként jelenik meg.

Konklúziók

A kutatás eredményei összességében azt igazolták, hogy a kidolgozott online fejlesztőprogram hatékonyan támogatja az általános iskolás tanulók tantárgyi szókincsének és szókincsalapú szövegértésének fejlődését. A kontrollcsoportos kísérleti elrendezés alapján a fejlesztőcsoport tanulói az utóteszten szignifikánsan jobb teljesítményt értek el, mint a kontrollcsoport tagjai, ami megerősíti a program pozitív hatását.

Az eredmények azt mutatták, hogy a fejlesztés minden vizsgált évfolyamon kedvező irányú változást eredményezett, ugyanakkor a program hatékonysága a tanulók kezdeti szókinccsszintjétől függően eltérő mértékben érvényesült. A legerőteljesebb fejlődés a közepes és magasabb teljesítményszintű tanulók körében mutatkozott, ami arra utal, hogy a program eredményes alkalmazásához bizonyos alapvető nyelvi és tanulási készségek megléte szükséges.

Összességében a vizsgálat empirikus bizonyítékot szolgáltat arra, hogy a gyakorisági alapon szervezett, online fejlesztőprogram hatékony eszköze lehet a tantárgyi szókinccs célzott fejlesztésének. Az eredmények egyúttal arra is rámutatnak, hogy a program további differenciálása hozzájárulhat ahhoz, hogy az alacsonyabb szókinccsszinttel rendelkező tanulók is nagyobb mértékben részesüljenek a fejlesztés előnyeiből.

Gyakorlati implikációk

A kutatás eredményei alapján a kidolgozott online tantárgyi szókinccsfejlesztő program hatékony eszköznek tekinthető a felső tagozatos tanulók tantárgyi szókinccsének és szókinccsalapú szövegértésének fejlesztésében. Az eredmények arra utalnak, hogy a program különösen a 6–8. évfolyamon alkalmazható eredményesen, amikor a tanulók metanyelvi tudása, önszabályozó tanulási képessége és tantárgyi nyelvi tapasztalata már megfelelő alapot biztosít a célzott fejlesztéshez. A program hatása legerőteljesebben a közepes és magasabb kezdeti szókinccsszinttel rendelkező tanulók körében mutatkozott meg, ezért ezekben a csoportokban a program önálló vagy részben önálló tanulási formában is hatékonyan alkalmazható.

Az alacsonyabb szókinccsszintű tanulók esetében célszerű a fejlesztőprogram használatát differenciált pedagógusi támogatással kiegészíteni. Számukra ajánlott a rövidebb feladatsorok alkalmazása, a gyakoribb visszacsatolás, a szójelentések explicit magyarázata, valamint a fokozatosan növekvő nehézségi szintek biztosítása. A program jól beilleszthető a tanórai munkába, például bemelegítő feladatként, gyakorló tevékenységként, házi feladatként vagy felzárkóztató és tehetséggondozó foglalkozások részeként. A hatékonyságot tovább növelheti, ha a pedagógus a digitális feladatokat közös megbeszéléssel, kontextusba ágyazott szóhasználatral és tantárgyi szövegfeldolgozással kapcsolja össze.

Összességében az eredmények azt sugallják, hogy a szógyakorisági alapon felépített, adaptív online fejlesztés olyan korszerű pedagógiai eszköz, amely rugalmasan alkalmazható a különböző teljesítményszintű tanulók körében, és hatékonyan támogatja a tantárgyi szókinccs tudatos és rendszeres fejlesztését.

Korlátok és jövőbeli kutatási irányok

Bár a kutatás eredményei megbízható empirikus bizonyítékot nyújtanak a fejlesztőprogram hatékonyságáról, több korlát is azonosítható, amelyek a következtetések értelmezését befolyásolhatják. Először is, a nagy elemszámú minta korlátozta az egyéni fejlődési folyamatok mélyebb vizsgálatát. A jövőbeli kutatásokban kisebb, longitudinális minták alkalmazása segíthetné az egyéni fejlődési pályák pontosabb feltárását. Másodszor, a kutatás nem vizsgálta a tanulók olvasáshoz való attitűdjét és motivációját, pedig ezek a tényezők jelentősen befolyásolhatják a szókinccs- és szövegértés-fejlődést. A jövőben érdemes lenne kvalitatív adatokat is gyűjteni (például interjúkat, önértékelő kérdőíveket), hogy jobban megértsük, miként élik meg a tanulók a fejlesztési folyamatot. Harmadrészt, a szocioökonómiai státusz (SES) nem került bevonásra az elemzésbe, így nem volt mód annak vizsgálatára, hogy a családi háttér milyen mértékben befolyásolja a program hatékonyságát. Ez a tényező a jövőben kulcsfontosságú le-

het a differenciált fejlesztési beavatkozások tervezése szempontjából. További korlátozó tényező a jelentős adatvesztés, amely elsősorban hiányzások és iskolai működési zavarok miatt következett be. A szigorú bevonási kritériumok miatt a minta mintegy 40%-a kiesett, ami torzíthatta az eredményeket és korlátozza az általánosíthatóságot. Végül, mivel a fejlesztőprogram szorosan illeszkedett az iskolai tantervhez, a különböző iskolák és pedagógusok közötti eltérő tanítási módszerek befolyásolhatták az eredményeket. A program hatékonysága így részben a pedagógiai kontextustól is függhetett. A jövőbeli vizsgálatokban érdemes lenne a tanári módszertani és kontextuális változókat is bevonni az elemzésbe.

Összességében a kutatás rámutatott arra, hogy a program hatékonyan támogatja a tanulók szókincsfejlődését és olvasási kompetenciáját, különösen a közepes és magas képességszintű tanulók körében. Ugyanakkor a jövőbeli fejlesztések során érdemes figyelembe venni a tanulói háttértényezőket, az egyéni különbségeket és a tanári gyakorlatokat, hogy a program még inkább adaptálható és eredményes lehessen különböző oktatási kontextusokban.

Köszönetnyilvánítás

A kutatás a Nemzeti Kutatási, Fejlesztési és Innovációs Hivatal OTKA K152413 és a Magyar Tudományos Akadémia KOZOKT2025-4 pályázata keretében valósult meg.

Irodalom

- Aitchison, J. (2012). *Words in the mind. An introduction to the mental lexikon*. Wiley Blackwell.
- Beck, I. L., McKeown, M. G., & Kucan, L. (2002). *Bringing words to life: Robust vocabulary instruction*. Guilford Press Book/Childcraft International.
- Behzad, T., Utkarsh, U., Abir, D., Zarezade, A., Schölkopf, B., & Gomez-Rodriguez, M. (2017). *Optimizing human learning*. arXiv. <https://doi.org/10.48550/arXiv.1712.01856>
- Bosch, L. J. van den, Segers, E., & Verhoeven, L. (2020). First and second language vocabulary affect early second language reading comprehension development. *Journal of Research in Reading*, 43(3), 290–308. <https://doi.org/10.1111/1467-9817.12304>
- Carver, R. P. (1994). Percentage of unknown vocabulary words in text as a function of the relative difficulty of the text: Implications for instruction. *Journal of Literacy Research*, 26, 413–437.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Lawrence Earlbaum Associates.
- Csapó, B., Ainley, J., Bennett, R. E., Latour, T., & Law, N. (2012a). Technological issues for computer-based assessment. In P. Griffin, B. McGaw, & E. Care (Eds.), *Assessment and teaching of 21st century skills* (pp. 143–230). Springer.
- Csapó, B., Lőrincz, A., & Molnár, G. (2012b). Innovative assessment technologies in educational games designed for young students. In D. Ifenthaler, D. Eseryel, & X. Ge (Eds.), *Assessment in game-based learning* (pp. 235–254). Springer. https://doi.org/10.1007/978-1-4614-3546-4_13
- Csapó, B., Molnár, G., & Nagy, J. (2014). Computer-based assessment of school readiness and early reasoning. *Journal of Educational Psychology*, 106(2), 639–650. <https://doi.org/10.1037/a0035756>
- Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in Psychology*, 10(1552). <https://doi.org/10.3389/fpsyg.2019.01522>
- Chen, M. H., Tseng, W. T., & Hsiao, T. Y. (2018). The effectiveness of digital game-based vocabulary learning: A framework-based view of meta-analysis. *British Journal of Educational Technology*, 49(1), 69–77. <https://doi.org/10.1111/bjet.12526>

- Colwell, J., Hutchison, A., & Woodward, L. (2023). Examining a phased planning approach to support elementary preservice teachers in disciplinary literacy-focused instruction. *Literacy Research: Theory, Method, and Practice*, 72(1), 161–178. <https://doi.org/10.1177/23813377231183419>
- Coxhead, A. (2000). A new academic word list. *TESOL Quarterly*, 34, 213–238. <https://doi.org/10.2307/3587951>
- Coxhead, A., & Nation, P. (2001). The specialized vocabulary of English for academic purposes. In J. Flowerdew & M. Peacock (Eds.), *Research perspectives on English for academic purposes* (pp. 252–267). Cambridge University Press.
- Daller, H., Milton, J., & Treffers-Daller, J. (Eds.). (2007). *Modelling and assessing vocabulary knowledge*. Cambridge University Press.
- Ding, J., Zhao, B., Huang, Y., Wang, Y., & Shi, Y. (2023). *GazeReader: Detecting unknown words using webcam for English as a second language (ESL) learners*. arXiv. <https://doi.org/10.48550/arXiv.2303.10443>
- Dujardin, E., Ecalte, J., Auphan, P., Bailloud, N., & Magnan, A. (2023). Vocabulary and reading comprehension: What are the links in 7- to 10 year-old children? *Scandinavian Journal of Psychology*, 64(5), 582–594. <https://doi.org/10.1111/sjop.12912>
- Dunn, L. M., & Dunn, D. M. (2007). Peabody Picture Vocabulary Test, Fourth Edition (PPVT™-4). NCS Pearson, Inc. <https://doi.org/10.1037/t15144-000>
- Gósy, M. (2005). *Pszicholingvisztika*. Osiris Kiadó.
- Hatcher, W. G., & Yu, W. (2018). A survey of deep learning: Platforms, applications and emerging research trends. *IEEE Access*, 6, 24411–24432. <https://doi.org/10.1109/access.2018.2830661>
- Hoffman, J. L., Teale, W. H., & Paciga, K. A. (2014). Assessing vocabulary learning in early childhood. *Journal of Early Childhood Literacy*, 14(4), 459–481. <https://doi.org/10.1177/1468798413501184>
- Hu, M., & Nation, I. S. P. (2000). Unknown vocabulary density and reading comprehension. *Reading in a Foreign Language*, 13(1), 403–430.
- Jarema, G., & Libben, G. (2007). Introduction: Matters of definition and core perspectives. In G. Jarema & G. Libben (Eds.), *The mental lexicon: Core perspectives* (pp. 1–5). https://doi.org/10.1163/9780080548692_002
- Józsa, K., & Steklács, J. (2012). Az olvasás tanításának tartalmi és tantervi szempontjai. In B. Csapó & V. Csépe (Eds.), *Tartalmi keretek az olvasás diagnosztikus értékeléséhez* (pp. 137–188). Nemzeti Tankönyvkiadó Zrt.
- Jurafsky, D., & Martin, J. (2013). *Speech and language processing* (2nd ed.). Pearson International. <https://elibrary.pearson.de/book/99.150005/>
- Kieffer, M. J., & Lesaux, N. K. (2012). Knowledge of words, knowledge about words: Dimensions of vocabulary in first and second language learners in sixth grade. *Reading and Writing*, 25(2), 347–373. <https://doi.org/10.1007/s11145-010-9272-9>
- Kremmel, B., & Schmitt, N. (2016). Interpreting vocabulary test scores: What do various item formats tell us about learners' ability to employ words? *Language Assessment Quarterly*, 13(4), 377–392. <https://doi.org/10.1080/15434303.2016.1237516>
- Lemhöfer, K., & Broersma, M. (2012). Introducing LexTALE: A quick and valid lexical test for advanced learners of English. *Behavior Research Methods*, 44, 325–343. <https://doi.org/10.3758/s13428-011-0146-0>
- Moje, E. B. (2008). Foregrounding the disciplines in secondary literacy teaching and learning: A call for change. *Journal of Adolescent & Adult Literacy*, 52(2), 96–107. <https://doi.org/10.1598/JAAL.52.2.1>
- Magyar, A., Habók, A., & Molnár, Gy. (2023). Számítógépes szókincsmérő teszt kisiskolás diákok részére: alkalmazhatóság, megbízhatóság, működés. *Magyar Pedagógia*, 123(1), 33–47. <https://doi.org/10.14232/mped.2023.1.33>
- Magyar, A., & Molnár, Gy. (2014). A szóolvasási készség adaptív mérését lehetővé tevő online tesztrendszer kidolgozása. *Magyar Pedagógia*, 114(4), 259–279. <https://doi.org/10.17670/MPed.2014.4.259>

- Magyar, A., & Molnár, Gy. (2015). A szóolvasási készség online mérésére kidolgozott adaptív és lineáris tesztrendszer összehasonlító hatékonyságvizsgálata. *Magyar Pedagógia*, 115(4), 403–428. <https://doi.org/10.17670/MPed.2015.4.403>
- Meara, P. (1980). Vocabulary acquisition: A neglected aspect of language learning. *Language Teaching*, 13(3–4), 221–246. <https://doi.org/10.1017/s0261444800008879>
- Molnár, Gy., & Magyar, A. (2018). A szóolvasó készség papíralapú és online értékelő-fejlesztő rendszere. In J. Nagy (Ed.), *Funkcionális analfabetizmus – Megelőző fejlesztési lehetőségek* (pp. 123–140). Mozaik Kiadó.
- Molnár, Gy. (2024). A mesterséges intelligencia hatása a mérés-értékelésre. *Educatio*, 33(1), 55–64. <https://doi.org/10.1556/2063.33.2024.1.6>
- Molnár, Gy., & Csapó, B. (2013). Az eDia online diagnosztikus mérési rendszer. In K. Józsa & J. B. Fejes (Eds.), *PÉK 2013. XI. Pedagógiai Értékelési Konferencia: CEA 2013. 11th Conference on Educational Assessment* (p. 82). Szegedi Tudományegyetem Neveléstudományi Doktori Iskola.
- Molnár, G., & Csapó, B. (2019a). A diagnosztikus mérési rendszer technológiai keretei: Az eDia online platform. *Iskolakultúra*, 29(4–5), 16–32. <https://doi.org/10.14232/iskult.2019.4.5.16>
- Molnár, G., & Csapó, B. (2019b). Making the psychological dimension of learning visible: Using technology-based assessment to monitor students' cognitive development. *Frontiers in Psychology*, 10, 1368. <https://doi.org/10.3389/fpsyg.2019.01368>
- Molnár, G., Pásztor, A., & Csapó, B. (2019). The eLea online training platform. In E. K. Molnár & K. Dancs (Eds.), *XVII. Pedagógiai Értékelési Konferencia - 17th Conference on Educational Assessment*. Program és összefoglalók - Programme and abstracts (p. 64). Szegedi Tudományegyetem Neveléstudományi Doktori Iskola.
- Nagy, J. (2004). A szóolvasó készség fejlődésének kritériumorientált diagnosztikus feltérképezése. *Magyar Pedagógia*, 104(2), 123–142.
- Nagy, W. E., & Townsend, D. (2012). Words as tools: Learning academic vocabulary as language acquisition. *Reading Research Quarterly*, 47(1), 91–108. <https://doi.org/10.1002/RRQ.011>
- Nation, I. S. P. (2001). *Learning vocabulary in another language*. Cambridge University Press. <https://doi.org/10.1017/CBO9781139524759>
- Nation, K. (2017). Nurturing a lexical legacy: Reading experience is critical for the development of word reading skill. *Npj Science of Learning*, 2, 3. <https://doi.org/10.1038/s41539-017-0004-7>
- Oakhill, J., Cain, K., & Elbro, C. (2019). Reading comprehension and reading comprehension difficulties. In D. A. Kilpatrick, R. M. Joshi, & R. K. Wagner (Eds.), *Reading development and difficulties: Bridging the gap between research and practice* (pp. 83–115). Springer International Publishing. https://doi.org/10.1007/978-3-030-26550-2_5
- OECD. (2023). *PISA 2022 results (volume I): The state of learning and equity in education*. OECD Publishing. <https://doi.org/10.1787/53f23881-en>
- Ouellette, G. P. (2006). What's meaning got to do with it: The role of vocabulary in word reading and reading comprehension. *Journal of Educational Psychology*, 98(3), 554–566. <https://doi.org/10.1037/0022-0663.98.3.554>
- Ouellette, G., & Shaw, E. (2014). Oral vocabulary and reading comprehension: An intricate affair. *L'Année Psychologique*, 114(4), 623–645. <https://doi.org/10.4074/S0003503314004023>
- Pásztor, A. (2013). Digitális játékok az oktatásban. *Iskolakultúra*, 23(9), 37–48.
- Pearson, P. D., Hiebert, E. H., & Kamil, M. L. (2007). Vocabulary assessment: What we know and what we need to learn. *Reading Research Quarterly*, 42(2), 282–296. <https://doi.org/10.1598/RRQ.42.2.4>
- Perfetti, C. A. (2007). Reading ability: Lexical quality to comprehension. *Scientific Studies of Reading*, 11(4), 357–383. <https://doi.org/10.1080/10888430701530730>
- Perfetti, C., & Stafura, J. (2014). Word knowledge in a theory of reading comprehension. *Scientific Studies of Reading*, 18(1), 22–37. <https://doi.org/10.1080/10888438.2013.827687>
- Porkoláb, Á., & Fekete, T. (2023). A mesterséges intelligencia alkalmazása a nyelvtanulásban. *Iskolakultúra*, 33(8), 67–80. <https://doi.org/10.14232/iskult.2023.8.67>
- Pusztai, F. & Csábi, Sz. (Eds.). (2003). *Magyar értelmező kéziszótár*. Akadémiai Kiadó.

- Schmitt, N. (2014). Size and depth of vocabulary knowledge: What the research shows. *Language Learning*, 64(4), 913–951. <https://doi.org/10.1111/lang.12077>
- Shanahan, T., & Shanahan, C. (2008). Teaching disciplinary literacy to adolescents: Rethinking content-area literacy. *Harvard Educational Review*, 78(1), 40–59. <https://doi.org/10.17763/haer.78.1.v62444321p602101>
- Snow, C. E. (2010). Academic language and the challenge of reading for learning about science. *Science*, 328(5977), 450–452. <https://doi.org/10.1126/science.1182597>
- Sparapani, N., Connor, C. M., McLean, L., Wood, T., Toste, J., & Day, S. (2018). Direct and reciprocal effects among social skills, vocabulary, and reading comprehension in first grade. *Contemporary Educational Psychology*, 53, 159–167. <https://doi.org/10.1016/j.cedpsych.2018.03.003>
- Szűts, Z., Törteli Telek, M., Móser, A., & Radics, K. (2022). *Az online eszközök által nyitott új utak a nyelvtanulásban: Szakmai, ismeretterjesztő összefoglaló*. <https://onlineplatformok.hu/files/bc221ac1-c747-474a-add0-4c2ff10b03ee.pdf>
- Terbe, E. (Ed.). (2019). *A szövegértés mint tanulási képesség fejlesztése*. Eötvös Loránd Tudományegyetem.
- Van Orman, D. S. J., Ardasheva, Y., Carbonneau, K. J., & Firestone, J. B. (2021). Examining the impacts of extended vocabulary instruction in mixed-English-proficiency science classrooms. *The Journal of Educational Research*, 114(1), 74–88. <https://doi.org/10.1080/00220671.2021.1881754>

ABSTRACT

TECHNOLOGY-BASED VOCABULARY DEVELOPMENT AND READING COMPREHENSION: RESULTS FROM A LARGE-SCALE STUDY

Andrea Magyar, Anita Habók & Gyöngyvér Molnár

Keywords: technology-based vocabulary development, reading comprehension, large-scale study, intervention program, subject-specific vocabulary

In upper primary education, reading and comprehension skills form the foundation for understanding various subjects and for independent learning. Among the many components of reading comprehension, knowledge of subject-specific vocabulary is essential for understanding texts and processing disciplinary materials. The focus of our research was the development of a technology-based intervention program designed to assess and improve students' subject-specific vocabulary across different disciplines. The effectiveness of the program was measured by using a pretest–posttest control group experimental design, involving a total of 1,562 students. Based on the pre-test results, there was no significant difference between the experimental group and the control group. A repeated-measures ANOVA revealed a significant main effect of time and a significant time $\times$ group interaction, indicating that the experimental group showed greater improvement than the control group. Analysis by grade level showed that the developmental effect was significantly lower in Grade 5 than in Grades 7 and 8, while no substantial difference was observed between Grades 6–8. When broken down by ability level, the improvement was strongest among students with medium and high vocabulary proficiency ($p = .006$ and $p = .036$), whereas the progress of lower-performing students did not exceed the level of natural development. The program thus proved to be an effective tool for developing vocabulary and reading performance, particularly among students who already possessed basic language competencies. However, the extent of the effect varied by grade and proficiency level, highlighting the importance of differentiated instruction and adaptive development.

Magyar Pedagógia, 126(2). 99–121. (2026)
<https://doi.org/10.14232/mped.2026.2.99>

Magyar Andrea:  <https://orcid.org/0000-0002-0759-3158>
MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport
H-6722 Szeged, Petőfi Sándor sgt. 32–34.
mandrea@edu.u-szeged.hu

Habók Anita:  <https://orcid.org/0000-0003-0904-8206>
Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola; MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport
H-6722 Szeged, Petőfi Sándor sgt. 32–34.
habok@edpsy.u-szeged.hu

Molnár Gyöngyvér:  <https://orcid.org/0000-0003-4890-6904>
Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola; MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport
H-6722 Szeged, Petőfi Sándor sgt. 32–34.
gymolnar@edpsy.u-szeged.hu



NARRATÍV IRODALMI ÁTTEKINTÉS A PEDAGÓGUSOK MUNKAHELYI STRESSZÉRŐL, MUNKAELÉGEDETTSÉGÉRŐL ÉS PSZICHOSZOMATIKUS STRESSZ-INDIKÁTORAIRÓL

Malhazjan Szergej¹, Pikó Bettina²

¹ Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola

² Szegedi Tudományegyetem, Magatartástudományi Intézet

Bevezetés

Az oktatási intézményekben zajló pedagógiai munka az elmúlt években jelentős átalakuláson ment keresztül, ami a pedagógusok mindennapi munkaterhelésének és pszichés igénybevételének növekedésével járt együtt. Hazai (Koltói & Kiss, 2024; Masa et al., 2022; Szabó & Jagodics, 2019; Szelezsánné Egyed, 2016) és nemzetközi (Arnold & Rahimi, 2025; Nwoko et al., 2023) kutatások rámutatnak, hogy az oktatási környezet változásai, a munkaterhelés növekedése, amelynek részét képezik a fokozódó adminisztratív elvárások, a tanulói sokféleség kezelése, valamint az érzelmi bevonódással járó pedagógiai helyzetek fokozott stresszterhelést jelentenek a pedagógusok számára. Ezek a terhelések hosszabb távon nemcsak a pszichés jóllétet, hanem a munkával való elégedettséget és az egészségi állapotot is kedvezőtlenül befolyásolhatják.

A munkahelyi stressz pedagógiai kontextusban nem csupán egyéni megélésként értelmezhető, hanem olyan komplex jelenségként is, amely szorosan összefügg az intézményi működéssel, a szervezeti kultúrával és a szakmai szerepelvárásokkal (Zhao et al., 2022). Empirikus vizsgálatok szerint a pedagógusok által megélt tartós munkahelyi stressz összefüggésbe hozható a munkával kapcsolatos negatív attitűdök erősödésével, a pályaelhagyási szándék növekedésével, valamint különböző pszichés és testi tünetek megjelenésével (Benevene et al., 2020; Herman et al., 2020). Mindez különösen releváns pedagógiai szempontból, mivel a pedagógusok jólléte közvetetten hatással lehet az oktatási folyamat minőségére és a tanulókkal való kapcsolatokra (Molnár & Fodor, 2024). A munkahelyi stressz és a munkaelégedettség kapcsolatát vizsgáló kutatások egyre inkább hangsúlyozzák, hogy a munkaelégedettség nem kizárólag a stressz következményeként értelmezhető, hanem bizonyos körülmények között védő tényezőként is működhet. Pedagógusok körében végzett vizsgálatok arra utalnak, hogy a pozitív iskolai klíma, a szakmai autonómia és a pedagógiai munka értelmességének megélése mérsékelheti a stressz negatív hatásait, és hozzájárulhat a jóllét fenntartásához (Granziera & Perera, 2019). Ugyanakkor a szakirodalom nem egységes a tekintetben, hogy a munkaelégedettség milyen szerepet tölt be a stressz és az egészségi kimenetek kapcsolatában. Míg egyes megközelítések a stressz következményeként, kimeneti változóként tekintenek rá, más kutatások szerint a munkaelégedettség önmagában is hozzájárulhat a stressz negatív hatásainak mérsékléséhez, és befolyásolhatja az egészségi kimenetek alakulását (Ganster & Rosen, 2013). E kettős értelmezés a pedagóguspályán különösen

releváns, amit a későbbi empirikus vizsgálatok eredményei is alátámasztanak (Jiang & Yuan, 2025; Woods et al., 2023).

A munkahelyi stressz egészségre gyakorolt hatásának vizsgálatában egyre nagyobb figyelmet kapnak a pszichoszomatikus stresszindikátorok, amelyek a tartós pszichés megterhelés testi tünetek formájában megjelenő következményeire utalnak. Pedagógusok és más segítő foglalkozásúak körében végzett kutatások következetesen azonosítják az alvászavarokat, a krónikus fáradtságot, a fejfájást és a mozgásszervi panaszokat, mint a stresszhez kapcsolódó leggyakoribb pszichoszomatikus tüneteket (Llorca-Pellicer et al., 2025; Zhang et al., 2011). Ezek a tünetek gyakran nem köthetők egyértelműen szervi elváltozásokhoz, mégis jelentős mértékben ronthatják a mindennapi működést és a munkaképességet. Ezért tanulmányunkban a pszichoszomatikus stresszindikátorok fogalmi és empirikus megközelítésének kiemelt jelentőséget tulajdonítunk a munkahelyi stressz és munkaelégedettség témakörének áttekintésében.

A jelen narratív áttekintés (narrative review) célja a téma szakirodalmának tematikus szintézise, valamint kritikus, összefoglaló értékelése, amely tematikus vagy kronológiai elrendezésben szintetizálja a meglévő ismereteket. A vonatkozó empirikus kutatások és elméleti tanulmányok szisztematikus értékelése helyett a téma átfogó bemutatására törekszik, és kutatási rések azonosítására fókuszál (Sukhera, 2022). Ennélfogva a források válogatása a szerzők kritikai szemléletét tükrözi, melynek célja az elméleti keretek felvázolásán túl az empirikus kutatási trendek felvázolása, kritikus elemzése és a kutatási hiányosságok azonosítása. A narratív irodalmi áttekintést az indokolja, hogy a pedagógusok és más segítő hivatású csoportok munkahelyi stresszére, munkaelégedettségére és pszichoszomatikus stressz-indikátoraira vonatkozó szakirodalom fogalmilag és módszertanilag heterogén, mivel a vizsgálatok eltérő definíciókat, mérőeszközöket, mintákat és egészségi kimeneteket alkalmaznak. A cél nem egyetlen, szűken körülhatárolt kérdés kimerítő és reprodukálható összegzése vagy hatásbecslése volt, hanem a fő elméleti kapcsolódások, empirikus trendek és kutatási hiányok értelmező szintézise, ezért a narratív áttekintés bizonyult adekvát megközelítésnek. A szakirodalom feltárása több lépésben történt. A nemzetközi publikációk a Scopus, Web of Science, PubMed, Google Scholar, a hazai közlemények pedig a MATARKA, REAL-MTAK, valamint releváns hazai folyóiratkereső felületek felhasználásával kerültek azonosításra. A keresések során angol és magyar nyelvű kulcsszavak kombinációit alkalmaztuk, például: teacher stress, occupational stress, job satisfaction, burnout, psychosomatic symptoms, teachers, helping professions, illetve pedagógus, munkahelyi stressz, munkaelégedettség, kiégés, pszichoszomatikus tünetek, jóllét. A beválasztás során azokat az empirikus tanulmányokat, áttekintéseket és elméleti közleményeket részesítettük előnyben, amelyek közvetlenül kapcsolódtak a pedagógusok vagy más segítő hivatású csoportok munkahelyi stresszéhez, munkaelégedettségéhez és/vagy pszichoszomatikus stressz-indikátoraihoz. Elsősorban frissebb publikációkra támaszkodtunk, ugyanakkor a fogalmi és elméleti keret szempontjából meghatározó korábbi munkákat is bevontuk. Nem kerültek be azok a közlemények, amelyek nem kapcsolódtak közvetlenül a vizsgált konstruktumokhoz, nem pedagógus vagy segítő foglalkozású mintára fókuszáltak, illetve nem nyújtottak érdemi empirikus vagy elméleti hozzájárulást a tanulmány kérdésfeltevéséhez.

Fogalmi és elméleti keret

Munkahelyi stressz a segítő hivatásokban és a pedagóguspályán

Az Egészségügyi Világszervezet (WHO) meghatározása szerint a munkahelyi stressz olyan pszichológiai és fiziológiai válaszként értelmezhető, amely akkor alakul ki, amikor a munkavégzéssel járó felelősség és követelmények tartósan meghaladják az egyén megküzdési képességét és erőforrásait (Maulik, 2017; World Health Organization, 2017). A kortárs megközelítések hangsúlyozzák, hogy a munkahelyi stressz nem kizárólag az objektív munkaterhelés következménye, hanem az egyéni és környezeti tényezők dinamikus kölcsönhatásában jön létre, magában foglalva a munkakörnyezet strukturális jellemzőit, a szervezeti elvárásokat, valamint az egyén szubjektív értelmezéseit és rendelkezésre álló erőforrásait (Ganster & Rosen, 2013; Nixon et al., 2011). A segítő szakmák esetében a munkahelyi stressz sajátos jelleget ölt, amely elsősorban a tartós emberi kapcsolatokról, az érzelmi bevonódásból, valamint a mások szükségleteire irányuló folyamatos figyelemből fakad. A pedagógusok, egészségügyi és szociális területen dolgozó szakemberek munkája gyakran jár együtt érzelmi munkavégzéssel, amely során a szakembereknek szabályozniuk kell saját érzelmi reakcióikat a munkavégzés során elvárt hivatásos szerepnek megfelelően. Az érzelmi bevonódással járó segítő munka hosszabb távon hozzájárulhat a stressz tartóssá válásához, különösen olyan szervezeti környezetekben, ahol az elvárások és a rendelkezésre álló erőforrások között tartós egyensúlytalanság áll fenn (Maslach & Leiter, 2016).

A munkahelyi stressz értelmezésében a segítő hivatások körében széles körben alkalmazott elméleti keretet nyújt a munkakövetelmények és erőforrások modellje a (JD–R-modell; munkahelyi követelmények–erőforrások modell), amely szerint a stressz kialakulása elsősorban a magas munkakövetelmények és az elégtelen erőforrások együttes jelenlétéhez kapcsolódik (Li et al., 2025). A modell korszerű alkalmazásai hangsúlyozzák, hogy a munkakövetelmények nem önmagukban károsak, hanem akkor vezetnek stresszhez és egészségromláshoz, ha azokat nem ellensúlyozzák megfelelő munkahelyi és személyes erőforrások (Bakker & Demerouti, 2017; Lesener et al., 2019). A pedagóguspályán a munkakövetelmények közé sorolhatók többek között a túlzott adminisztratív terhek, az időnyomás, az osztálytermi fegyelmezési nehézségek, a tanulói és szülői elvárások, valamint az oktatáspolitikai változásokból fakadó bizonytalanság. Ezzel szemben az olyan erőforrások, mint a szakmai autonómia, a testületi támogatás, a vezetői visszajelzés vagy a munka értelmességének megélése, védő szerepet tölthetnek be a stressz káros hatásaival szemben (Kovács et al., 2024).

A pedagógusokra vonatkozó gyakorlati kutatások következetesen rámutatnak arra, hogy a munkahelyi stressz egyik legjelentősebb forrása a szerepkonfliktusok és a szereptúlterhelés jelenléte (Mihálka & Pikó, 2018). A pedagógusok gyakran egyszerre tapasztalják meg a tanítási feladatokkal, a nevelési felelősséggel, az adminisztratív elvárásokkal és az intézményi megfelelési kényszerekkel járó terhelést, ami hozzájárulhat a stressz fokozódásához és a pszichés kimerülés kialakulásához (Aloe et al., 2014; Skaalvik & Skaalvik, 2017). A segítő hivatások más területein, mint például az egészségügyben vagy a szociális ellátórendszerben, hasonló stresszorok azonosíthatók, kiegészülve a fokozott időnyomással, az érzelmileg megterhelő klienshelyzetekkel és az erőforráshiánnyal (Ganster & Rosen, 2013; Stosic et al., 2022). A munkahelyi stressz mérésére a segítő hivatások körében elsősorban önbeszámoló kérdőíveket alkalmaznak, amelyek a stressz észlelt mértékét, a munkával kapcsolatos stresszorokat és a stresszhez kapcsolódó pszichés reakciókat vizsgálják (Ganster & Rosen, 2013). A pedagógusokra fókuszáló kutatások egy része általános munkahelyi stresszskálákat alkalmaz, míg más vizsgálatok kifejezetten a pedagóguspálya sajátosságaira kidolgozott

mérőeszközöket használnak. A mérési megközelítések sokfélesége megnehezíti az eredmények közvetlen összehasonlítását, és hangsúlyozza, hogy a munkahelyi stressz fogalma és annak mérési keretei nem egységesek a szakirodalomban (Harvey et al., 2017; Hassard et al., 2018; Kyriacou, 2001).

Pszichoszomatikus stresszindikátorok a segítő hivatásokban

A munkahelyi stressz egészségre gyakorolt hatásának vizsgálatában kiemelt szerepet töltenek be a pszichoszomatikus stresszindikátorok, amelyek a stresszhez kapcsolódó testi tünetek és funkcionális panaszok formájában jelenhetnek meg (Marcatto et al., 2024). A segítő hivatásokban és a pedagógusi munkában a munkahelyi stressz gyakran nem csupán pszichológiai distresszben, hanem jól körülírható pszichoszomatikus indikátorokban is megjelenik. A pedagógusokra szabott mérési keretek közül a stresszforrások és a hozzájuk kapcsolódó tünetek feltárására gyakran alkalmazzák a Teacher Stress Inventory többdimenziós rendszerét is (Jin et al., 2008).

A segítő jellegű egészségügyi minták hasonló együttjárást erősítenek meg, bár eltérő mérési megoldásokkal. Zhang és munkatársai (2011) pekingi népegészségügyi szakembereknél négy pszichoszomatikus panaszterületet azonosítottak (szemdiszkomfort; fejfájás/szédülés; inszomnia/alvászavar; depresszió/érzelmi instabilitás), és regressziós eredményeik szerint a fennálló panaszok nemcsak demográfiai és életmódbeli jellemzőkkel, hanem a munkafeltételekkel és a munkahelyi környezettel is összefüggésben álltak. Különösen a tartósan nagy munkaterhelés és a munka által kiváltott stressz mutatott erős kapcsolatot a pszichoszomatikus panaszokkal, míg a támogató munkahelyi légkör és a munka iránti érdeklődés inkább védő tényezőként jelent meg, ami arra utal, hogy a szervezeti klíma képes lehet csökkenteni a stressz egészségre gyakorolt hatásait (Zhang et al., 2011).

A kínai ápolók körében Gu és munkatársai (2019) a pszichoszomatikus jóllétet standardizált mérőeszkővel külön dimenziókra bontva vizsgálták, így a depressziót, a szorongást, az alvásminőséget és a szomatikus tünetterhelést is elkülönítve mérték. Eredményeik alapján a munkastresszorok különböző típusai nem ugyanúgy kapcsolódtak ezekhez a kimenetekhez, például a munkaterhelés és az időnyomás elsősorban a szorongással járt együtt, míg az interperszonális nehézségek és a menedzsmenthez köthető problémák több területen is megjelentek, köztük a szorongás, a depresszió és a PHQ 15-tel mért, az elmúlt hónapra vonatkozó szomatikus tünetek esetében is (Gu et al., 2019). Ez összességében azt sugallja, hogy a segítő szakmákban a pszichoszomatikus kimenetek értelmezése akkor a leginformatívabb, ha a stresszforrásokat és a panaszokat többdimenziós rendszerben vizsgáljuk, nem pedig egyetlen, globális stresszmutató alapján (Gu et al., 2019).

A rendelkezésre álló kutatások alapján a pszichoszociális munkahelyi terhelések, mint például a fokozott munkahelyi megterhelés és stressz, a mozgásszervi panaszok kockázatával is összefügghetnek; longitudinális vizsgálatok metaanalitikus összegzése szerint munkahelyi feszültség (job strain) esetén emelkedett rizikó figyelhető meg mozgásszervi fájdalmakra (Amiri & Behnezhad, 2020).

A pszichoszomatikus stresszindikátorok mérésére a segítő hivatásokkal foglalkozó vizsgálatok nagy része továbbra is önbeszámoló kérdőíveket használ, amelyek jellemzően a testi tünetek gyakoriságát, intenzitását vagy a tünetekhez kapcsolódó szubjektív megterhelést ragadják meg. Bár egyes kutatások kiegészítik ezt objektívebb indikátorokkal (pl. fiziológiai mutatókkal), a terület jelentős részében a szubjektív egészségmutatók dominálnak, ami egyfelől a mindennapi működés szempontjából releváns, másfelől módszertani heterogenitást és

összehasonlíthatósági korlátokat is eredményezhet (Niedhammer et al., 2021). Eltérő tünetlisták, skálák és időbeli keretek használata megnehezíti az eredmények közvetlen összehasonlíthatóságát és az általános következtetések levonását. A pszichoszociális munkatényezők és az egészségi kimenetek vizsgálatában alkalmazott eltérő mérési megközelítések szintén megnehezítik az eredmények összehasonlítását, ami indokolja az egységesebb fogalmi keretek és mérési eljárások kialakításának igényét (Niedhammer et al., 2021). A következőkben röviden összefoglaljuk a tanulmányban alkalmazott kulcsfogalmakat és azok értelmezési kereteit, mielőtt rátérnénk az empirikus kutatások főbb eredményeinek áttekintésére (ld. 1. táblázat).

1. táblázat

Fogalmi keretek összefoglalása

Fogalom	Meghatározás és jellemzők	Pedagógusok és segítő szakmák sajátosságai
Munkahelyi stressz	Tartós egyensúlytalanság a munkakövetelmények és az egyén rendelkezésre álló erőforrások között, amely pszichés és testi terheléssel jár (Ganster & Rosen, 2013).	A segítő szakmákban fokozott érzelmi munka, magas felelősség és tartós interperszonális bevonódás jellemzi; pedagógusoknál gyakori a szerepkonfliktus és az adminisztratív túlterhelés (Maslach & Leiter, 2016).
Munkaelégedettség	A munkával kapcsolatos általános pozitív értékelés, amely magában foglalja az autonómia, a kompetencia és a munkahelyi kapcsolódás megélését, ezáltal az értelmes munka megtapasztalását (Martela & Riekk, 2018).	Pedagógusoknál kiemelt szerepet kap a tanulók fejlődésének megélése és a szakmai autonómia; segítő szakmákban a másokon segítségértéke fontos védő tényező (Klassen & Chiu, 2010).
Pszichoszomatikus stresszindikátorok	A tartós pszichés megterheléshez kapcsolódó testi tünetek, amelyek nem mindig köthetők egyértelmű szervi elváltozásokhoz (Marcatto et al., 2024).	Gyakori az alvászavar, krónikus fáradtság, fejfájás és mozgásszervi fájdalom; segítő szakmákban ezek magas prevalenciával fordulnak elő (Amiri et al., 2024; Llorca-Pellicer et al., 2025; Zhang et al., 2011).

Forrás: saját szerkesztés.

Az 1. táblázat összefoglalja a tanulmányban használt kulcsfogalmak fő értelmezési kereteit, és jól láthatóvá teszi, hogy a stressz, a pszichoszomatikus indikátorok és a munkaelégedettség egymással szorosan összefüggő, de eltérő elemzési szinteken megragadható konstruktumok. A továbbiakban a munkaelégedettség jelentőségét tekintjük át, külön figyelmet fordítva arra, hogy ez a tényező milyen szerepet tölthet be a pedagógusok és más segítő hivatású csoportok munkahelyi stresszének és jóllétének alakulásában.

A munkaelégedettség jelentősége a segítő hivatásokban

A munkaelégedettség a munkavégzéssel összefüggő szubjektív értékelések összességeként értelmezhető, amely kiterjed az egyén munkakörülményeire, feladataira és szakmai kapcsolataira (Brown & Lent, 2008). A szakirodalom a munkaelégedettséget a munkahelyi jóllét egyik meghatározó tényezőjeként jeleníti meg, amely kapcsolatban áll a munkateljesítménnyel, a szervezeti elköteleződéssel, valamint a pszichés és testi egészséggel is (Faragher et al., 2005; Ray, 2022). A segítő hivatások esetében a munkaelégedettség különösen összetett konstrukciónak tekinthető, mivel ezekben a szakmákban a munkavégzés értelme, társadalmi hasznossága és a másokon való segítség lehetősége kiemelt szerepet játszik az elégedettség megélésében. Pedagógusok körében végzett kutatások következetesen rámutatnak arra, hogy a munkaelégedettséget nem kizárólag a munkafeltételek objektív jellemzői határozzák meg, hanem olyan pszichoszociális tényezők is, mint a szakmai autonómia, a kollegiális támogatás, a vezetői visszajelzések minősége és a pedagógiai munka értelmességének szubjektív megélése (Klassen & Chiu, 2010; Skaalvik & Skaalvik, 2017). Ezek a tényezők a pedagógusok jóllétének fenntartásában és a munkával kapcsolatos pozitív attitűdök kialakulásában is szerepet játszanak. A munkaelégedettség és a munkahelyi stressz kapcsolatát vizsgáló empirikus kutatások többsége arra utal, hogy a magasabb szintű munkahelyi stressz alacsonyabb munkaelégedettséggel jár együtt, különösen olyan munkakörnyezetekben, ahol a munkakövetelmények tartósan meghaladják az egyén rendelkezésére álló erőforrásokat (Faragher et al., 2005; Ganster & Rosen, 2013). Ugyanakkor a szakirodalomban nem egységes az oksági irány megítélése a munkahelyi stressz és a munkaelégedettség kapcsolatában. Míg egyes megközelítések a stresszt tekintik a munkaelégedettség csökkenésének elsődleges forrásaként, más modellek szerint a munkaelégedettség szintje befolyásolhatja azt, hogy az egyén hogyan értelmezi és éli meg a munkahelyi követelményeket (Jiang & Yuan, 2025; Woods et al., 2023). A munkakövetelmények és erőforrások modellje alapján a munkaelégedettség olyan pozitív, munkával kapcsolatos állapotként értelmezhető, amely összefüggésben áll a rendelkezésre álló erőforrások szintjével, és közvetett módon befolyásolhatja a munkahelyi stresszhez kapcsolódó pszichés és testi kimeneteket. A modell értelmezése szerint a magasabb szintű munkaelégedettség nem szünteti meg a munkakövetelményeket, ugyanakkor hozzájárulhat ahhoz, hogy az egyén a munkahelyi kihívásokat kevésbé élje meg fenyegetőként, és hatékonyabban alkalmazza megküzdési erőforrásait (Bakker & Demerouti, 2017). Ennek megfelelően a munkaelégedettség a szakirodalomban nem kizárólag kimeneti változóként, hanem bizonyos összefüggésekben közvetítő vagy moderáló tényezőként is szerepet kap a munkahelyi stressz és a jóllétet érintő kimenetek kapcsolatában. A pedagóguspályán a munkaelégedettség alakulása szoros összefüggést mutat a kiégéssel és a pályaelhagyási szándékkal. Empirikus eredmények szerint az alacsonyabb munkaelégedettség együtt járhat az érzelmi kimerültséggel, a deperszonalizációval és a szakmai hatékonyság csökkenésének gyakoribb megélésével, míg a magasabb szintű elégedettség védő tényezőként járulhat hozzá a pedagógusok jóllétének és pályán maradásának fenntartásához (Maslach & Leiter, 2016; Skaalvik & Skaalvik, 2017). Hasonló összefüggések más segítő hivatások esetében is megfigyelhetők, ami arra utal, hogy a munkaelégedettség a segítő szakmákban általános pszichológiai erőforrásként értelmezhető (Hricova et al., 2020).

Empirikus kutatások áttekintése

A munkahelyi stressz és pszichoszomatikus tünetek a pedagógusok körében

A pedagóguspályán megjelenő munkahelyi stressz és az ahhoz kapcsolódó pszichoszomatikus tünetek vizsgálata a 2000-es évektől kezdve kiemelt figyelmet kapott a nemzetközi szakirodalomban, amelynek során a *coping*, a társas támogatás, a reziliencia is fontos szempontot képviseltek a védelemben és a megelőzésben. Egy, a hongkongi tanárok körében végzett vizsgálatban Jin és munkatársai (2008) tíz jellegzetes panaszt vettek számba, többek között az ingerlékenységet, a szorongást, a fejfájást, az alvászavart, a szívdobogásérzést, a koncentrációs nehézségeket és a feledékenységet, és azt kérdezték a résztvevőktől, hogy az elmúlt három hónapban milyen gyakran tapasztalták ezeket. Eredményeik szerint valamennyi azonosított stresszforrás pozitív kapcsolatban állt a pszichoszomatikus tünetekkel, miközben a tanítással összefüggő munkaterhelés bizonyult a legerősebb összefüggésű prediktornak, ami arra utal, hogy a napi oktatási rutinból fakadó intenzív követelmények közvetlenül hozzájárulhatnak a tanárok testi-lelki panaszterhéhez. Egy német középiskolai tanárok körében végzett kutatásban szignifikáns kapcsolatot találtak a stresszel való megbirkózás, azaz a *coping* készség, valamint a pszichoszomatikus tünetek között; különösen a társas megküzdési formák és a társas támogatás jelentettek fontos védőfaktort a kiegész és a pszichoszomatikus tünetek ellen (Bauer et al., 2006). Finnországban egy korai kutatásban pszichofiziológiai vizsgálatokkal is igazolták a munkahelyi stressz és a pszichoszomatikus egészség kapcsolatát, így pl. a vérnyomás, az izomtónus vagy endokrin paraméterek (stresszhormonok, adrenalin, noradrenalin) segítségével vetették össze a tanév és a nyári szünet értékeit (Ritvanen et al., 2003).

A pedagógusokra fókuszáló kutatások egy része arra is felhívja a figyelmet, hogy bizonyos munkakörnyezethez kapcsolódó stresszorok különösen erősen kapcsolódhatnak a pszichoszomatikus tünetek kialakulásához. Ilyen tényezők lehetnek például az osztálytermi fegyelmezési nehézségek, a tanulói magatartási problémák, a túlzott adminisztratív terhek, valamint az intézményi támogatás hiánya, amelyek gyakran kumulatív módon vannak jelen a pedagógusok mindennapi munkájában (Bauer et al., 2006; Klassen & Chiu, 2010; Kyriacou, 2001). E stresszorok együttes fennállása hozzájárulhat a tartós stressz kialakulásához, és növelheti a pszichoszomatikus tünetek megjelenésének kockázatát.

Az utóbbi években megjelent friss, pedagógusokra fókuszáló szintetizáló kutatások tovább erősítik a stresszhez és kiegészhez kapcsolódó egészségi kimenetek jelentőségét. Egy átfogó szisztematikus áttekintés például arra jutott, hogy a tanári kiegész konzisztensen együtt jár különböző fizikai egészségi panaszokkal, többek között fejfájással, emésztőrendszeri tünetekkel, alvásproblémákkal és általános kimerültséggel, ami arra utal, hogy a tartós munkahelyi stressz testi szinten is mérhető következményekkel járhat a pedagógusok körében (Madigan et al., 2023). Ez az eredmény megerősíti a szakirodalomban régóta megjelenő azon feltételezést, hogy a pedagóguspálya sajátos terhelései összefüggésbe hozhatók az egészségi állapotbeli tünetek gyakoribb megjelenésével (Kyriacou, 2001).

Perui általános és középiskolában dolgozó pedagógusok 20,9%-ában súlyos, 45,2%-ában közepes, és 33,9%-ukban alacsony szintű pszichoszomatikus tüneti gyakoriságot találtak, és leggyakoribb tünetként a fáradtság, a fejfájás és az alvászavarok szerepeltek. A nem és a foglalkozási státusz szerint azonban jelentős eltérések voltak a tüneti gyakoriságokban (Estrada-Araoz et al., 2023). Egy több, mint 14 ezer kínai pedagógust vizsgáló kutatásban a kiegészét szoros kapcsolatba hozták az alvászavarokkal, azonban a reziliencia moderáló hatást fejtett ki az összefüggésre, azaz védőfaktorként funkcionált (Yang et al., 2023). Egy arab és izraeli pedagógusok körében végzett vizsgálat szintén igazolta, hogy a munkahelyi stressz

hatásai nem kizárólag pszichés szinten jelentkeznek, hanem testi tünetek formájában is megjelenhetnek, különösen akkor, ha a stressz tartósan fennáll (Ghadban et al., 2023). A tüneti gyakoriság szoros összefüggést mutatott a tartós stressz okaiként szereplő munkahelyi erőszakkal, valamint a kiégés mértékével.

Az empirikus pedagógusvizsgálatok arra is rámutatnak, hogy a munkahelyi stressz és az egészségi állapot kapcsolata nem egységes, és jelentős egyéni különbségeket mutat. Egy friss vizsgálat szerint a pedagógusok stresszszintje és érzelmi egészsége szoros összefüggést mutat a megküzdési stratégiák jellegével, ami arra utal, hogy az egyéni erőforrások és stresszkezelési módok módosíthatják a stresszhez kapcsolódó pszichés és testi kimenetek megjelenését (Emeljanovas et al., 2023). Ezek az eredmények összhangban állnak azokkal az elméleti megközelítésekkel, amelyek a munkahelyi stressz egészséghatásait komplex, többtényezős folyamatként értelmezik (Schonfeld & Chang, 2017).

A hazai szakirodalom is egyértelműen jelzi, hogy a pedagóguspálya mentális megterhelése nem pusztán nemzetközi tendencia. Berecz és Mandel (2022) a bölcsődei és óvodai pedagógusok mentális állapotát áttekintve a stressz, a kiégés és a munkahelyi pszichoterror összekapcsolódó jelenlétét emelik ki. Gyarmati és munkatársai (2020) tanulmánya szerint a pedagóguspálya komoly kiégési faktorokkal terhelt, ezért a tünetek korai felismerése és a prevenció különösen fontos. Tóthpál-Halasi és Pikó (2024) eredményei alapján az intrinzik és identifikált motiváció védő jellegű tényezőként értelmezhető, mivel alacsonyabb kiábrándultsággal és depresszióval, valamint magasabb jólléttel járt együtt. Ezzel szemben az amotiváció kedvezőtlenebb mintázatot mutatott, hiszen magasabb kiábrándultsággal, kimerültséggel és depresszióval, illetve alacsonyabb jólléttel kapcsolódott össze. Patai és Németh (2026) vizsgálata szerint a pedagógusok rendszeresen szembesülnek a stressz és a kiégés jeleivel, amelyek megelőzésében a vezetői és szervezeti támogatás, az együttműködés, az önreflexió és a pihenés is meghatározó szerepet játszik. A pályán maradás motivációs oldaláról Koltói és Kiss (2024) rámutatnak, hogy az elhivatottság, a társas beágyazottság, valamint a szakmai és szervezeti elkötelezettség csökkentheti a kilépési szándékot. Szabó és Fehérvári (2023) eredményei szerint a pályán maradást elsősorban a belső motivációs tényezők és a pálya iránti elköteleződés támogatják, miközben a magasabb kiégést mutató pedagógusoknál a kényszer szerepe erősebbé válik. Emellett Szabó és munkatársai (2020) négy nehézségkomponenst azonosítottak, a diákok viselkedését, a vezető hozzáállását, a túlterheltséget és a személyes kimerülést, amelyek közül különösen a túlterheltség mutatott szoros kapcsolatot a kiégéssel. Elekes és munkatársai (2025) 555 fős hazai pedagógusmintán azt találták, hogy a testi tünetek megjelenése szorosan összefügg a szubjektív egészségérzettel és a kiégés mértékével; a kimerülés és a kiábrándultság közepesen erős együttjárást mutatott a testi tünetekkel. Ugyanakkor fontos módszertani korlátot jelent, hogy e vizsgálatok többsége önbeszámolós mérőeszközökre épül, ami a szubjektív torzítás lehetőségét hordozza magában, és nem teszi lehetővé az oksági összefüggések egyértelmű feltárását.

Munkahelyi stressz, munkaelégedettség és az egészségi kimenetek kapcsolata pedagógusok körében

A pedagógusok körében végzett empirikus kutatások jelentős része a munkahelyi stressz és a munkaelégedettség kapcsolatát vizsgálja, és több tanulmány arra utal, hogy a munkaelégedettség a stresszhez kapcsolódó kedvezőtlen kimenetekkel szemben erőforrás-jellegű, illetve bizonyos körülmények között feltételhez kötött védő szerepet is betölthet (Klassen & Chiu, 2010; Skaalvik & Skaalvik, 2017; Woods et al., 2023). Pedagógusmintákon végzett vizsgálá-

tok következetesen azt mutatják, hogy a magasabb munkahelyi stressz alacsonyabb munkaelégedettséggel jár együtt, különösen olyan munkakörnyezetekben, ahol a pedagógusok tartós túlterheltséget, időnyomást és alacsony kontrollt élnek meg (Klassen & Chiu, 2010; Skaalvik & Skaalvik, 2017). A munkaelégedettség és az általános jólléti, illetve egészségi mutatók kapcsolatát a tágabb munkapszichológiai szakirodalom metaanalitikus eredményei is alátámasztják, jelezve, hogy a munkaelégedettség számos pszichés, valamint önértékelt testi egészségi kimenettel együtt járhat (Faragher et al., 2005). A pedagógusok körében készült empirikus vizsgálatok arra is rámutatnak, hogy az alacsonyabb munkaelégedettség és a munkával kapcsolatos negatív pszichés állapotok gyakran társulnak a pszichés kimerültség és a stresszhez kapcsolódó egészségi panaszok gyakoribb megjelenésével. Kutatások szerint a pedagógusok körében gyakran együtt fordulnak elő a stresszhez és kiégéshez társuló testi tünetek, így az alvásproblémák, a krónikus kimerültség és a fájdalomtünetek, valamint a munkával kapcsolatos negatív pszichés állapotok (Madigan et al., 2023). Ezek az eredmények összhangban állnak azzal az értelmezéssel, hogy a munkaelégedettség csökkenése és a stressz testi szinten történő megjelenése gyakran együtt fordul elő, ugyanakkor a keresztmetszeti vizsgálati elrendezések korlátozzák az oksági irányok egyértelmű megállapítását.

A munkaelégedettség lehetséges védő szerepének empirikus vizsgálata pedagógusmintákon egyre gyakrabban jelenik meg moderációs és mediációs modellek formájában. Egy nagymintás vizsgálat például kimutatta, hogy a megküzdési stratégiák moderálják a munkahelyi stressz és a munkaelégedettség kapcsolatát: hatékonyabb megküzdés esetén a stressz negatív hatása az elégedettségre mérsékeltebb mértékben érvényesül (Woods et al., 2023). A vizsgálati eredmények arra utalnak, hogy az egyéni pszichológiai erőforrások fontos szerepet játszanak abban, miként kapcsolódik össze a stresszélmény és a munkával való elégedettség pedagógusok körében. Más empirikus kutatások mediációs mechanizmusokra is rávilágítottak. Egy friss tanulmány szerint a munkaelégedettség és a kiégés közötti kapcsolatot a pedagógusok pszichológiai tőkéje közvetítheti, miközben a fizikai aktivitás moderáló tényezőként befolyásolja a közvetett hatás erősségét (Jiang & Yuan, 2025). Hasonlóan, pedagógusmintán végzett kutatások kimutatták, hogy a munkahelyi stressz és a kiégés kapcsolatában olyan köztes változók, mint a munka–magánélet konfliktus vagy az énhatékonyság, szintén meghatározó szerepet játszhatnak (Zhao et al., 2022). Mindez arra utal, hogy a munkaelégedettség és a kapcsolódó pszichológiai erőforrások nem pusztán kimeneti változóként, hanem a stressz és jóllét közötti összefüggések aktív alakítóiként is értelmezhetők pedagógiai kontextusban.

Pedagógusok és más segítő hivatások összehasonlítása a munkahelyi stressz, a munkaelégedettség és a pszichoszomatikus tünetek szempontjából

A pedagógusok munkahelyi stresszének és egészségi kimeneteinek értelmezése egyre gyakrabban jelenik meg tágabb, segítő szakmákra kiterjedő kontextusban, más segítő hivatások tapasztalataival összefüggésben. Empirikus és szisztematikus vizsgálatok rámutatnak, hogy a pedagógusok által megélt munkahelyi stressz és az ahhoz kapcsolódó pszichés és testi tünetek (pl. fejfájás, immunválasz jelzők) hasonló mintázatokat mutatnak azokkal a tünetekkel, amelyeket a segítő és egészségügyi területen dolgozó szakembereknél írtak le (Al Shuaibi, 2025; Chunga Diaz et al., 2024; Madigan et al., 2023). Ezek az eredmények arra engednek következtetni, hogy a segítő hivatások közös pszichoszociális jellemzői, mint például a tartós interperszonális bevonódás és a magas felelősség, összefüggésbe hozhatók a munkahelyi stressz fokozott kockázatával.

Bár a stresszorok jellege eltérhet, a stressz megélésének intenzitása és tartóssága mindkét csoportban hasonló egészségi kockázatokkal járhat. A munkaelégedettség szerepét vizsgáló

kutatások rámutatnak arra, hogy a tanárok munkahelyi elégedettsége szoros összefüggést mutat a mentális egészséggel és a jólléttel (Tria, 2023), miközben más segítő szakmákban, például az egészségügyi dolgozók körében végzett empirikus vizsgálatok szerint a munkahelyi stressz és a munkaelégedettség kapcsolata szoros összefüggést mutat a kiégés és jóllét kimeneteivel, hangsúlyozva a munka környezeti és szervezeti tényezőinek szerepét (Aljawarneh et al., 2025).

Pedagógusok esetében a kutatások gyakran hangsúlyozzák a szakmai autonómiát, a munka értelmességének megélését és az iskolai környezet támogató jellegét, mint a jóllét és a munkaelégedettség szempontjából fontos erőforrásokat (Admiraal & Kittelsen Røberg, K.-I., 2023; Skaalvik & Skaalvik, 2014). Más segítő hivatásokban, különösen az egészségügyi és gondozási területeken – ezzel szemben a munkaszervezés egyes jellemzői, például a beosztások strukturáltsága és előre jelezhetősége, valamint a vezetői és intézményi támogatás szerepe az érzelmileg megterhelő helyzetek kezelésében jelenik meg hangsúlyosan (Andersen et al., 2025; Jung & Kim, 2025).

Empirikus vizsgálatok szerint a segítő szakmákban jelentkező tartós munkahelyi stressz fiziológiai és pszichoszomatikus tünetekkel is jár. Egészségügyi dolgozók körében nagy arányban számoltak be krónikus fáradtságról és kiégésről (Wang et al., 2025), továbbá fejfájás, kimerültség és ingerlékenység is megjelent nővéreknél a stressz következményeként (Gmayi-naam et al., 2024). Más vizsgálatok pedig alvászavarokat, például insomniát és cirkadián-ritmus zavarokat is azonosítottak a munkahelyi stresszhez kapcsolódóan (Pachi et al., 2025). Bár egyes tanulmányok eltéréseket jeleznek a tünetek erősségében az egyes szakmacsoportok között, a pszichoszomatikus panaszok alapvető jellege több segítő hivatás esetében hasonló, ami a közös egészségi kockázatok jelenlétére hívja fel a figyelmet.

A pedagógusokat és más segítő hivatásúakat érintő kutatások közötti főbb hasonlóságokat és különbségeket a munkahelyi stressz, a munkaelégedettség és a pszichoszomatikus tünetek szempontjából a 2. táblázat foglalja össze.

2. táblázat

A pedagógusok és más segítő hivatások összehasonlítása kutatási szempontok alapján

Szempont	Pedagógusok	Más segítő szakmák (pl. egészségügy, szociális munka)
Tipikus stresszorok	Szerepkonfliktus, adminisztratív terhek, időnyomás, fegyelmezési nehézségek, szülői elvárások.	Időnyomás, műszakos munkarend, érzelmileg megterhelő esetek, krónikus erőforráshiány.
Védő tényezők / erőforrások	Szakmai autonómia, támogató iskolai klíma, kollegiális együttműködés, a munka értelmességének megélése.	Kiszámítható beosztás, vezetői támogatás, pihenőidő biztosítása, szervezeti támogatás.
Gyakori pszichés és testi tünetek	Alvászavar, kimerültség, fejfájás, mozgásszervi fájdalmak, kiégés tünetei.	Hasonló tünetek, kiegészülve fokozott fizikai kifáradással és egészségügyi kockázatokkal.

Szempont	Pedagógusok	Más segítő szakmák (pl. egészségügy, szociális munka)
Kiemelt sajátosságok	Erős érzelmi bevonódás a tanulókkal, nevelési felelősség, magas hivatástudat.	Fizikai terhelés, élet- és egészségvédelmi felelősség, egészségügyi kockázatok, intenzív munkaszervezési kihívások.

Forrás: saját szerkesztés.

A 2. táblázat összefoglalóan bemutatja a pedagógusok és más segítő hivatásúak körében leggyakoribb stresszorokat, védő tényezőket, valamint néhány pszichoszomatikus és testi tünetet, amelyek a körükben leggyakrabban előfordulnak. Jellemző például a fokozott fizikai terhelés vagy az egészségügyi kockázatok megjelenése az egészségügyben dolgozók esetében, míg a nevelési felelősség és az erős érzelmi bevonódás a pedagógusok számára jelent kihívást.

Kutatási ellentmondások és módszertani hiányok

A pedagógusok és más segítő hivatású csoportok munkahelyi stresszével, munkaelégedettségével és pszichoszomatikus tüneteivel foglalkozó szakirodalom jelentős mennyiségű empirikus eredményt kínál, ugyanakkor számos ellentmondást és módszertani korlátot is magában hordoz. Ezek a hiányosságok nemcsak az eredmények értelmezését nehezítik meg, hanem korlátozzák a pedagógiai gyakorlat számára releváns, általánosítható következtetések levonását is.

Az egyik leggyakrabban felmerülő módszertani kérdés az önkéntes mintavétel alkalmazása és a reprezentativitás hiánya, amely különösen a pedagógusok körében végzett kérdőíves vizsgálatok esetében jelent kihívást (Herman et al., 2020). További visszatérő nehézséget jelent a munkahelyi stresszhez és a tanári jólléthez kapcsolódó konstruktumok fogalmi és mérési heterogenitása. A különböző vizsgálatok eltérő definíciókat és mérőeszközöket alkalmaznak, ami megnehezíti az eredmények összehasonlítását és az általánosítható következtetések levonását. A friss szisztematikus áttekintések is rámutatnak a tanári jóllét fogalmának konceptuális sokféleségére és az operacionalizációk heterogenitására, amely akadályozhatja az empirikus eredmények összevetését (Gámez-Genovart et al., 2025; Kurrle & Warwas, 2025). Emellett a pedagógusokra fókuszáló kutatásokban párhuzamosan jelennek meg az általános stressz- és kiégést mérő eszközök, valamint kifejezetten a tanári pálya sajátosságaira kidolgozott skálák. E mérőeszközök pszichometriai validálása, valamint a különböző skálák közötti empirikus összehasonlíthatóság vizsgálata a szakirodalomban visszatérő módszertani kihívásként jelenik meg (Chen et al., 2022; Martínez-Líbano, 2025; Wang et al., 2024).

További jelentős korlátot jelent a kutatások túlnyomóan keresztmetszeti jellege. Bár ezek a vizsgálatok értékes információkat nyújtanak az együttjárások feltárásában, nem teszik lehetővé az oksági irányok egyértelmű meghatározását. Így továbbra sem tisztázott kellő mértékben, hogy a munkahelyi stressz milyen mechanizmusokon keresztül vezet pszichoszomatikus tünetek kialakulásához, illetve, hogy a munkaelégedettség milyen mértékben tekinthető következménynek, közvetítő tényezőnek vagy védőfaktornak ezekben a folyamatokban. A longitudinális vizsgálatok hiánya különösen problematikus a pedagóguspálya esetében, ahol a stressz és a terhelés gyakran hosszú időn keresztül halmozódik.

A mintavételi sajátosságok szintén korlátozzák az eredmények általánosíthatóságát. Számos kutatás önkéntes részvételen alapul, ami torzíthatja a mintát, mivel a fokozottan érintett vagy éppen kevésbé terhelt pedagógusok nagyobb valószínűséggel vesznek részt a vizsgálatokban. Emellett a pedagógusminták gyakran nem reprezentálják az oktatási rendszer teljes spektrumát, például az eltérő iskolatípusokat, életkori csoportokat vagy regionális sajátosságokat, ami tovább nehezíti az eredmények pedagógiai szintű értelmezését.

A pszichoszomatikus stresszindikátorok vizsgálatával kapcsolatban további módszertani kihívást jelent, hogy a kutatások többsége önbeszámoló eszközökre támaszkodik. Bár ezek a módszerek fontos információkat nyújtanak az egyén szubjektív megéléséről, korlátozottan alkalmasak a testi tünetek objektív értékelésére. A friss, pedagógusokra fókuszáló szakirodalom hangsúlyozza, hogy a többforrású vagy objektív mérési megközelítések alkalmazása az oktatáskutatásban továbbra is viszonylag ritka (Kurrle & Warwas, 2025; Madigan et al., 2023).

Az intervenciók kutatások területén is jelentős hiányosságok azonosíthatók. Bár az elmúlt években megjelentek olyan vizsgálatok, amelyek stresszkezelő vagy jóllétfókuszú programok hatását elemzik pedagógusok körében, ezek jelentős része kis mintán, rövid utánkövetéssel és heterogén módszertani megközelítésekkel dolgozik. Emiatt a beavatkozások hosszú távú hatékonyságáról és pedagógiai beágyazhatóságáról továbbra is korlátozott ismeretek állnak rendelkezésre (Avola et al., 2025; von der Embse et al., 2019).

Végül meg kell említeni, hogy a jövőbeli elemzési irányokat tekintve a téma további feldolgozása során indokolt lehet egy szisztematikus irodalomelemzés elkészítése is. A jelen kéziratban tudatosan megtartottuk a narratív áttekintés műfaját, mivel célunk a heterogén szakirodalom értelmező szintézise, a fogalmi kapcsolódások tisztázása, valamint a kutatási hiányok azonosítása volt. A szisztematikus elemzéshez a fogalmi és módszertani heterogenitás miatt azonban szükséges egy szűkebb fókusz és koncentráltabb szelekciós módszer alkalmazása. Úgy gondoljuk azonban, hogy jelen tanulmány a fogalmi és módszertani összegzéssel és a kritikus értékeléssel hozzájárulhat a további elemzések sikeréhez.

Következtetések és pedagógiai implikációk

A pedagógusok munkahelyi stresszével, munkaelégedettségével és pszichoszomatikus stresszindikátoraival foglalkozó szakirodalom narratív áttekintése arra hívja fel a figyelmet, hogy a pedagóguspálya egészségvédelme nem kizárólag egyéni megküzdési kérdésként értelmezhető, hanem szoros összefüggésben áll az oktatási intézmények szervezeti működésével és a pedagógiai munka strukturális feltételeivel.

A bemutatott empirikus eredmények alapján a munkahelyi stressz, a munkaelégedettség és a testi-lelki jóllét közötti kapcsolatok komplex rendszert alkotnak, amelyben egyéni, szervezeti és rendszerszintű tényezők egyaránt szerepet játszanak. Pedagógiai szempontból különösen fontos felismerés, hogy a munkahelyi stressz tartós fennállása nemcsak a pedagógusok egészségi állapotát érinti, hanem közvetett módon hatással lehet a pedagógiai munka minőségére, a tanulókkal való kapcsolatokra és az oktatási folyamat eredményességére is (Collie et al., 2012; Madigan et al., 2023). A pszichoszomatikus tünetek megjelenése, mint például az alvászavarok, a krónikus fáradtság vagy a fájdalomtünetek, hosszabb távon csökkentheti a pedagógusok munkaképességét és pályán maradási esélyeit, ami intézményi és rendszerszinten is kihívásokat vet fel.

A szakirodalom alapján a munkaelégedettség olyan kulcsfontosságú tényezőként értelmezhető, amely mérsékelheti a munkahelyi stressz kedvezőtlen hatásait, és hozzájárulhat a pedagógusok jóllétének fenntartásához (Skaalvik & Skaalvik, 2014; Tria, 2023). Ennek pedagógiai

implikációja, hogy az oktatási intézmények egészségvédelmi és jóllétfókuszú megközelítései nem szorítkozhatnak kizárólag egyéni stresszkezelő technikákra, hanem szükségszerűen ki kell terjedniük a munkahelyi környezet, a szervezeti kultúra és a szakmai autonómia kérdéseire is. A támogató iskolai klíma, a kollegiális együttműködés és a pedagógiai munka értelmességének megélése olyan védő tényezők lehetnek, amelyek erősítik a munkaelégedettséget és csökkentik a stressz egészségkárosító hatásait (Aldridge & Fraser, 2016).

A bemutatott kutatási eredmények arra is rámutatnak, hogy a pedagógusok egészségvédelme szempontjából indokolt a preventív szemlélet erősítése. Ez magában foglalhatja a pedagógusképzés és továbbképzés során a stresszel és jólléttel kapcsolatos témák hangsúlyosabb megjelenítését, valamint olyan intézményi szintű beavatkozások kialakítását, amelyek elősegítik a munkahelyi terhelések kiegyensúlyozását. Az online vagy rövid intervenciók programok ígéretes lehetőséget kínálnak, ugyanakkor hatékonyságuk pedagógiai kontextusban történő hosszú távú vizsgálata továbbra is fontos kutatási feladat (Avola et al., 2025).

Összességében elmondható, hogy a pedagógusok és más segítő hivatásúak munkahelyi stressze, munkaelégedettsége és egészségmutatói szoros összefüggésben állnak, ezért a jövőben elengedhetetlen a holisztikus szemléletű megközelítés, az egyéni és szervezeti tényezők együttes kezelése a szakemberek jóllétének biztosítása érdekében.

Irodalom

- Admiraal, W., & Kittelsen Røberg, K.-I. (2023). Teachers' job demands, resources and their job satisfaction: Satisfaction with school, career choice and teaching profession of teachers in different career stages. *Teaching and Teacher Education*, 125, 104063. <https://doi.org/10.1016/j.tate.2023.104063>
- Al Shuaili, A. S. T. (2025). Occupational stress and job satisfaction: Unveiling challenges in teachers' and doctors' work environments. *World Journal of Education*, 15(1), 37–52. <https://doi.org/10.5430/wje.v15n1p37>
- Aldridge, J. M., & Fraser, B. J. (2016). Teachers' views of their school climate and its relationship with teacher self-efficacy and job satisfaction. *Learning Environments Research*, 19(2), 291–307. <https://doi.org/10.1007/s10984-015-9198-x>
- Aljawarneh, Y., Al-Bashaireh, A., Albikawi, Z., Alkouri, O., Alrida, N. A., Abuadas, M., Alobaidly, A., Alyileili, M., Alyileili, S., Alyileili, W., Alhefetti, F., Alhefeti, A., Alshehri, A., Alharbi, A., & Qaladi, O. (2025). The influence of occupational stress and job satisfaction on burnout among healthcare workers in the UAE: A cross-sectional study. *PLOS One*, 20(10), e0335430. <https://doi.org/10.1371/journal.pone.0335430>
- Aloe, A. M., Shisler, S. M., Norris, B. D., Nickerson, A. B., & Rinker, T. W. (2014). A multivariate meta-analysis of student misbehavior and teacher burnout. *Educational Research Review*, 12, 30–44. <https://doi.org/10.1016/j.edurev.2014.05.003>
- Amiri, S., & Behnezhad, S. (2020). Is job strain a risk factor for musculoskeletal pain? A systematic review and meta-analysis of 21 longitudinal studies. *Public Health*, 181, 158–167. <https://doi.org/10.1016/j.puhe.2019.11.023>
- Amiri, S., Mahmood, N., Mustafa, H., Javaid, S. F., & Khan, M. A. B. (2024). Occupational risk factors for burnout syndrome among healthcare professionals: A global systematic review and meta-analysis. *International Journal of Environmental Research and Public Health*, 21(12), 1583. <https://doi.org/10.3390/ijerph21121583>
- Andersen, L. P., Pihl-Thingvad, J., & Andersen, D. R. (2025). How superiors support employees to manage emotional demands: A qualitative study. *International Journal of Environmental Research and Public Health*, 22(5), 670. <https://doi.org/10.3390/ijerph22050670>

- Arnold, B., & Rahimi, M. (2025). Teachers' working conditions, wellbeing and retention: An exploratory analysis. *Australian Educational Researcher*, 52, 1947–1973. <https://doi.org/10.1007/s13384-024-00794-1>
- Avola, P., Soini-Ikonen, T., Jyrkiäinen, A., & Pentikäinen, V. (2025). Interventions to teacher well-being and burnout: A scoping review. *Educational Psychology Review*, 37(1), 11. <https://doi.org/10.1007/s10648-025-09986-2>
- Bakker, A. B., & Demerouti, E. (2017). Job demands–resources theory: Taking stock and looking forward. *Journal of Occupational Health Psychology*, 22(3), 273–285. <https://doi.org/10.1037/ocp0000056>
- Bauer, J., Stamm, A., Virnich, K., Wissing, K., Müller, U., Wirsching, M., Schaarschmidt, U. (2006). Correlation between burnout syndrome and psychological and psychosomatic symptoms among teachers. *International Archives of Occupational and Environmental Health*, 79(3), 199–204. <https://doi.org/10.1007/s00420-005-0050-y>
- Benevene, P., De Stasio, S., & Fiorilli, C. (2020). Editorial: Well-being of school teachers in their work environment. *Frontiers in Psychology*, 11, 1239. <https://doi.org/10.3389/fpsyg.2020.01239>
- Berez, N., & Mandel, K. M. (2022). Bölcsoédei és óvodai pedagógusok mentális állapota. *Educatio*, 31(1), 131–139. <https://doi.org/10.1556/2063.31.2022.1.11>
- Brown, S. D., & Lent, R.W. (Eds.). (2008). *Handbook of counselling psychology* (4th edition). John Wiley & Sons.
- Chen, J. J., Li, Z., Rodrigues, W., & Kaufman, S. B. (2022). Thriving beyond resilience despite stress: A psychometric evaluation of the newly developed Teacher Stress Scale and Teacher Thriving Scale. *Frontiers in Psychology*, 13, 862342. <https://doi.org/10.3389/fpsyg.2022.862342>
- Chunga Diaz, T. O., Santiago, G. B., Fuentes Calcino, A. R., Mèndez Ilizarbe, G. S., Quispilay Joyos, G. E., & Leiva Torres, J. G. (2024). Work stress in health professionals: A systematic review. *Revista de Gestão Social e Ambiental*, 18(4), e04575. <https://doi.org/10.24857/rgsa.v18n4-018>
- Collie, R. J., Shapka, J. D., & Perry, N. E. (2012). School climate and social–emotional learning: Predicting teacher stress, job satisfaction, and teaching efficacy. *Journal of Educational Psychology*, 104(4), 1189–1204. <https://doi.org/10.1037/a0029356>
- Elekes, Gy., Svraka, B., Gombás, J., & Gyarmathy, É. (2025). A kiégés és az egészség dimenziói a hazai pedagógusok körében. In V. Bócsi & K. Kovács (Eds.), *Új kutatások a neveléstudományokban 2024. A nevelés és oktatás kihívásai a válságok korában*. MTA Pedagógiai Tudományos Bizottság, Debreceni Egyetem Gyermeknevelési és Gyógypedagógiai Kar, Debreceni Egyetem BTK Nevelés- és Művelődéstudományi Intézete. https://www.edu.u-szeged.hu/ces2025/download/uj_kihivasok_2024_konfktotet_honlap_2.pdf#page=12
- Emeljanovas, A., Sabaliauskas, S., Mežienė, B., & Istomina, N. (2023). The relationships between teachers' emotional health and stress coping. *Frontiers in Psychology*, 14, 1276431. <https://doi.org/10.3389/fpsyg.2023.1276431>
- Estrada-Araoz, E. G., Quispe, J. A. B., Reyes, B. V., Ponce, W. C. C., Limachi, D. J. I., Aquino, V. S., Valverde, Y. P., & Herrera, R. Q. (2023). Evaluation of psychosomatic symptoms associated with stress in teachers after returning to face-to-face classes. *International Journal of Learning, Teaching and Educational Research*, 22(4), 362–376. <https://doi.org/10.26803/ijlter.22.4.21>
- Faragher, E. B., Cass, M., & Cooper, C. L. (2005). The relationship between job satisfaction and health: A meta-analysis. *Occupational and Environmental Medicine*, 62(2), 105–112. <https://doi.org/10.1136/oem.2002.006734>
- Gámez-Genovart, M., Oliver-Trobat, M. F., & Rosselló-Ramon, M. R. (2025). Research on teacher wellbeing: A systematic review. *Teaching and Teacher Education*, 168, 105264. <https://doi.org/10.1016/j.tate.2025.105264>
- Ganster, D. C., & Rosen, C. C. (2013). Work stress and employee health: A multidisciplinary review. *Journal of Management*, 39(5), 1085–1122. <https://doi.org/10.1177/0149206313475815>
- Ghadban, M., Berkowitz, R., & Enosh, G. (2023). Teachers' workplace victimization, job burnout, and somatic and posttraumatic symptoms: A structural equation modeling examination. *Education Sciences*, 13(2), 179. <https://doi.org/10.3390/educsci13020179>

- Gmayinaam, V. U., Nortey, A. N., Sedode, S., Apedo, S. K., Kye-Duodu, G., Kwabla, P., Osei, E., & Ananga, M. K. (2024). Work-related stress among nurses: A comparative cross-sectional study of two government hospitals in Ghana. *BMC Public Health*, 24, 2257. <https://doi.org/10.1186/s12889-024-19757-3>
- Granziera, H., & Perera, H. N. (2019). Relations among teachers' self-efficacy beliefs, engagement, and work satisfaction: A social cognitive perspective. *Contemporary Educational Psychology*, 58, 75–84. <https://doi.org/10.1016/j.cedpsych.2019.02.003>
- Gu, B., Tan, Q., & Zhao, S. (2019). The association between occupational stress and psychosomatic wellbeing among Chinese nurses: A cross-sectional survey. *Medicine*, 98(22), e15836. <https://doi.org/10.1097/MD.00000000000015836>
- Gyarmati, Z., Müller, A., & Biró, M. (2020). A kiegész problémája a pedagógus pályán. *Gradus*, 7(3), 110–116. <https://doi.org/10.47833/2020.3.ART.006>
- Harvey, S. B., Modini, M., Joyce, S., Milligan-Saville, J. S., Tan, L., Mykletun, A., & Mitchell, P. B. (2017). Can work make you mentally ill? A systematic meta-review of work-related mental health risks. *Occupational and Environmental Medicine*, 74(4), 301–310. <https://doi.org/10.1136/oemed-2016-104015>
- Hassard, J., Teoh, K. R. H., Visockaite, G., Dewe, P., & Cox, T. (2018). The cost of work-related stress to society: A systematic review. *Journal of Occupational Health Psychology*, 23(1), 1–17. <https://doi.org/10.1037/ocp0000069>
- Herman, K. C., Prewett, S. L., Eddy, C. L., Savala, A., & Reinke, W. M. (2020). Profiles of middle school teacher stress and coping: Concurrent and prospective correlates. *Journal of School Psychology*, 78, 1–15. <https://doi.org/10.1016/j.jsp.2019.11.003>
- Hricova, M., Nezkusilova, J., & Raczova, B. (2020). Perceived stress and burnout: The mediating role of psychological, professional self-care and job satisfaction as preventive factors in helping professions. *European Journal of Mental Health*, 15(1), 3–22. <https://doi.org/10.5708/EJMH.15.2020.1.1>
- Jiang, Y., & Yuan, Q. (2025). How job satisfaction affects teachers' occupational burnout: The mediating role of psychological capital moderated by physical exercise. *Acta Psychologica*, 260, 105594. <https://doi.org/10.1016/j.actpsy.2025.105594>
- Jin, P., Yeung, A. S., Tang, T.-O., & Low, R. (2008). Identifying teachers at risk in Hong Kong: Psychosomatic symptoms and sources of stress. *Journal of Psychosomatic Research*, 65(4), 357–362. <https://doi.org/10.1016/j.jpsychores.2008.03.003>
- Jung, Y. J., & Kim, H. (2025). Job satisfaction, quality of life, and turnover intention among nurses: A comparative study of pattern-based and rotating shift schedules. *Healthcare*, 13(20), 2551. <https://doi.org/10.3390/healthcare13202551>
- Klassen, R. M., & Chiu, M. M. (2010). Effects on teachers' self-efficacy and job satisfaction: Teacher gender, years of experience, and job stress. *Journal of Educational Psychology*, 102(3), 741–756. <https://doi.org/10.1037/a0019237>
- Koltói, L., & Kiss, P. (2024). A szakmai és szervezeti elköteleződés szerepe a munkahelyről való kilépés szándékában tanárok körében. *Magyar Pszichológiai Szemle*, 79(3), 611–634. <https://doi.org/10.1556/0016.2024.00100>
- Kovács, K., Borbély, Sz., Dobay, B., Halasi, Sz., Vajda, I., & Hideg-Fehér, G. (2024). Közép-kelet-európai oktatók jóllétét befolyásoló tényezők. In K. Kovács & K. Pallay (Eds.), *Oktatók a Kárpát-medencei felsőoktatásban* (pp. 15–31). CHERD-HUNGARY. https://real.mtak.hu/224069/7/15_realoktataskutatok_20_beliv_nyomdai.pdf
- Kurrlé, L. M., & Warwas, J. (2025). Teacher well-being—A conceptual systematic review (2020–2023). *Education Sciences*, 15(6), 766. <https://doi.org/10.3390/educsci15060766>
- Kyriacou, C. (2001). Teacher stress: Directions for future research. *Educational Review*, 53(1), 27–35. <https://doi.org/10.1080/00131910120033628>
- Lesener, T., Gusy, B., & Wolter, C. (2019). The job demands–resources model: A meta-analytic review. *Work & Stress*, 33(1), 76–103. <https://doi.org/10.1080/02678373.2018.1529065>

- Li, Y., Chen, C., & Yuan, Y. (2025). Evolving the job demands-resources framework to JD-R 3.0: The impact of after-hours connectivity and organizational support on employee psychological distress. *Acta Psychologica*, 253, 104710. <https://doi.org/10.1016/j.actpsy.2025.104710>
- Llorca-Pellicer, M., Gil-LaOrden, P., Prado-Gascó, V. J., & Gil-Monte, P. R. (2025). The role of psychosocial risks in burnout, psychosomatic disorders, and job satisfaction: Lineal models vs a QCA approach in non-university teachers. *Psychology & Health*, 40(4), 681–695. <https://doi.org/10.1080/08870446.2023.2253258>
- Madigan, D. J., Kim, L. E., Glandorf, H. L., & Kavanagh, O. (2023). Teacher burnout and physical health: A systematic review. *International Journal of Educational Research*, 119, 102173. <https://doi.org/10.1016/j.ijer.2023.102173>
- Marcatto, F., Patriarca, E., Bramuzzo, D., Lucci, E., & Filon, F. L. (2024). Investigating the role of organizational stress in nurses' psychosomatic complaints: Evidence from a study in northeastern Italy. *AIMS Public Health*, 11(2), 420–431. <https://doi.org/10.3934/publichealth.2024021>
- Martela, F., & Riekk, T. J. J. (2018). Autonomy, competence, relatedness, and beneficence: A multicultural comparison of the four pathways to meaningful work. *Frontiers in Psychology*, 9, 1157. <https://doi.org/10.3389/fpsyg.2018.01157>
- Martínez-Libano, J. (2025). Teachers' Emotional Exhaustion Scale (TEES): Development and psychometric validation. *Education Sciences*, 15(6), 742. <https://doi.org/10.3390/educsci15060742>
- Masa, A., Fiedler, O., Derzsi-Horváth, M., Szabó A. A., Tobak, O., & Deutsch, K. (2022). Pedagógusok mentális egészségének komplex vizsgálata. *Sport- és Egészségtudományi Füzetek*, 6(4), 73–91. <https://publicatio.bibl.u-szeged.hu/26840/1/MasaAndrea2022.pdf>
- Maslach, C., & Leiter, M. P. (2016). Understanding the burnout experience. *World Psychiatry*, 15(2), 103–111. <https://doi.org/10.1002/wps.20311>
- Maulik, P. K. (2017). Workplace stress: A neglected aspect of mental health wellbeing. *Indian Journal of Medical Research*, 146(4), 441–444. https://doi.org/10.4103/ijmr.IJMR_1298_17
- Mihálka, M., & Pikó, B. (2018). Study of work–family conflict (WFC), burnout and psychosocial health among Hungarian educators. *Central European Journal of Occupational and Environmental Medicine*, 24(1–2), 83–95. https://nnk.gov.hu/cejoem/Volume24/Vol24No1-2/24_1-2_Article_08.pdf
- Molnár, A., & Fodor, Sz. (2024) Pedagógusok karaktererősségeinek és munkahelyi jóllétének feltáró elemzése. *Iskolakultúra*, 34(9), 39–60. <https://doi.org/10.14232/iskkult.2024.9.39>
- Niedhammer, I., Bertrais, S., & Witt, K. (2021). Psychosocial work exposures and health outcomes: A meta-review of 72 literature reviews with meta-analysis. *Scandinavian Journal of Work, Environment & Health*, 47(7), 489–508. <https://doi.org/10.5271/sjweh.3968>
- Nixon, A. E., Mazzola, J. J., Bauer, J., Krueger, J. R., & Spector, P. E. (2011). Can work make you sick? A meta-analysis of the relationships between job stressors and physical symptoms. *Work & Stress*, 25(1), 1–22. <https://doi.org/10.1080/02678373.2011.569175>
- Nwoko, J. C., Emeto, T. I., Malau-Aduli, A. E. O., & Malau-Aduli, B. S. (2023). A systematic review of the factors that influence teachers' occupational wellbeing. *International Journal of Environmental Research and Public Health*, 20(12), 6070. <https://doi.org/10.3390/ijerph20126070>
- Pachi, A., Sikaras, C., Melas, D., Alikanioti, S., Soultanis, N., Ivanidou, M., Ilias, I., & Tselebis, A. (2025). Stress, anxiety and depressive symptoms, burnout and insomnia among Greek nurses one year after the end of the pandemic: A moderated chain mediation model. *Journal of Clinical Medicine*, 14(4), 1145. <https://doi.org/10.3390/jcm14041145>
- Patai, F., & Németh, G. (2026). Empirikus kutatás a pedagógusi munkakör kiegészi okainak jobb megértésére. *Új Munkaügyi Szemle*, 7(1), 30–41. <https://doi.org/10.58269/umsz.2026.1.3>
- Ray, T. K. (2022). Work related well-being is associated with individual subjective well-being. *Industrial Health*, 60(3), 242–252. <https://doi.org/10.2486/indhealth.2021-0122>
- Ritvanen, T., Louhevaara, V., Helin, V. L., Halonen, T., & Hänninen, O. (2003). Psychophysiological stress in high school teachers. *International Journal of Occupational Medicine and Environmental Health*, 16(3), 255–264.
- Schonfeld, I. S., & Chang, C.-H. (2016). *Occupational health psychology. Work, stress, and health*. Springer. <https://doi.org/10.1891/9780826199683>

- Skaalvik, E. M., & Skaalvik, S. (2014). Teacher self-efficacy and perceived autonomy: Relations with teacher engagement, job satisfaction, and emotional exhaustion. *Psychological Reports, 114*(1), 68–77. <https://doi.org/10.2466/14.02.PR0.114k14w0>
- Skaalvik, E. M., & Skaalvik, S. (2017). Motivated for teaching? Associations with school goal structure, teacher self-efficacy, job satisfaction and emotional exhaustion. *Teaching and Teacher Education, 67*, 152–160. <https://doi.org/10.1016/j.tate.2017.06.006>
- Stosic, M. D., Blanch-Hartigan, D., Aleksanyan, T., Duenas, J., & Ruben, M. A. (2022). Empathy, friend or foe? Untangling the relationship between empathy and burnout in helping professions. *The Journal of Social Psychology, 162*(1), 89–108. <https://doi.org/10.1080/00224545.2021.1991259>
- Sukhera, J. (2022). Narrative reviews: Flexible, rigorous, and practical. *Journal of Graduate Medical Education, 14*(4), 414–417. <https://doi.org/10.4300/JGME-D-22-00480.1>
- Szabó, É., Bátor, B., Bobor, P., Korláth, P., Szappanos, C., & Jagodics, B. (2020). Mi van a kiégés-fogékonyság mögött? Pedagógusok munkahelyi nehézségeinek feltárása kvalitatív és kvantitatív eljárással. *Iskolakultúra, 30*(3), 3–17. <https://doi.org/10.14232/ISKKULT.2020.3.3>
- Szabó, É., & Fehérváry, G. (2023). Hivatás vagy kényszer? A gyakorló pedagógusok pályaelköteleződési motivációjának vizsgálata. *Iskolakultúra, 33*(5), 3–20. <https://doi.org/10.14232/iskcult.2023.5.3>
- Szabó, É., & Jagodics, B. (2019). Teacher burnout in the light of workplace, organizational, and social factors. *Hungarian Educational Research Journal, 9*(3), 539–559. <https://doi.org/10.1556/063.9.2019.3.44>
- Szelezsánné Egyed, D. (2016). A pedagógusok körében megjelenő stressz, pszichoterror és kiégés jelensége; prevenció és kezelési lehetőségek a köznevelési intézményben. *Opus et Educatio, 3*(5). <https://doi.org/10.3311/ope.134>
- Tóthpál-Halasi, O., & Pikó, B. (2024). Pedagógusok kiégésének és jóllétének vizsgálata a munkamotiváció dimenzióinak tükrében. *Iskolakultúra, 34*(9), 61–75. <https://doi.org/10.14232/iskcult.2024.9.61>
- Tria, J. Z. (2023). Job satisfaction among educators: A systematic review. *International Journal of Professional Development, Learners and Learning, 5*(2), ep2310. <https://doi.org/10.30935/ijpddl/13212>
- von der Embse, N. P., Ryan, S. V., Gibbs, T., & Mankin, A. (2019). Teacher stress interventions: A systematic review. *Psychology in the Schools, 56*(8), 1328–1343. <https://doi.org/10.1002/pits.22279>
- Wang, J., Li, J., Wu, J., & Fan, X. (2024). Measurement invariance of scores on the Teacher Stress Scale: An international sample of PreK–12 teachers. *Psychology in the Schools, 62*(1), 149–166. <https://doi.org/10.1002/pits.23317>
- Wang, Q., Ding, J., Dai, Y., Yang, S., & Zhou, Z. (2025). Occupational stress, burnout, and fatigue among healthcare workers in Shanghai, China: A questionnaire-based cross-sectional survey. *Healthcare, 13*(13), 1600. <https://doi.org/10.3390/healthcare13131600>
- Woods, S., Sebastian, J. C., Herman, K. C., Huang, F. L., Reinke, W. M., & Thompson, A. M. (2023). The relationship between teacher stress and job satisfaction as moderated by coping. *Psychology in the Schools, 60*(10), 3459–3474. <https://doi.org/10.1002/pits.22857>
- World Health Organization (2017). *Occupational health. Stress at the workplace*. http://www.who.int/occupational_health/topics/stressatwp/en
- Yang, Z., Wang, D., Fan, Y., Ma, Z., Chen, X., Zhang, Y., & Fan, F. (2023). Relationship between sleep disturbance and burnout among Chinese urban teachers: Moderating roles of resilience. *Sleep Medicine, 108*, 29–37. <https://doi.org/10.1016/j.sleep.2023.06.002>
- Zhang, X., Smith, D. R., Zheng, Y., & Wang, R.-S. (2011). Occupational stress and psychosomatic complaints among health professionals in Beijing, China. *Work, 40*(2), 239–245. <https://doi.org/10.3233/WOR-2011-1224>
- Zhao, W., Liao, X., Li, Q., Jiang, W., & Ding, W. (2022). The relationship between teacher job stress and burnout: A moderated mediation model. *Frontiers in Psychology, 12*, 784243. <https://doi.org/10.3389/fpsyg.2021.784243>

ABSTRACT

A NARRATIVE REVIEW OF TEACHERS' OCCUPATIONAL STRESS, JOB SATISFACTION AND PSYCHOSOMATIC STRESS INDICATORS

Szergej Malhazjan & Bettina Pikó

Keywords: occupational stress, job satisfaction, psychosomatic stress indicators, teachers, helping professions

Professionals working in helping professions—particularly teachers—are exposed to increased psychological demands, which may have adverse long-term effects on psychological well-being, job satisfaction, and overall health. Therefore, examining the relationships between occupational stress, job satisfaction, and stress-related psychosomatic symptoms is of particular importance from both pedagogical and preventive perspectives. The present study provides a narrative literature review of the main research directions concerning teachers and other helping professionals. Its aim is to explore how the literature conceptualizes and measures occupational stress, how it positions job satisfaction within the relationship between stress and health outcomes, and which psychosomatic stress indicators are most commonly applied in empirical investigations. The review also highlights inconsistencies and gaps within the existing body of research, with particular attention to the specific characteristics of the teaching profession. The findings may contribute to the development of preventive approaches and institution-level interventions aimed at protecting teachers' health and well-being.

Magyar Pedagógia, 126(2). 123–140. (2026)
<https://doi.org/10.14232/mped.2026.2.123>

Malhazjan Szergej:  <https://orcid.org/0009-0001-6125-9708>
Szegedi Tudományegyetem, Neveléstudományi Doktori Iskola
H-6722 Szeged, Petőfi Sándor sgt. 32–34.
malhazjan.szergej@edu.u-szeged.hu

Pikó Bettina:  <https://orcid.org/0000-0002-3072-9615>
Szegedi Tudományegyetem, Magatartástudományi Intézet
H-6722 Szeged, Szentháromság u. 5.
fuzne.piko.bettina@med.u-szeged.hu



A Magyar Pedagógia folyóirat 2020-as évfolyamának számaitól
kizárólag online formában jelenik meg.

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: Börcsökkné Soós Edit.

Megjelent 4,8 (B/5) ív terjedelemben.

HU ISSN 0025–0260

KÖZLÉSI FELTÉTELEK

A *Magyar Pedagógia* a tágan értelmezett neveléstudomány és a határtudományok minden területéről közöl tanulmányokat, empirikus vizsgálat eredményeit közlő cikket, egy kutatási terület eredményeinek szintetizáló bemutatását, és az oktatás problémáival foglalkozó elméleti elemzést egyaránt.

A *Magyar Pedagógia* csak eredeti, másutt még nem publikált tanulmányokat közöl. A kézirat benyújtásával a szerző vállalja, hogy írását korábban még nem jelentette meg, párhuzamosan más folyóirathoz nem nyújtja be. A *Magyar Pedagógiában* való megjelenés szempontjából nem számít előzetes publikációnak a zárt körben, kézíratos sokszorosításként való terjesztés (belső kiadvány, kutatási zárójelentés, konferencia előadás stb.). A megjelent tanulmányok szerzői megőrzik azt a jogukat, hogy tanulmányukat a *Magyar Pedagógiában* való megjelenés után másutt (gyűjteményes kötetben, más nyelven stb.) újra közöljék.

A kéziratokat magyar, vagy különlegesen indokolt esetben angol nyelven lehet benyújtani. Az elfogadott angol nyelvű kéziratok fordításáról a szerkesztőség gondoskodik. A tanulmányok optimális terjedelme 10–20 nyomtatott oldal (30000–60000 betű). Az angol nyelvű abstract számára 150–250 szavas összegzést kell mellékelni angol nyelven.

A beérkezett kéziratokat a szerkesztőség a tudományos folyóiratoknál megszokott bírálati eljárás keretében véleményezi. A folyóirat témakörébe eső cikkek közlésének kizárólagos szempontja a munka színvonala.

A kéziratokat elektronikus formában a magyarpedagogia.hu címen elérhető szerkesztőségi rendszerbe kell feltölteni.

AIMS AND SCOPE

Established in 1892 and published quarterly, *Magyar Pedagógia* is the journal of the Educational Committee of the Hungarian Academy of Sciences. It publishes original reports of empirical research, theoretical contributions and synthetic research reviews within the field of Education in the broadest sense and the related social sciences. The journal publishes articles in Hungarian accompanied by an abstract in English. *Magyar Pedagógia* seeks to provide a forum for communication between the Hungarian and international research communities. Therefore, the Editorial Board encourages international authors to submit their manuscripts for consideration.

Submitted manuscripts will be subjected to a peer review process. Selection is based exclusively on the scientific quality of the work. Only original manuscripts will be considered. Manuscripts which have been published previously or are currently under consideration elsewhere will not be reviewed for publication in *Magyar Pedagógia*. However, authors retain their rights to reprint their article after it has appeared in this journal.

Manuscripts should be preferably in Hungarian or in English. Papers should be between 10–20 printed pages (ca. 30000–60000 characters) and accompanied by a 150–250 word abstract. Manuscripts submitted in English should be prepared in accordance with the Publicational Manual of APA.

Manuscripts must be submitted to the editorial system at magyarpedagogia.hu.

RESEARCH PAPERS

Erika Majoros: Trend Scaling Methods: Linking Recent and Older IEA Studies on Mathematics and Science	83
Andrea Magyar, Anita Habók & Gyöngyvér Molnár: Technology-Based Vocabulary Development and Reading Comprehension: Results From a Large-Scale Study	99
Szergej Malhazjan & Bettina Pikó: A Narrative Review of Teachers' Occupational Stress, Job Satisfaction and Psychosomatic Stress Indicators	123