



A TESZTMEGOLDÁSI MOTIVÁCIÓ KÉRDŐÍVES ÉS LOGADAT ALAPÚ MÉRÉSÉNEK ÖSSZEHASONLÍTÓ ELEMZÉSE ALACSONY TÉTEL RENDELKEZŐ INTERAKTÍV PROBLÉMAMEGOLDÓ KÖRNYEZETBEN

Csányi Róbert * és Molnár Gyöngyvér **

** SZTE Neveléstudományi Doktori Iskola*

*** SZTE Neveléstudományi Intézet, MTA-SZTE Digitális Tanulási Technológiák Kutatócsoport*

Az oktatási folyamat során alkalmazott értékelési eljárások célja a tanulók tudásszintjének, az elsajátított ismeretek különböző helyzetekben való alkalmazásának meghatározása. Ehhez olyan mérőeszközök szükségesek, amelyek megfelelően reprezentálják a mérni kívánt tudásterületeket, illetve illeszkednek a diákok tudás- és képességszintjéhez. A tesztek eredményei tanulói, tanulócsoporti, intézményi, illetve az oktatási rendszer szintjén használhatók fel a mért tulajdonságok jellemzésére.

A mérés-értékelés célja, megközelítése erőteljes változáson esett át az elmúlt években (Molnár, 2021). A szummatív értékelés mellett egyre inkább előtérbe kerül a személyre szabott tanulást és tanítást támogató, tanulói szintű visszacsatolásra fókuszáló diagnosztikus értékelés. A diagnosztikus értékelés pedig jó alapul szolgál a tanulók fejlesztő értékeléséhez (Csapó & Molnár, 2019).

A tanulók teszten nyújtott teljesítményét nemcsak tudás- és képességszintjük határozza meg, hanem erőteljesen befolyásolhatják különböző affektív tényezők, többek között a motiváció részleges vagy teljes hiánya (Wise et al., 2014). Wise és DeMars (2005) elemzése alapján a nem motivált diákok több mint fél szórással gyengébb eredményt értek el a teszteken, mint motivált társaik. Ezt támasztják alá Finn (2015), Schüttpelz-Brauns és munkatársai (2018), illetve Wise és Kong (2005) kutatási eredményei is, melyek a motiváltabb tesztmegoldók magasabb teljesítményét jelezték. Lee és Chen (2011) kutatási eredményei szerint nem mutatható ki jelentős összefüggés a tesztelt személyek motivációja és képességszintje között, azaz nem igaz az az állítás, hogy a magasabb képességszintűek esetleg erőteljesebb motivációval rendelkeznek, illetve az alacsonyabb tudásszintűek motivációja alacsonyabb lenne, vagy fordítva. A diákok tudás- és képességszintje, valamint a teszteredményt befolyásoló motiváció kapcsolata nem lineáris.

A tesztek tétje: alacsony és magas téttel rendelkező tesztek

A tesztelt személy számára a teszt tétje definíció szerint azt jelenti, hogy milyen szintű következményekkel jár a teszten nyújtott teljesítmény a tesztelt személy előrehaladása, jövője szempontjából, azaz mekkora tétje van a teszten nyújtott teljesítménynek (Csapó et al., 2008) a tanulmányai, végső értékelése, továbbtanulása, előrehaladása tekintetében. Az alacsony téttel rendelkező tesztek eredménye nincs jelentős hatással az adott személy előmenetelére, míg a magas téttel rendelkező teszteken nyújtott teljesítmény meghatározó következményekkel bír. A magas téttel rendelkező tesztek közé tartozik például a nyelvvizsga, az érettség, míg a diákok számára alacsony téttel rendelkező tesztek az országos kompetenciamérések és a nemzetközi oktatási felmérések. Fontos hangsúlyozni, hogy ugyan a tesztmegoldó számára nincs releváns következménye egy alacsony téttel rendelkező teszt eredményének, azonban ezek az oktatási rendszer több szintje számára szolgálhatnak adatot, és az egyéni pontszámokból magasabb rendszerszinten megszülető pontszámok, teljesítménymutatók már komoly következményekkel bírhatnak intézményi szinten, vagy oktatási rendszer szintjén (Lindner et al., 2019).

A tesztek tétje jelentősen befolyásolja a tesztelt személyek tesztmegoldási motivációját: a tét növekedésével növekszik a kifejtett erőfeszítés, de ezzel párhuzamosan növekszik annak a valószínűsége is, hogy a tesztmegoldók nem megengedett segédeszközöket használnak, vagy a túl nagymértékű izgulás negatívan hathat teljesítményükre. A teszt tétjének, szerepének csökkenésével arányosan csökkenhet a tesztmegoldási motiváció, ami hatással lehet a tesztmegoldók teljesítményére (Rios, 2021).

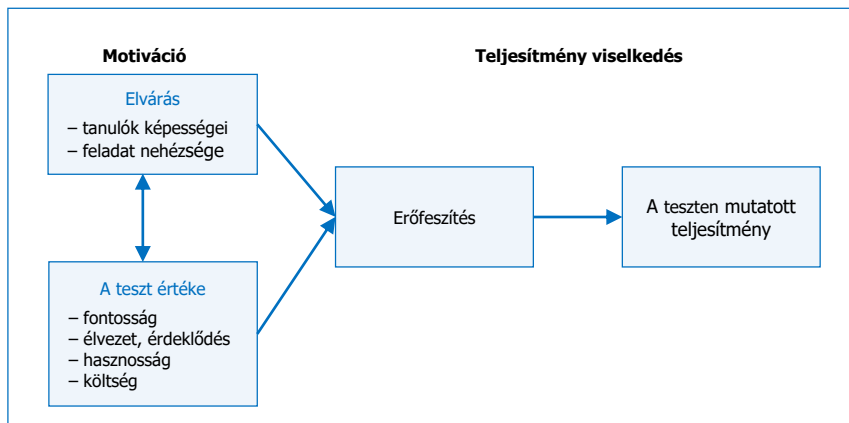
A tesztmegoldási motiváció: elvárás-érték elmélet

A tesztmegoldási motiváció értelmezésére gyakran használt megközelítés az elvárás-érték elmélet¹ (*Expectancy-value theory*; Wise & DeMars, 2005). A modell szerint a motiváció mértékét egyrészt az elvárás és az azon alapuló siker valószínűsége, másrészt a teszt értéke határozza meg. Az elvárás-érték elmélet tesztmegoldási motivációra alkalmazott modelljét az 1. ábra mutatja.

A tesztmegoldók saját teljesítményükkel kapcsolatos elvárásait egyrészt a saját képességeikről alkotott véleményük, másrészt a feladatok nehézsége határozza meg. A teszt hasznossága kapcsán Eccles és Wigfield (2002) négy komponenst különböztetett meg: (1) a teszt fontossága, (2) a teszt iránti érdeklődés, annak élvezete, (3) a teszt hasznossága és (4) a teszt költsége (pl. a feladatra fordított idő, tesztizorongás). Fontos hangsúlyozni, hogy a tesztmegoldási motiváció nem azonos a tanulási motivációval.

¹ A magyarul általában elváráselméletként fordított expectancy theory of motivation jellemzően Vroom (1964) munkájára utal. Az általunk hivatkozott expectancy-value theoryt jellemzően elvárás-érték elméletként, vagy érték-elvárás elméletként fordítják, lásd erről Abari és Polonyi (2017), Mezei (2013), É. Molnár (2009).

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben



1. ábra

*Elvárás-érték elmélet a tesztmegoldási motiváció kontextusában
(Penk & Richter, 2017 alapján)*

A tesztmegoldási motivációt számos tényező befolyásolja, ezek közül vizsgálatunk szempontjából a tesztet megoldók személye bír relevanciával. Goldhammer és munkatársai (2016) felnőttek vizsgálata során (PIAAC 2012) a férfiak esetében összességében alacsonyabb tesztmegoldási motivációt mutattak ki, mint a nőknél. Azonban a tesztmegoldási motiváció országonként és tanulmányi területek szerint is változott. Néhány ország esetében a problémamegoldást vizsgáló feladatoknál fordított volt a helyzet, a nők voltak kevésbé motiváltak. Wise és DeMars (2010) egyetemisták között vizsgálták a tesztmegoldási motivációt. Szignifikánsan alacsonyabb tesztmegoldási motivációt mutattak ki a férfiaknál, valamint a magasabb évfolyamba járó hallgatók között.

A tesztmegoldási motiváció mérési lehetőségei

A tesztmegoldási motiváció mérésére alkalmazott módszerek közül két nagy csoportot állítottunk fókuszba: (1) önértékelő kérdőíveken alapuló módszerek, (2) válaszdíő alapú módszerek. Az önértékelő kérdőívek a vizsgázók saját megítélésén alapulnak, a válaszdíő alapú megközelítés a számítógépes tesztek naplóadatait (*logdata*) használja. A felhasznált logadat a feladatokkal töltött idő, ami kiegészülhet egyéb adatokkal (pl. kattintások száma, típusa).

Az önértékelő kérdőívek a legrégebb óta és legszélesebb körben alkalmazott megoldások a tesztmegoldási motiváció mérésére, melyek jellemzően Likert-skálán mérik a tesztmegoldó motivációjának összetevőit. Az önértékelő kérdőívek fontos előnye, hogy használatuk egyszerű, hagyományos, papíralapú tesztelés során is megvalósítható, valamint feldolgozásuk is könnyen megvalósítható. Számos korlátjuk közül kiemelendő, hogy szubjektív, illetve nem tudhatjuk, mennyire volt őszinte a tesztmegoldók válasza, hiszen azt sok tényező befolyásolhatja (Wise & Kong, 2005).

A válaszdó alapú módszerek az informatikai eszközök oktatásban való elterjedésével párhuzamosan indultak fejlődésnek. Ezen módszerek hatalmas előnye, hogy a tesztek megoldása során számos olyan kontextuális adatot tudunk rögzíteni (pl. klikkelések, feladatokkal eltöltött idő, előre- és hátraugrás, szemmozgás), amelyek a hagyományos papír-alapú értékelési rendszerekkel elképzelhetetlenek voltak, és ezek elemzése mélyebb összefüggésekre világíthat rá (Tóth et al., 2017). A válaszdó alapú módszerek azon a feltételezésen alapulnak, hogy az alacsony tesztmegoldási erőfeszítéssel rendelkező résztvevők kevesebb időt töltenek a feladatok megoldásával, ezért gyorsabban válaszolnak a kérdésekre, mint a magasabb szinten motiváltak (Wise & Kong, 2005).

Célok, kutatási kérdések

A kutatás során a tesztmegoldási motiváció (*test-taking motivation, TTM*) többféle módszerrel történő vizsgálatát, valamint a különböző adatfelvételi technikákon alapuló eredmények összehasonlító elemzését tűztük ki célul. A kutatás során a következő kérdésekre kerestünk választ: (1) Milyen összefüggés van a tesztmegoldási motiváció, a feladatokkal töltött idő, a kattintások száma és a teszten nyújtott teljesítmény között? (2) Nemek és életkor tekintetében van-e különbség a tesztmegoldási motiváció átlagos mértékében? (3) Változik-e a tesztmegoldási motiváció mértéke a teszt előrehaladása során? (4) Milyen tanulói profilok azonosíthatók a logadatok, illetve az önértékelő kérdőív válaszai alapján? (5) Az önértékelő kérdőíves és a logadat alapú módszerek mennyire feleltethetők meg egymásnak?

Módszerek

Minta

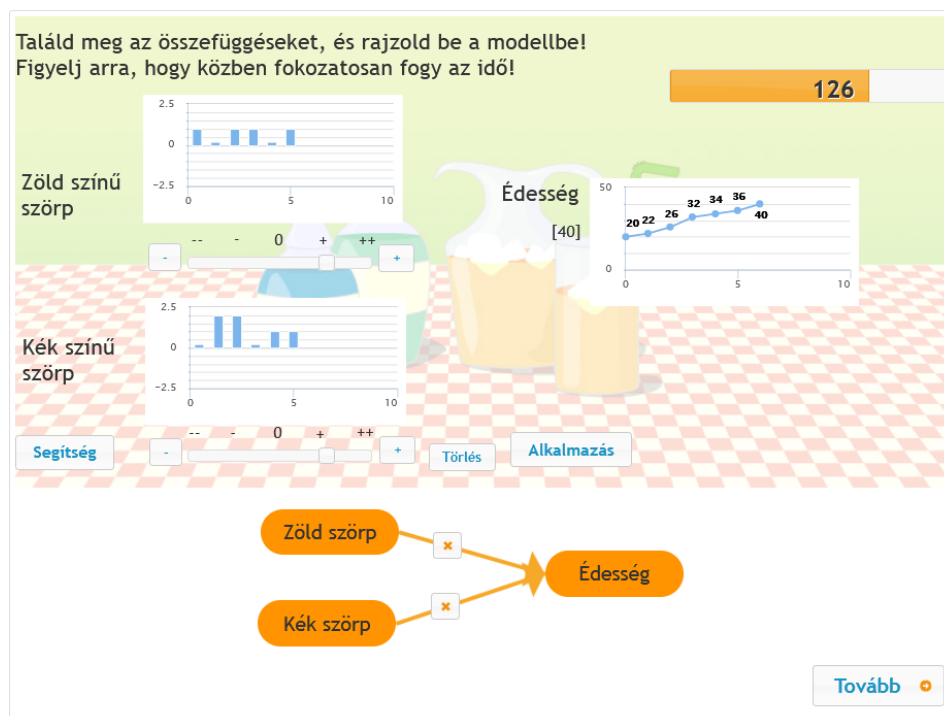
Az empirikus vizsgálatot a kizárólag gondolkodási képességek mérésére fókuszáló, alacsony tétellel bíró diagnosztikus mérések keretein belül valósítottuk meg 2021 tavaszán. A mintába 554 fő 6., 7. és 8. osztályos tanuló került, közülük 187 fő 6. osztályos (33,75%), 228 fő 7. osztályos (41,16%) és 139 fő 8. osztályos (25,09%) volt. Nemi megoszlás szerint 256 fő fiú (46,21%), 296 fő lány (53,43%), valamint 2 fő (0,36%) ezt az adatot nem adta meg. A vizsgálatban részt vevő tanulók átlagéletkora 13,64 év (SD=0,91) volt.

Mérőeszköz

Az adatfelvétel során a tanulónak 10, előzetes tudást nem igénylő interaktív problémát kellett feltérképezniük és megoldaniuk. A problémák mindegyike a MicroDYN modellen alapult (Greiff et al., 2013). Az interaktív problémák keretein belül a tanulók legfeljebb három bemeneti és három kimeneti változóval találkoztak. A problémamegoldás

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

első fázisában a változók között lévő összefüggéseket kellett kitalálniuk úgy, hogy szabadon változtathatták a bemeneti változók értékeit, majd megfigyelhették a változtatások hatását a kimeneti változók értékeire. Az így kapott és értelmezett információk alapján fel kellett rajzolniuk a kapcsolato(ka)t a bemeneti és a kimeneti változók között (Molnár & Csapó, 2018). A 2. ábrán bemutatott példában („Anyukádtól két újfajta szörpöt kaptál, amiket összekeverve még finomabb szörpöket tudsz magadnak kikeverni. Találd ki, hogyan befolyásolja a kikevert szörp édességét a zöld és kék színű szörp mennyisége!”) a tanulóknak először két bemeneti és egy kimeneti változó közötti összefüggést kellett felfedezniük. A tanulók a zöld és kék szörp értékeinek változtatásával tudták megfigyelni az édesség változását, majd a megfigyelt összefüggéseket felrajzolni a problémaszcenárió alatti modellbe.



2. ábra

Az interaktív problémamegoldás első szakasza: két bemeneti és egy kimeneti változó közötti kapcsolat feltérképezése

A feladatmegoldás következő fázisában („A következő oldalon megnézheted, hogy a valóságban hogyan befolyásolja a kikevert szörp édességét a zöld és a kék színű szörp mennyisége. Hogyan lehetne a kék és a zöld szirup mennyiségét úgy beállítani, hogy maximum négy lépésben elérjük a kikevert szörp előre meghatározott édességi fokát?”) a valódi összefüggések megismerése után a bemeneti változók értékeinek változtatásával

maximum négy lépésben el kellett érniük a kimeneti változók előre meghatározott értékeit. A példa esetében a zöld és a kék szörp megfelelő arányú adagolásával maximum négy lépésben beállítani a kikevert szörp előre meghatározott édességi fokát.

A tanulmányban az első fázis, az explorációs rész logadatait elemeztük. Ennek oka, hogy ebben a fázisban nem maximalizáltuk a lehetséges kattintások mennyiségét, illetve több idő állt a diákok rendelkezésére. Mindennek következtében mind az idő, mind a kattintás adatok szórása jelentősebbnek bizonyult, jobban differenciált, erőteljesebben megmutatja a különbségeket, mint a második fázis adatai, ahol rövidebb időkorlát alkalmazása mellett maximum négy beállítást próbálhattak ki a tanulók.

Önértékelő kérdőív

A tesztmegoldási motivációt mérő kérdőív kidolgozásához a következő, széles körben használt kérdőíveket elemeztük: (1) Online Motivációs Kérdőív (*Online Motivation Questionnaire; OMQ*) (Crombach et al., 2003), (2) Kérdőív a jelenlegi motivációról (*Questionnaire of Current Motivation; QCM*) (Freund et al., 2011), (3) Erőfeszítés hőmérő (*Effort-thermometer*) (Butler & Adams, 2007), (4) Tesztmegoldási erőfeszítés rövid skála (*Test-taking Effort Short Scale; TESS*) (Schüttpelz-Brauns et al., 2018), valamint (5) Hallgatói vélemény skála (*Student Opinion Scale; SOS*) (Thelk et al., 2009).

A kérdőív összeállítása során kiemelt hangsúlyt fektettünk azon részterületekre, amelyek relevánsak lehetnek az általunk vizsgált alacsony tétellel rendelkező tesztek esetében. Továbbá szem előtt tartottuk azt is, hogy a tesztbe épített hosszabb kérdőív kitöltése túl nagy megterheléssel járhat a tanulók számára, ami önmagában is csökkentheti motivációjukat. A kidolgozott kérdőív végső verziója hat részterületet vizsgál: (1) erőfeszítés (*effort, EFF*), (2) érdekesség (*interest, INT*), (3) fontosság (*importance, IMP*), (4) nehézség (*difficulty, DIF*), (5) önhatékonyság (*self-efficacy, SEF*) és (6) kihívás (*challenge, CHL*). A diákoknak válaszaikat egy ötfokú Likert-skálán kellett jelölniük (1=egyáltalán nem igaz – 5=teljes mértékben igaz). Annak érdekében, hogy ne csak egy statikus képet kapjunk a diákok tesztmegoldási motivációjáról, hanem nyomon tudjuk követni annak változását a teszt előrehaladása során, az önértékelő kérdőívet összesen öt alkalommal töltötték ki a tanulókkal. Az első kitöltés a próbafeladat után, még az „éles teszt” megoldása előtt történt, utána három alkalommal a teszt közben, majd a végén, a feladatok megoldása után.

A kérdőív reliabilitása megfelelőnek bizonyult, az egyes adatfelvételi pontokon a Cronbach- α értékei rendre 0,67, 0,71, 0,75, 0,78 és 0,79. Ez alapján a diákok önértékelése a fent említett hat szempont szerint a tesztben előrehaladva egyre homogénebbé vált.

Eljárások

A tesztmegoldási motiváció mérését két, egymástól különböző megközelítéssel végeztük el. Egyrészt a résztvevők saját megítélésén alapuló önértékelő kérdőívet alkalmaztunk, másrészt az eDia-rendszerből (Molnár & Csapó, 2019) származó logadatokat (a tanulók feladatokkal töltött idejét és a kattintások számát) elemeztük. A teszt kiközvetítése 2021 májusában és júniusában valósult meg.

Logadatok

A feladatok kiközvetítése az eDia-rendszerben történt (Csapó & Molnár, 2019), ami lehetőséget biztosít a feladatok megoldása során keletkező kontextuális adatok gyűjtésére. Ezek közül két adatot vontunk be az elemzésbe, (1) a feladatok megoldásával töltött időt (*time-on-task*, *TOT*) és (2) a kattintások számát (*CLICK*). A gyűjtött logadatokból a tanulók saját megítélésétől függetlenül lehet következtetni a tesztmegoldási motivációra. A válaszidő alapú módszerek azon a feltételezésen alapulnak, hogy az alacsony tesztmegoldási erőfeszítéssel rendelkező résztvevők kevesebb időt töltenek a feladat megoldásával, ezért gyorsabban válaszolnak a kérdésekre, mint a magasabb szinten motiváltak (Wise & Kong, 2005). Ehhez hasonlóan a kattintások alacsonyabb száma szintén alacsonyabb szintű motivációt jelez (Sahin & Colvin, 2020).

A válaszidő alapú módszerek esetén a mért indikátor az itemhez tartozó válaszidő, az az időtartam, amit a tesztmegoldó a feladattal tölt. Emellett az adott itemhez tartozóan meg kell határozni egy időküszöbértéket. Amennyiben az adott itemhez tartozó válaszidő a küszöbértéknél rövidebb, akkor azt motiválatlan válasznak tekintjük, míg ha nagyobb vagy egyenlő, akkor motiváltak. Az i itemhez és j vizsgázóhoz tartozó motivált, vagy más néven megoldási viselkedés (*solution behavior*, SB_{ij}) mérésére Wise és Kong (2005) az alábbi összefüggést vezette be:

$$SB_{ij} = \begin{cases} 1, & \text{ha } RT_{ij} \geq T_i \\ 0, & \text{ha } RT_{ij} < T_i \end{cases}$$

ahol T_i = az i itemhez tartozó küszöbérték, RT_{ij} = az i itemhez és j vizsgázóhoz tartozó válaszidő. Szintén Wise és Kong (2005) vezette be a *válaszidő erőfeszítés* (*response time effort*, *RTE*) kifejezést. Az *RTE* az adott résztvevőre vonatkoztatva a motivált válaszok átlagos értéke, vagyis a befektetett erőfeszítés mértéke. Az *RTE* a j vizsgázóra vonatkoztatva

$$RTE_j = \frac{\sum SB_{ij}}{k}$$

ahol k = az itemek száma.

A küszöbérték meghatározására többféle módszert alkalmaztunk, majd ezek alapján számoltuk ki az egyes tesztmegoldókra vonatkozó *RTE*-t. A legegyszerűbb és legrégebb óta alkalmazott megoldás a konstans küszöbérték alkalmazása. Ennek lényege, hogy minden item esetén egy adott, előre meghatározott küszöbértéket alkalmazunk. Ennél kifinomultabb megoldás az itemspecifikus küszöbértékek alkalmazása. Ezek meghatározásának az alapja az a feltételezés, hogy az egyes itemek megoldásához szükséges minimális idő minden item esetén más és más. Amíg a tesztmegoldók képesek gyorsan megoldani egy egyszerű számolási feladatot, addig egy bonyolult problémamegoldó feladat elolvasása, értelmezése és megoldása sokkal több időt vesz igénybe (Goldhammer et al., 2016). Ez alapján nem ugyanazt a küszöbértéket határozzuk meg az összes itemhez, hanem itemenként különbözőt. A kutatásunk során hatféle küszöbértéket vizsgáltunk.

A konstans küszöb módszerek közül a két leggyakrabban használtat alkalmaztuk, a három másodperces (3s) és az öt másodperces (5s) küszöböt (Wise & Kong, 2005). A normatív küszöb módszer (*normative threshold*; *NT10*) (Wise & Ma, 2012) a feladatokkal

töltött idő átlagához viszonyítva határozza meg a küszöbértéket. Az NT10 küszöbérték a vizsgázók által az itemre fordított átlagos idő 10%-a, de maximum 10 másodperc. Például, ha egy adott item esetén az átlagos feladattal töltött idő 38s, akkor az itemhez tartozó küszöbérték 3,8s, viszont ha az átlagos feladattal töltött idő 160s, akkor a küszöbérték nem 16s lesz, hanem 10s. A fenti szabály alapján alkalmaztuk az NT15 és az NT20 küszöbértéket is.

A nullát meghaladó helyes válaszarány (*proportion correct greater than zero; $P+>0\%$*) módszer a feleletalkotó itemek esetén használható. A feleletválasztós tesztek esetén a véletlen találgatás során nullánál nagyobb valószínűséggel születik helyes válasz (pl. egy 4 válaszlehetőséget tartalmazó teszt esetén 0,25). A feleletalkotós itemek esetén, ahol a tesztmegoldóknak saját válaszokat kell adniuk, a helyes válasz elérésének véletlenszerű esélye nulla. A $P+>0\%$ küszöb meghatározásához az adott válaszüzökhöz tartozó helyes válaszok arányát egy másodperces időközönként kell kiszámítani. A küszöb az a leg-
rövidebb válaszüzö, amelynél a helyes válaszok aránya nagyobb, mint nulla (Goldhammer et al., 2016).

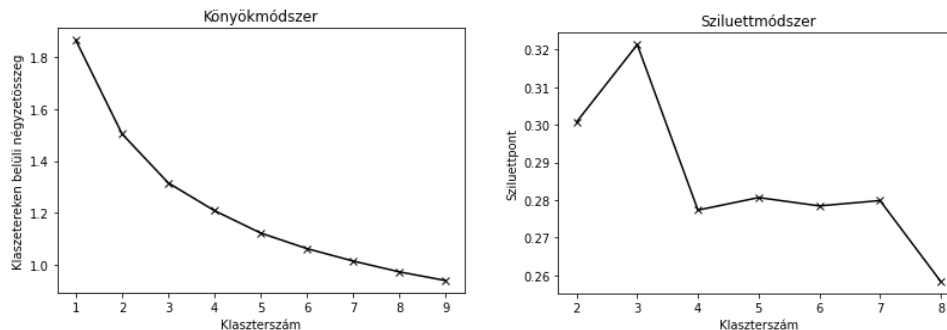
A logadat alapú módszerek közül az optimális kiválasztása érdekében validációs kritériumokat alkalmaztunk. Egy érvényes indikátornak jól el kell különítenie a motiválatlan válaszokat a motivált válaszoktól. Ennek alapja az a feltételezés, hogy a helyes válaszok között nagyobb valószínűséggel kell előfordulniuk a motivált válaszoknak, mint a motiválatlanoknak (Goldhammer et al., 2016).

Tanulói profilok

A tanulói profilok azonosításához k-közép klaszteranalízis elvégzésével alakítottuk ki a tanulói csoportokat. A klaszteranalízis elvégzése során az egyik legfontosabb eldöntendő kérdés az optimális klaszterszám meghatározása. Erre hagyományosan a könyökmódszert (*elbow method*) szokták használni, melynél az optimális klaszterszám az ábrázolt értékek legnagyobb mértékű meredekségváltozásánál található. A módszer hátránya, hogy számos esetben nehezen olvasható le a grafikonról ez az érték, így nem lehet egyértelműen meghatározni az optimális klaszterszámot. Emiatt az elemzések során a sziluettmódszert (*silhouette method*) alkalmaztuk, ahol az optimális klaszterszám a sziluettérték maximuma alapján azonosítható (Shi et al., 2021). A 3. ábra mutatja a könyök- és a sziluettmódszer alkalmazásának összehasonlító elemzését.

A könyökmódszer alapján nehezebben olvasható le az a töréspont, ahol a legnagyobb a meredekség változása (3. ábra). Ennek oka az, hogy több helyen van jelentősebb töréspont: a 2, a 3 és a 6 klaszterszámok esetén is, és ezek közül nehéz a legnagyobbat kiválasztani. A sziluettmódszer alapján egyértelműen azonosítható, hogy a sziluettérték maximuma háromnál van, így ez az érték a klaszterek optimális száma. Az optimális klaszterszám meghatározását minden elemzés esetén elvégeztük, az optimális klaszterszám minden esetben három. A számolásokat és adatelemzéseket az SPSS 25 verziójával, az ábrákat az Excel 2016 verziójával, az optimális klaszterszám meghatározását Python-programmal végeztük el.

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben



3. ábra
A könyök- és a sziluettmódszer esetén adódó optimális klaszterszámok

Eredmények

Logadat alapú módszerek összehasonlítása

Első lépésként a logadat alapú módszereket hasonlítottuk össze, majd ezek közül a validációs kritériumok alapján választottuk ki a legmegfelelőbbet. A tesztmegoldási motiváció mérésére alkalmazott logadat alapú módszerek összehasonlítását az 1. táblázat tartalmazza.

1. táblázat. A motiválatlan válaszok aránya feladatonként és módszerenként, valamint a helyes válaszok aránya

Módszer	Motiválatlan válaszok százalékos aránya feladatonként										Átlag
	1.	2.	3.	4.	5.	6.	7.	8.	9.	10.	
3s	0,00	0,18	0,00	0,18	0,00	0,00	0,18	0,00	0,00	0,18	0,07
5s	0,00	0,36	0,54	0,72	0,36	0,00	0,18	0,55	0,91	1,46	0,51
NT10	0,72	0,36	0,54	0,36	0,18	0,00	0,18	0,00	0,00	0,18	0,25
NT15	1,62	2,17	1,81	1,81	0,36	0,73	0,18	0,55	0,91	1,46	1,16
NT20	3,25	3,61	3,44	3,80	2,36	2,00	2,92	1,46	2,92	2,38	2,81
P+>0%	0,00	1,81	0,90	4,88	0,36	0,00	7,66	17,88	40,69	34,37	10,86
Helyes válaszok aránya	71,30	52,53	28,57	32,55	45,55	50,27	30,47	33,39	4,93	6,95	35,73

Megjegyzés: a módszerként alkalmazott küszöbértékek: 3s: három másodperc; 5s: öt másodperc; NT10: normatív küszöb 10; NT15: normatív küszöb 15; NT20: normatív küszöb 20; P+>0%: nullát meghaladó helyes válaszarány

Feladatonként és módszerenként összehasonlítottuk a motiválatlannak minősített válaszok arányát (1. táblázat). Ezek átlagos mértéke az egyes módszerek alapján jelentősen eltér, 0,07%-tól 10,86%-ig terjed. A legkevesebb választ a 3s módszer, míg a legtöbbet a P+>0% módszer azonosította motiválatlannak. Az egyes feladatok szintjén még jelentősebbek a különbségek. A 9. feladat esetén a 3s módszer egyetlen választ sem azonosított motiválatlanként, míg a P+>0% módszer a válaszok 40,69%-át. A motiválatlan válaszok aránya mellett az 1. táblázat tartalmazza az egyes feladatokhoz tartozó helyes válaszok arányát is, ami azt mutatja, hogy a teszt vége felé nehezedtek a feladatok.

Validációs kritériumok

A 2. táblázat azt mutatja, hogy a hat módszer (3s, 5s, NT10, NT15, NT20, P+>0%) esetében hogyan alakult a motiváltak és a motiválatlannak minősített helyes válaszok aránya. Minden módszer esetén a helyes válaszok arányát a feladatokra adott válaszok eredményeinek átlagolásával kaptuk meg. A P+>0% módszer esetén a motiválatlan helyes válaszok arányának a módszer elvéből adódóan nullának kell lennie. A P+>0% módszer esetén legmagasabb a motivált és motiválatlan helyes válaszok arányának különbsége, azaz ez a módszer különíti el legjobban a motivált válaszokat a motiválatlanoktól. Emiatt a további elemzéseket ezzel a módszerrel végeztük el.

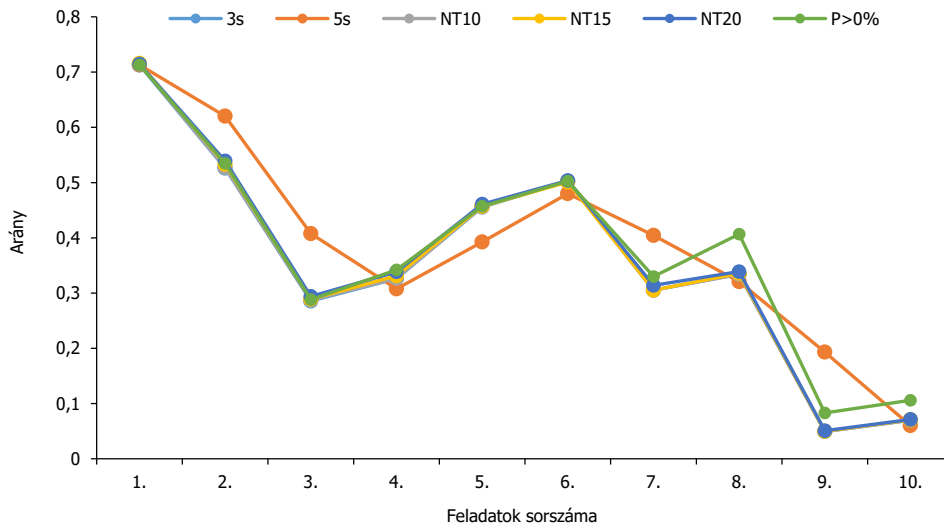
2. táblázat. A motivált és a motiválatlan helyes válaszok aránya és ezek különbsége az egyes módszerek esetén

<i>Módszer</i>	<i>Helyes válaszok aránya (motivált)</i>	<i>Helyes válaszok aránya (motiválatlan)</i>	<i>Különbség</i>
3s	0,36	0,00	0,36
5s	0,36	0,00	0,36
NT10	0,36	0,14	0,22
NT15	0,36	0,17	0,19
NT20	0,36	0,15	0,21
P+>0%	0,40	0,00	0,40

Megjegyzés: a módszerként alkalmazott küszöbértékek: 3s: három másodperc; 5s: öt másodperc; NT10: normatív küszöb 10; NT15: normatív küszöb 15; NT20: normatív küszöb 20; P+>0%: nullát meghaladó helyes válaszarány

A 4. ábra mutatja itemenként a motivált, az 5. ábra a motiválatlan helyes válaszok arányát az egyes módszerek alapján. A motivált helyes válaszok arányában az egyes módszerek között viszonylag kis különbségek vannak, míg a motiválatlan helyes válaszok arányában jelentős különbségek adódnak. A motiválatlan helyes válaszok aránya néhány feladat és bizonyos módszerek alkalmazása esetén igen magas értéket mutat. Ennek oka az, hogy voltak módszerek, amelyek mindössze néhány választ azonosítottak motiválatlannak, és amennyiben azok között volt helyes válasz, akkor ezek aránya nagyon magas lett.

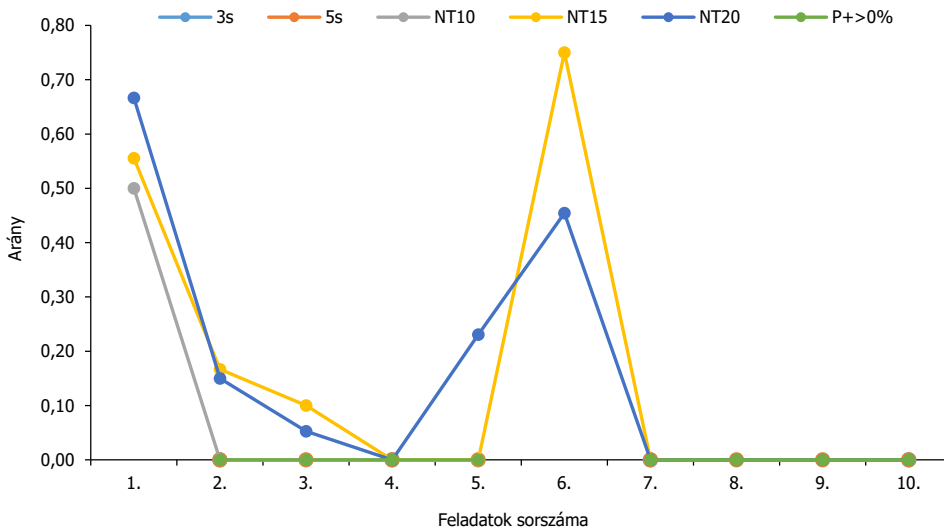
A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben



Küszöbértékek: **3s**: három másodperc; **5s**: öt másodperc; **NT10**: normatív küszöb 10; **NT15**: normatív küszöb 15; **NT20**: normatív küszöb 20; **P+>0%**: nullát meghaladó helyes válaszarány

4. ábra

A motivált helyes válaszok aránya az egyes módszerek alapján



Küszöbértékek: **3s**: három másodperc; **5s**: öt másodperc; **NT10**: normatív küszöb 10; **NT15**: normatív küszöb 15; **NT20**: normatív küszöb 20; **P+>0%**: nullát meghaladó helyes válaszarány

5. ábra

A motiválatlan helyes válaszok aránya az egyes módszerek alapján

Tesztmegoldási motiváció az önértékelő kérdőívre adott válaszok alapján

A tesztmegoldási motiváció mérésére alkalmazott önértékelő kérdőív elemzése érdekében meghatároztuk a kérdőív itemeire adott válaszok és az elért pontszám együttjárását (3. táblázat). Két esetben fordul elő negatív irányú kapcsolat, a nehézség és a kihívás teljesítménnyel való kapcsolatában, vagyis minél nehezebb, minél nagyobb kihívást jelent egy feladat megoldása, annál rosszabb eredményt érnek el a tanulók. A többi esetben pozitív kapcsolat áll fenn, azaz az önértékelő kérdőív alapján magukat motiváltabbnak tartó tanulók tendenciaszerűen jobb eredményeket értek el a teszten.

3. táblázat. A tanulók önértékelő kérdőívre adott válaszainak az elért pontszámmal való korrelációja

Független változó	<i>r</i>	<i>F</i>	<i>R</i> ²	<i>p</i>
Motiváció (TTM)	0,11	6,48	0,01	0,01
Erőfeszítés (EFF)	0,31	56,90	0,09	<0,01
Érdekesség (INT)	0,10	5,90	0,01	0,02
Fontosság (IMP)	0,19	21,36	0,04	<0,01
Nehézség (DIF)	-0,19	20,33	0,04	<0,01
Önhatékonyság (SEF)	0,29	51,29	0,09	<0,01
Kihívás (CHL)	-0,24	34,66	0,06	<0,01

A tesztmegoldási motiváció, a feladattal töltött idő, a kattintások száma és a teljesítmény kapcsolata

Az önértékelő kérdőíves és a logadat alapú módszerek összehasonlítása érdekében meghatároztuk az egyes változók közötti korrelációkat. A 4. táblázat az önértékelő kérdőíven és a logadatokon alapuló tesztmegoldási motiváció (a feladatokkal eltöltött idő, a kattintások száma), valamint a teszten nyújtott teljesítmény közötti korrelációs együtthatókat tartalmazza.

Minden vizsgált esetben szignifikáns az összefüggés (4. táblázat). A feladatokkal töltött idő és a kattintások száma közötti együttjárás bizonyult a legerősebbnek ($r=0,65$, $p<0,01$). Az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció szignifikánsan alacsonyabb korrelációt mutatott ($Z=7,79$, $p<0,01$) a teljesítménnyel ($r=0,11$, $p<0,05$), mint a logadat alapú ($r=0,49$, $p<0,01$). A kattintások száma szignifikánsan erősebb korrelációt mutat ($Z=8,60$, $p<0,01$) a teljesítménnyel ($r=0,41$, $p<0,01$), mint a feladatokkal eltöltött idő ($r=0,11$, $p<0,01$).

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

4. táblázat. A korreláció nagysága az egyes változók között

Változó	Korreláció			
	TTM	TOT	CLICK	SCORE
TTM	–			
TOT	0,20**	–		
CLICK	0,18**	0,65**	–	
SCORE	0,11*	0,11**	0,41**	–
RTE_P+>0%	0,21**	0,54**	0,57**	0,49**

Megjegyzés: * $p < 0,05$, ** $p < 0,01$ (TTM: tesztmegoldási motiváció, TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám, RTE_P+>0%: válaszdő erőfeszítés a nullát meghaladó helyes válaszarány módszer esetén)

A háttérváltozók szerepe

A tesztmegoldási motiváció nemek és életkor szerinti összehasonlítása érdekében független mintás t-próbával, illetve varianciaanalízissel végeztünk vizsgálatot. Nem mutatható ki szignifikáns különbség a fiúk és a lányok motivációjában az önértékelő kérdőívre adott válaszok alapján ($t=1,83$, $p=0,07$), míg a logadatok elemzése alapján a lányok motivációja szignifikánsan magasabb ($t=-2,13$, $p=0,03$). A lányok szignifikánsan több időt töltöttek el a feladatok megoldásával ($t=-2,45$, $p=0,02$), és szignifikánsan többet kattintottak ($t=-2,00$, $p=0,05$) – ennek ellenére az elért pontszámokban nem volt szignifikáns különbség.

Szignifikáns különbség mutatható ki az egyes évfolyamok között a vizsgált változók alapján. Ennek részletes adatait az 5. táblázat tartalmazza.

5. táblázat. Az évfolyamok közötti különbségek a tesztmegoldási motiváció, a feladatokkal töltött idő, a kattintások száma és a teszten nyújtott teljesítmény alapján

Változók	6. évfolyam		7. évfolyam		8. évfolyam		F*	p	Elkülönülő évfolyamok*
	M	SD	M	SD	M	SD			
TTM	3,27	0,74	3,25	0,71	2,86	0,76	15,28	<0,01	{8}<{6, 7}
TOT	481,29	192,96	446,37	195,38	345,62	147,01	22,57	<0,01	{8}<{6, 7}
CLICK	35,89	21,88	31,10	21,09	28,38	18,91	5,49	<0,01	{8}<{6}
SCORE	3,29	2,79	3,25	2,72	4,39	3,34	7,71	<0,01	{6,7}<{8}
RTE_P+>0%	0,92	0,12	0,89	0,14	0,87	0,15	4,65	0,01	{8}<{6}

Megjegyzés: * A szignifikánsan elkülönülő évfolyamok esetén a < a szignifikáns különbség ($p < 0,05$) irányát jelöli. Az évfolyamok közötti összehasonlító oszlop a TTM és az RTE_P+>0% esetén a Tukey-B-próba, a többi esetben a Dunnett-T3-próba alapján a szignifikánsan elkülönülő évfolyamokat mutatja (TTM: tesztmegoldási motiváció, TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám, RTE_P+>0%: válaszdő erőfeszítés a nullát meghaladó helyes válaszarány módszer esetén).

Az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció szerint a 6. és 7. osztályosok között nem mutatható ki különbség, de a 8. osztályosok motiváltsága jelentősen alacsonyabb volt (5. táblázat). A logadatok alapján meghatározott tesztmegoldási motiváció szintje tekintetében a 6. és a 8. osztályosok között volt jelentős különbség. A problémák feltérképezésével töltött időben sem volt jelentős különbség a 6. és a 7. osztályosok között, míg a 8. évfolyamos diákok szignifikánsan kevesebb időt töltöttek a problémák feltérképezésével. A kattintások számában a 6. és a 8. osztályosok között volt jelentősebb a különbség. A 8. osztályosok összességében szignifikánsan jobb eredményt értek el a teszten, mint a 6. és a 7. osztályosok.

A tesztmegoldási motiváció változása a teszt előrehaladása során

A tesztmegoldási motiváció változásának vizsgálatához az egyes mérési időpontok átlagát vettük figyelembe, majd ezeket ismételt méréses ANOVA módszerrel hasonlítottuk össze (6. táblázat).

6. táblázat. Az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció változása a tesztfeladatokban történő előrehaladás során

Mérési alkalom	TTM		ANOVA		Szignifikánsan elkülönülő mérési időpontok
	M	SD	F	p	
1.	3,32	0,71			
2.	3,18	0,82			
3.	3,14	0,87	14,57	< 0,01	{1}>{2, 3, 4}≥{3, 4, 5}
4.	3,10	0,92			
5.	3,07	0,94			

Megjegyzés: A csoportok közötti összehasonlító oszlopban szereplő számok az egyes mérési időpontok eredményeire vonatkoznak. A > a szignifikáns különbség ($p < 0,05$) irányát jelöli (TTM: tesztmegoldási motiváció).

A teszt előrehaladása során szignifikáns különbséget tapasztaltunk (6. táblázat) a tesztmegoldási motivációban (Wilk- $\lambda = 0,90$, $F(4, 543) = 14,57$, $p < 0,01$, $\eta^2 = 0,10$). A páronkénti vizsgálatok során az első és a többi mérési eredmény, valamint a második és az ötödik mérési eredmény között volt szignifikáns különbség, ami arra utal, hogy a tesztmegoldási motiváció szignifikánsan csökkent a tesztben történő előrehaladás közben.

A $P > 0\%$ módszerrel meghatározott tesztmegoldási motiváció változásának vizsgálatához a feladatokhoz tartozó megoldási viselkedés (SB) értékeket vettük figyelembe, majd ezeket ismételt méréses ANOVA módszerrel hasonlítottuk össze. A tesztben történő előrehaladás során szignifikáns a különbség a tesztmegoldási motivációban (Wilk- $\lambda = 0,56$, $F(8, 539) = 53,47$, $p < 0,001$, $\eta^2 = 0,44$). A páronkénti vizsgálatok részletes eredményeit a 7. táblázat tartalmazza, ami alapján a tesztmegoldási motiváció folyamatos, szignifikáns csökkenése mutatható ki – azaz nincs egy állandó, egyénre jellemző, konstans érték.

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

7. táblázat. A $P+>0\%$ alapján meghatározott tesztmegoldási motiváció változása a teszt előrehaladása során

Mérési alkalom	$P+>0\%$		ANOVA		Szignifikánsan elkülönülő mérési időpontok
	M	SD	F	p	
1.	1,00	0,00			
2.	0,99	0,12			
3.	0,99	0,10			
4.	0,95	0,21			
5.	1,00	0,06	53,47	< 0,01	{1, 2, 3, 5, 6}>
6.	1,00	0,00			{4, 7}>{8}>{10}>{9}
7.	0,92	0,27			
8.	0,82	0,38			
9.	0,59	0,49			
10.	0,66	0,48			

Megjegyzés: A csoportok közötti összehasonlító oszlopban szereplő számok az egyes mérési időpontok eredményeire vonatkoznak. A > a szignifikáns különbség ($p<0,05$) irányát jelöli ($P+>0\%$: nullát meghaladó helyes válaszarány módszer).

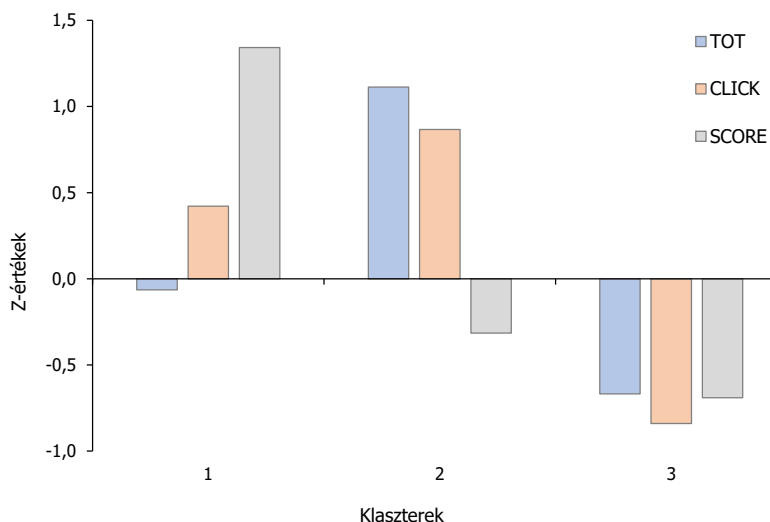
Tanulói profilok azonosítása

Tanulói profilok kizárólag a logadatok alapján

Az első klaszterelemzés során kizárólag a mért független változókat, azaz a feladatokkal eltöltött időt és a kattintások számát, valamint az elért pontszámot vettük figyelembe. A 6. ábrán a változók standardizált értékeinek (Z-értékek) átlagát adtuk meg klaszterek szerint, a 8. táblázat tartalmazza a változók átlagait és szórásértékeit. A varianciaanalízis eredményei alapján szignifikáns különbség van a három klaszter között ($p<0,01$). Az F-értékek alapján leginkább a pontszám, legkevésbé a feladattal töltött idő szerint vannak eltérések a klaszterek átlagai között. A varianciaanalízis utóelemzéseivel vizsgálva a klaszterek közötti különbségeket, a varianciák nem homogének (a Levene-érték minden esetben $p<0,01$), így Dunnett-T3 próbát alkalmaztunk.

Az első klaszterbe 160 tanuló tartozik, a minta 29%-a. Az ebben a klaszterben lévő tanulók közepes mennyiségű idő alatt közepesen sokat próbálkoztak a feladatok megoldásával, azaz közepesen sok kattintást hajtottak végre. A három klaszter közül ők érték el a legjobb eredményt, ők a legjobb képességűek.

A második klaszterbe 151 tanuló tartozik (28%). E tanulók töltötték a legtöbb időt a feladatok megoldásával, ők azok, akik a legtöbbet próbálkoztak a feladatok megoldásával. Az eredményük rosszabb, mint az első klaszterbe tartozóké, de jobb, mint a harmadiké. Az ebbe a klaszterbe tartozó tanulók tehát szignifikánsan hosszabb idő alatt, több próbálkozással szignifikánsan rosszabb eredményt értek el, mint az első klaszter tanulói, azaz gyengébb képességűek az első klaszter tanulóinál.



TOT: feladatokkal töltött idő; **CLICK:** kattintások száma; **SCORE:** elért pontszám

6. ábra

A feladattal eltöltött idő, a kattintások száma és az elért pontszám alapján létrejövő tanulói profilok

A harmadik klaszterbe 236 tanuló tartozik (43%). Rájuk az jellemző, hogy a legkevesebb idő alatt a legkevesebbet próbálkoztak a feladatok megoldásával, azaz a legkevesebbet kattintottak a feladatok megoldása során. A három klaszter közül a legrosszabb eredményt érték el. Az ebben a klaszterben lévő tanulókról állíthatjuk azt, hogy a legkevésbé motiváltak.

8. táblázat. A logadatok alapján létrejött tanulói profilok jellemzői

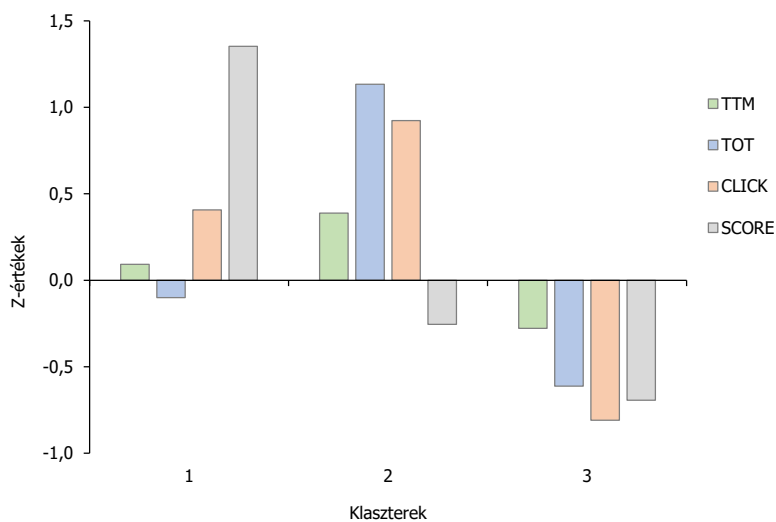
Változó	1. klaszter (N=160)		2. klaszter (N=151)		3. klaszter (N=236)		F*	Elkülönülő klaszterek**
	M	SD	M	SD	M	SD		
1. TOT	-0,06	0,56	1,11	0,89	-0,67	0,60	314,80	{3} < {1} < {2}
2. CLICK	0,42	0,59	0,87	0,92	-0,84	0,48	353,01	{3} < {1} < {2}
3. SCORE	1,34	0,50	-0,32	0,66	-0,69	0,33	872,76	{3} < {2} < {1}

Megjegyzés: *Minden F-érték $p < 0,01$ szinten szignifikáns. **A szignifikánsan elkülönülő klaszterek esetén a <a szignifikáns különbség ($p < 0,05$) irányát jelöli. A klaszterek közötti összehasonlító oszlop a Dunnett-T3-próba alapján a szignifikánsan elkülönülő klasztereket mutatja (TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám).

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

Tanulói profilok: a logadatok a motivációval kiegészítve

A következő klaszterelemzés során a korábbi változók mellett bevontuk az elemzésbe az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció (TTM) értékét is. A 7. ábrán a változók standardizált értékeinek (Z-értékek) átlagát adtuk meg klaszterek szerint, a 9. táblázat tartalmazza a változók átlagait és szórásértékeit. A varianciaanalízis eredményei alapján szignifikáns különbség van a három klaszter között ($p < 0,01$). Az F-értékek alapján megállapítható, hogy leginkább a kattintások száma, legkevésbé a motiváció szerint vannak eltérések a klaszterek átlagai között. A varianciaanalízis utóelemzésével vizsgálva a klaszterek közötti különbségeket, a varianciák nem homogének, így Dunnett-T3 próbát alkalmaztunk.



TTM: tesztmegoldási motiváció az önértékelő kérdőív alapján; **TOT:** feladatokkal töltött idő; **CLICK:** kattintások száma; **SCORE:** elért pontszám

7. ábra

Az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció, a feladattal eltöltött idő, a kattintások száma és az elért pontszám alapján létrejövő tanulói profilok

Az új változó bevonásával az egyes klaszterekbe tartozó tanulók száma minimálisan, legföljebb néhány százalékkal változott meg. Az önértékelő kérdőíven alapuló tesztmegoldási motiváció nagysága összhangban van a logadatokkal az egyes tanulói csoportok esetén. Az első klaszterbe tartozó legjobb képességű tanulók közepes mértékű motivációval rendelkeznek. Számukra nincs is szükség túl nagy motivációra, mivel közepes számú kattintással, közepes mennyiségű idő alatt is a legjobb eredményt érték el. A második klaszterbe tartozó, kevésbé jó képességű tanulók mutatták a legnagyobb mértékű motivációt.

ciót. Ez összhangban van azzal, hogy a legtöbb idő alatt a legtöbbet kattintottak. A harmadik klaszterbe tartozó, legkevésbé motivált tanulók a legkevesebb idő alatt a legkevesebb kattintást hajtották végre a probléma feltárása közben.

9. táblázat. A logadatok és a motiváció alapján létrejött tanulói profilok jellemzői

Változó	1. klaszter (N=156)		2. klaszter (N=146)		3. klaszter (N=245)		F*	Elkülönülő klaszterek**
	M	SD	M	SD	M	SD		
1. TTM	0,09	0,86	0,39	0,80	-0,28	1,10	22,87	{3} < {1} < {2}
2. TOT	-0,10	0,54	1,13	0,89	-0,61	0,66	288,56	{3} < {1} < {2}
3. CLICK	0,41	0,58	0,92	0,91	-0,81	0,51	359,37	{3} < {1} < {2}
4. SCORE	1,35	0,49	-0,26	0,69	-0,69	0,33	845,98	{3} < {2} < {1}

Megjegyzés: *Minden F-érték $p < 0,01$ szinten szignifikáns. **A szignifikánsan elkülönülő klaszterek esetén a <a szignifikáns különbség ($p < 0,05$) irányát jelöli. A klaszterek közötti összehasonlító oszlop a Dunnett-T3-próba alapján a szignifikánsan elkülönülő klasztereket mutatja (TTM: tesztmegoldási motiváció, TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám).

A klaszterelemzés eredményei alapján az önértékelő kérdőíven alapuló tesztmegoldási motiváció jól magyarázható a logadatokkal. A részletesebb elemzések érdekében az önértékelő kérdőív egyedi válaszai alapján is elvégeztük a klaszteranalízist. Ezek során a korábbi változók mellé rendre a tesztmegoldás során kifejtett erőfeszítést (EFF), a feladatok érdekességét (INT), a fontosságot (IMP), a feladatok nehézségét (DIF), az önhatékonyaságot (SEF) és a kihívást (CHL) vontuk be negyedik változóként.

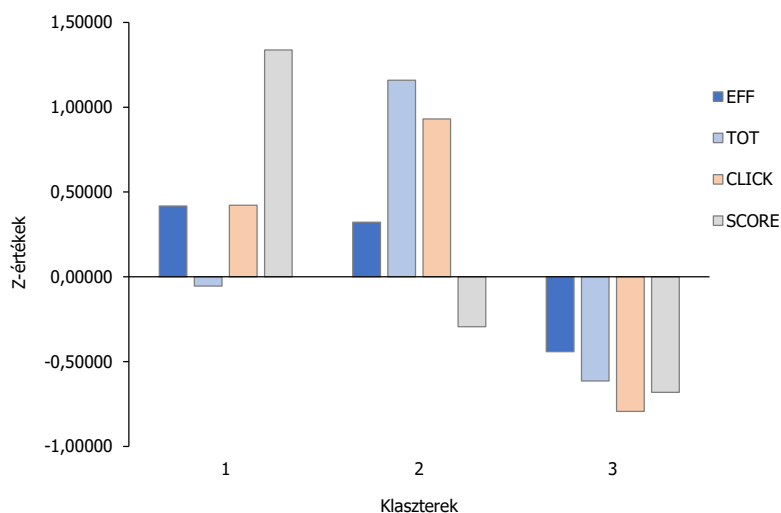
Tanulói profilok: a logadatok a motiváció egyes összetevőivel kiegészítve

A hat elemzés közül kettőt emelünk ki, az erőfeszítés és az önhatékonyaság változóval kiegészített klaszteranalízist. A 8. ábrán az előbbi elemzéshez a változók standardizált értékeinek (Z-értékek) átlagát adtuk meg klaszterek szerint, a 10. táblázat tartalmazza a változók átlagait és szórásértékeit. A varianciaanalízis eredményei alapján szignifikáns különbség van a három klaszter között ($p < 0,01$). Az F-értékek alapján leginkább a pontszám, legkevésbé az erőfeszítés szerint vannak eltérések a klaszterek átlagai között. A varianciaanalízis utóelemzéseivel vizsgálva a klaszterek közötti különbségeket, a varianciák nem homogének, így Dunnett-T3 próbát alkalmaztunk.

Az egyes klaszterekbe tartozó tanulók száma ebben az esetben is csak kismértékben változott a korábbiakhoz képest (8. ábra, 10. táblázat). A harmadik klaszterbe tartozó tanulók szignifikánsan alacsonyabbra értékelték a kifejtett erőfeszítésüket, mint a másik két klaszter tanulói. Ezzel összhangban van az, hogy ők kattintottak a legkevesebbet a legrövidebb idő alatt. Az első két klaszterbe tartozó tanulók átlagosan ugyanakkora mértékűnek értékelték a feladatmegoldás során kifejtett erőfeszítésüket, miközben az első klaszter tanulói szignifikánsan kevesebb idő alatt kevesebb kattintást végeztek el. Az első két klaszter esetén nincs összhang az önértékelő kérdőív válaszai és a logadatok között. Ennek

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

egyik lehetséges magyarázata, hogy az első klaszterbe tartozó legjobb képességű tanulók a társadalmi elvárás miatt magasabbra értékelték az erőfeszítésüket a valóságnál. A másik lehetséges magyarázat, hogy a második klaszterbe tartozó, kevésbé jó képességű tanulók, akiknek gyengébb képességeik miatt általában több erőfeszítés szükséges a feladatok megoldásához, alulértékelték a teszten nyújtott erőfeszítéseiket.



EFF: tesztmegoldási erőfeszítés az önértékelő kérdőív alapján; **TOT:** feladatokkal töltött idő; **CLICK:** kattintások száma; **SCORE:** elért pontszám

8. ábra

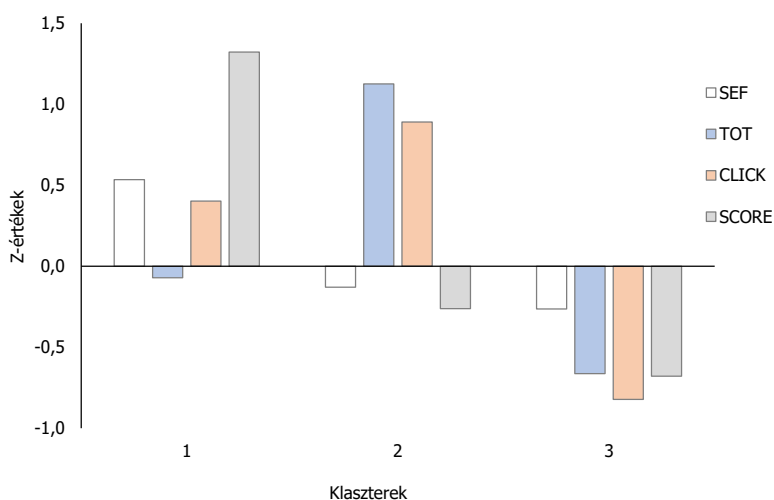
Az erőfeszítés, a feladattal eltöltött idő, a kattintások száma és az elért pontszám alapján létrejövő tanulói profilok

10. táblázat. A logadatok és az erőfeszítés alapján létrejött tanulói profilok jellemzői

Változó	1. klaszter (N=160)		2. klaszter (N=139)		3. klaszter (N=248)		F*	Elkülönülő klaszterek**
	M	SD	M	SD	M	SD		
1. EFF	0,42	0,84	0,32	0,82	-0,44	1,01	54,08	{3} < {1, 2}
2. TOT	-0,06	0,56	1,16	0,90	-0,62	0,65	288,45	{3} < {1} < {2}
3. CLICK	0,42	0,59	0,93	0,93	-0,79	0,52	344,50	{3} < {1} < {2}
4. SCORE	1,34	0,50	-0,30	0,67	-0,68	0,35	853,57	{3} < {2} < {1}

Megjegyzés: *Minden F-érték $p < 0,01$ szinten szignifikáns. **A szignifikánsan elkülönülő klaszterek esetén a <a szignifikáns különbség ($p < 0,05$) irányát jelöli. A klaszterek közötti összehasonlító oszlop a Dunnett-T3-próba alapján a szignifikánsan elkülönülő klasztereket mutatja (EFF: tesztmegoldási erőfeszítés az önértékelő kérdőív alapján; TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám).

A 9. ábrán az önhatékonyság változóval kiegészített elemzéshez a változók standardizált értékeinek (Z-értékek) átlagát adtuk meg klaszterek szerint, a 11. táblázat tartalmazza a változók átlagait és szórásértékeit. A varianciaanalízis eredményei alapján megállapítható, hogy szignifikáns különbség van a három klaszter között ($p < 0,01$). Az F-értékek alapján leginkább a pontszám, legkevésbé az önhatékonyság szerint vannak eltérések a klaszterek átlagai között.



SEF: önhatékonyság az önértékelő kérdőív alapján; **TOT:** feladatokkal töltött idő;
CLICK: kattintások száma; **SCORE:** elért pontszám)

9. ábra

Az önhatékonyság, a feladattal eltöltött idő, a kattintások száma és az elért pontszám alapján létrejövő tanulói profilok

Az első klaszterbe tartozó tanulók szignifikánsan magasabbra értékelték a teszten nyújtott teljesítményüket, mint a másik két klaszter tanulói (9. ábra, 11. táblázat). Ezzel összhangban van az, hogy valóban ők érték el a legjobb teljesítményt. A második és a harmadik klaszter tanulói átlagosan ugyanolyannak értékelték a teszten nyújtott teljesítményüket, miközben a harmadik klaszter tanulói szignifikánsan rosszabb eredményt értek el, mint a második klaszter tanulói. A második és harmadik klaszter esetén nincs összhang az önértékelő kérdőív válaszai és a logadatok között, ami feltételezhetően a tanulók nem teljesen reális önértékeléséből adódik.

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

11. táblázat. A logadatok és az önhatékonyság alapján létrejött tanulói profilok jellemzői

Változó	1. klaszter (N=156)		2. klaszter (N=151)		3. klaszter (N=240)		F*	Elkülönülő klaszterek**
	M	SD	M	SD	M	SD		
1. SEF	0,54	0,87	-0,13	0,86	-0,26	1,03	35,97	{2, 3} < {1}
2. TOT	-0,07	0,55	1,13	0,88	-0,66	0,60	325,55	{3} < {1} < {2}
3. CLICK	0,40	0,57	0,89	0,93	-0,82	0,50	349,49	{3} < {1} < {2}
4. SCORE	1,32	0,55	-0,26	0,71	-0,68	0,36	704,69	{3} < {2} < {1}

Megjegyzés: *Minden F-érték $p < 0,01$ szinten szignifikáns. **A szignifikánsan elkülönülő klaszterek esetén a <math>p < 0,05</math> szignifikáns különbség ($p < 0,05$) irányát jelöli. A klaszterek közötti összehasonlító oszlop az önhatékonyság esetén a Tukey-B-próba, a többi változó esetén a Dunnett-T3-próba alapján a szignifikánsan elkülönülő klasztereket mutatja (SEF: önhatékonyság az önértékelő kérdőív alapján; TOT: feladatokkal töltött idő; CLICK: kattintások száma; SCORE: elért pontszám).

Diszkusszió

A kutatás alapvető céljaként a tesztmegoldási motiváció vizsgálatát tűztük ki ugyanazon mintán többféle módszerrel. A kutatások nagy része egyféle alapelv szerint vizsgálja a tesztmegoldási motivációt, a feltárt szakirodalmak igen kis részében találtunk egyszerre többféle módszerrel végrehajtott kutatást. Ezt támasztja alá Silm és munkatársainak (2020) metaanalízise is, melyben a feldolgozott tanulmányok hozzávetőleg 10%-a alkalmazott többféle megközelítést, döntő többségük egyféle módon mérte a tesztmegoldási motivációt. A kutatásunk során egyrészt önértékelő kérdőívet alkalmaztunk a tesztmegoldási motiváció mérésére, másrészt logadat alapú módszereket. A válaszdő alapú módszerek eredményei között jelentős különbségek adódtak, a motiválatlannak minősített válaszok átlagos mértéke 0,07%-tól 10,86%-ig terjedt. A válaszdő alapú módszerek közül az alkalmazott validációs kritérium alapján a $P > 0\%$ módszer bizonyult a legpontosabbnak, ami összhangban van a Goldhammer és munkatársainak (2016) tanulmányában közölt eredménnyel.

A vizsgált változók (önértékelő kérdőív, illetve a válaszdő alapú tesztmegoldási motiváció, a feladatokkal töltött idő, a kattintások száma és a teszten elért pontszám) között minden esetben szignifikáns korrelációt tapasztaltunk. Az önértékelő kérdőív alapján meghatározott tesztmegoldási motiváció szignifikánsan alacsonyabb korrelációt mutat a teljesítéssel ($r=0,11$), mint a tesztmegoldási erőfeszítés ($r=0,31$). Utóbbi értéke nagy egyezést mutat a Silm és munkatársai (2020) által végzett metaanalízis eredményével ($r=0,33$). Lényegesen erősebb korreláció adódott a logadat alapú módszerrel meghatározott válaszadási erőfeszítés (RTE_ $P > 0\%$) és a teljesítmény között ($r=0,49$), ami szignifikánsan alacsonyabb ($Z=-8,71$, $p < 0,01$) a Silm és munkatársai (2020) által végzett metaanalízis átlagos értékénél ($r=0,72$). Az adatok összességében azt mutatják, hogy az önértékelő kérdőívek és a logadat alapú módszerek nem teljesen ugyanazt mérik.

A teszt során alkalmazott interaktív problémamegoldó feladatok jellegéből adódik, hogy azokat nem lehet fejben megoldani, a sikeres teljesítéshez ki kell próbálni a változók közötti lehetséges összefüggéseket. Azok a tanulók voltak képesek jó eredményt elérni a teszten, akik megfelelő számú próbálkozást hajtottak végre a feladatokon. A kattintások száma szignifikánsan nagyobb mértékű korrelációt mutatott ($Z=8,60$, $p<0,01$) a teljesítménnyel ($r=0,41$), mint a feladatokkal eltöltött idő ($r=0,11$). Ennek oka, hogy a jó képességű problémamegoldók rövid idő alatt voltak képesek nagy számú próbálkozást végrehajtani, míg a rosszabb képességűeknek ehhez sokkal több idő kellett. Nagyon kevés próbálkozással viszont nem lehetett sikeresen megoldani a feladatokat.

A fiúk és a lányok motivációjában az önértékelő kérdőív alapján nem volt kimutatható szignifikáns különbség, míg a válaszügy alapú módszer alapján a fiúk motivációja szignifikánsan alacsonyabb volt ($p=0,03$). A 6. évfolyamos tanulók mindkét módszer szerint szignifikánsan motiváltabbak voltak, mint a 8. osztályosok. Ezek az eredmények szintén összhangban állnak a korábbi kutatásokkal, melyek a fiúk/férfiak, illetve az idősebbek alacsonyabb tesztmegoldási motivációját mutatták ki (Goldhammer et al., 2016; Wise & DeMars, 2010).

A tesztmegoldási motiváció csökkenését mutattuk ki a teszt előrehaladása során mind az önértékelő kérdőív, mind a logadatok eredményei alapján. Ez szintén összhangban van számos korábbi kutatási eredménnyel, melyek szerint a tesztmegoldók tesztmegoldási erőfeszítései a teszt kezdetétől a végéig általában csökkenő tendenciát mutatnak (Attali, 2016; List et al., 2017; Nuutila et al., 2021; Penk & Richter, 2017; Wise & Gao, 2017; Wolgast et al., 2020).

Tanulói profilok meghatározása segítségével azonosítottuk a tanulók egyes csoportjait, amit többféle változó bevonásával végeztünk el. Először kizárólag a mért független változókat, azaz a feladatokkal eltöltött időt és a kattintások számát, valamint az elért pontszámot vontuk be az elemzésbe, a továbbiakban ezen változók mellé a tesztmegoldási motivációt, illetve annak egyes összetevőit. Az optimális klaszterszám minden esetben három volt. Az első esetben – amikor kizárólag mért adatok elemzésével történt a tanulói profilok meghatározása – az első klaszterbe kerültek a legjobb képességű tanulók (160 fő, 29%), akik átlagos mennyiségű idő alatt közepesen sokat próbálkoztak a feladatok megoldásával, miközben a legjobb eredményt érték el. A második klaszterbe gyengébb képességű tanulók tartoznak (151 fő, 28%), mint az elsőbe, mivel több próbálkozást szignifikánsan több idő alatt teljesítettek, miközben szignifikánsan rosszabb eredményt értek el. A harmadik klaszterbe a legkevésbé motivált tanulók tartoznak (236 fő, 43%), náluk a kattintások száma és a feladatokkal töltött idő is szignifikánsan kevesebb volt, mint a másik két klaszterbe tartozó tanulók esetében, és a legrosszabb eredményt érték el. Ez a modell jól magyarázhatóan jellemzi mind az egyes tanulói csoportokat, mind a tanulók összességét.

A következő esetben a fenti változók mellé bevontuk az önértékelő kérdőív alapján meghatározott tesztmegoldási motivációt. Az új változó bevonásával az egyes klaszterekben tartozó tanulók száma minimálisan változott meg. A legjobb képességű tanulók közepes mértékű motivációja azzal magyarázható, hogy a képességeik miatt közepes mértékű motiváció mellett is a legjobb eredményt érték el. A kevésbé jó képességű tanulók legnagyobb mértékű motivációja összhangban van azzal, hogy a legtöbb idő alatt a legtöbbet próbálkoztak a teszt során. A legkevésbé motivált tanulók a legkevesebb idő alatt

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

a legkevesebb kattintást hajtották végre a probléma feltárása közben. Ezek alapján az önértékelő kérdőíven alapuló tesztmegoldási motiváció jól magyarázható a logadatokkal.

A részletesebb elemzések érdekében az önértékelő kérdőív összetevői alapján is elvégeztük a klaszteranalízist, aminek során több esetben nem voltak összhangban az önértékelő kérdőív válaszai a logadatokkal. Az egyik ilyen példa, hogy a legjobb képességű és a gyengébb képességű tanulók az önértékelő kérdőív alapján hasonló mértékű erőfeszítést fejtettek ki a teszt során, miközben a gyengébb képességűek szignifikánsan több idő alatt és többet próbálkoztak a feladatok megoldásával. A másik példa, hogy a gyengébb képességűek és a legkevésbé motiváltak az önértékelő kérdőív alapján ugyanolyannak értékelték a teszten nyújtott teljesítményüket, miközben szignifikáns különbség volt köztük. Ezek az eredmények arra utalnak, hogy a tanulók válaszai nem teljes mértékben tükrözik a valóságot, melyek a tesztmegoldási motivációt vizsgáló önértékelő kérdőívek korlátait jelzik. A tanulók valóságot nem teljes mértékben tükröző válaszainak okai lehetnek egyrészt a társadalmi elvárások, aminek hatására a tanulók egy része nem a valós érzéseit, gondolatait rögzíti a válaszadás során, hanem azt, amit elvárnak tőle. Egy másik lehetséges magyarázat a tanulók egy részének nem megfelelő önismerete, önértékelése.

Jelen kutatás fontos eredménye, hogy az önértékelő kérdőívekre adott válaszok nem minden esetben vannak összhangban a tanulók konkrét viselkedési mintázatával. Több esetben ellentmondást mutattunk ki a tanulók egy részének válaszai és az általuk a feladatok megoldása során nyújtott viselkedés között. A tesztmegoldási motiváció alacsony tétellel rendelkező tesztek esetén jelentős mértékben befolyásolhatja a teszten nyújtott teljesítmények értékelését. A számítógép alapú tesztek esetében a logadatok elemzése fontos segítséget nyújthat az eredmények érvényességének biztosításában.

Limitáció

A kutatás egyik korlátjának a viszonylag kis mintaszám tekinthető, ami miatt az eredmények nem általánosíthatók. A másik fontos korlát, hogy a teszt kizárólag interaktív problémamegoldást mérő feladatokat tartalmazott, amelyek a számítógéppel való interakciók, kattintások, próbálkozások alapján oldhatók meg sikeresen. Emiatt ugyanezen elemzéseket nem lehetne alkalmazni sok más feladattípus, például egy többszörös választásos teszt során, ahol minden item esetén egy kattintással meg lehet adni a helyes választ.

Konklúzió

A kutatás megerősíti számos korábbi tanulmány eredményét, melyek közül a legfontosabb, hogy az önértékelésen és a logadatokon alapuló tesztmegoldási motiváció az alapvető kérdésekre hasonló válaszokat ad, azonban nincs teljes összhang közöttük. A kettő közötti különbséget támasztja alá a kétféle módon mért tesztmegoldási motiváció és a teljesítmény közötti korreláció jelentős eltérése. Ugyancsak erre utalnak a klaszterelemzés eredményei is. Az önértékelő kérdőív alapú tesztmegoldási motiváció összességében jól magyarázható a logadatokkal, azonban az összetevők elemzése során több esetben ellent-

mondást tapasztaltunk. Az önértékelő kérdőív válaszai és a logadatok közötti ellentmondás okai lehetnek – többek között – a társadalmi elvárások, valamint a tanulók nem reális önértékelése.

A klaszterelemzés alapján felállított tanulói profilokból az előzőeken túl arra is lehet következtetni, hogy a válasziidő alapú módszereknek is vannak korlátai. A feladattal eltöltött idő és a kattintások száma összességében együttjárást mutat a teljesítménnyel, ugyanakkor az egyes tanulói profilok között szignifikáns különbségek adódtak. További kutatási irány lehet a válasziidő alapú módszerek pontosságának a növelésére a kattintások számának bevonása az elemzésekbe.

Köszönetnyilvánítás

A tanulmány az Innovációs és Technológiai Minisztérium Kooperatív Doktori Program Doktori Hallgatói Ösztöndíj Programjának a Nemzeti Kutatási, Fejlesztési és Innovációs Alapból finanszírozott szakmai támogatásával készült. A kutatást az OTKA K135727, illetve az MTA Közoktatás-fejlesztési Kutatási Pályázata támogatta (KOZOKT2021-16).

Irodalom

- Abari, K., Polonyi, T., & Tünde, É. (2017). A gamifikáció lehetőségei a nyelvtanulásban. In T. Polonyi (Ed.), *Digitális tanulás és tanítás* (pp. 159–187). Debreceni Egyetemi Kiadó.
- Attali, Y. (2016). Effort in low-stakes assessments: What does it take to perform as well as in a high-stakes setting? *Educational and Psychological Measurement*, 76(6), 1045–1058. doi: [10.1177/0013164416634789](https://doi.org/10.1177/0013164416634789)
- Butler, J., & Adams, R. J. (2007). The impact of differential investment of student effort on the outcomes of international studies. *Journal of Applied Measurement*, 8(3), 279–304.
- Crombach, M. J., Boekaerts, M., & Voeten, M. J. M. (2003). Online measurement of appraisals of students faced with curricular tasks. *Educational and Psychological Measurement*, 63(1), 96–111. doi: [10.1177/0013164402239319](https://doi.org/10.1177/0013164402239319)
- Csapó, B., & Molnár, G. (2019). Online diagnostic assessment in support of personalized teaching and learning: The eDia system. *Frontiers in Psychology*, 10. Article 15222. doi: [10.3389/fpsyg.2019.01522](https://doi.org/10.3389/fpsyg.2019.01522)
- Csapó, B., Molnár, G., & R. Tóth, K. (2008). A papír alapú tesztetől a számítógépes adaptív tesztelésig: a pedagógiai mérés-értékelés technikájának fejlődési tendenciái. *Iskolakultúra*, 18(3–4), 3–16.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, 53(1), 109–132. doi: [10.1146/annurev.psych.53.100901.135153](https://doi.org/10.1146/annurev.psych.53.100901.135153)
- Finn, B. (2015). Measuring motivation in low-stakes assessments. *ETS Research Report Series*, 2015(2), 1–17. doi: [10.1002/ets2.12067](https://doi.org/10.1002/ets2.12067)
- Freund, P. A., Kuhn, J. T., & Holling, H. (2011). Measuring current achievement motivation with the QCM: Short form development and investigation of measurement invariance. *Personality and Individual Differences*, 51(5), 629–634. doi: [10.1016/j.paid.2011.05.033](https://doi.org/10.1016/j.paid.2011.05.033)
- Goldhammer, F., Martens, T., Christoph, G., & Lüdtke, O. (2016). Test-taking engagement in PIAAC. *OECD Education Working Papers*, 133, 0–67. doi: [10.1787/5jlzfl6fhxs2-en](https://doi.org/10.1787/5jlzfl6fhxs2-en)
- Greiff, S., Wüstenberg, S., Molnár, G., Fischer, A., Funke, J., & Csapó, B. (2013). Complex problem solving in educational contexts—Something beyond g: Concept, assessment, measurement invariance, and construct validity. *Journal of Educational Psychology*, 105(2), 364–379. doi: [10.1037/a0031856](https://doi.org/10.1037/a0031856)

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

- Lee, Y.-H., & Chen, H. (2011). A review of recent response-time analyses in educational testing. *Psychological Test and Assessment Modeling*, 53(3), 359–379. http://www.psychologie-aktuell.com/fileadmin/download/ptam/3-2011_20110927/06_Lee.pdf
- Lindner, M. A., Lüdtke, O., & Nagy, G. (2019). The onset of rapid-guessing behavior over the course of testing time: A matter of motivation and cognitive resources. *Frontiers in Psychology*, 10, Article 1533. doi: 10.3389/fpsyg.2019.01533
- List, M. K., Robitzsch, A., Lüdtke, O., Köller, O., & Nagy, G. (2017). Performance decline in low-stakes educational assessments: different mixture modeling approaches. *Large-Scale Assessments in Education*, 5, Article 15. doi: 10.1186/s40536-017-0049-3
- Mezei, G. (2013). Motivációs stratégiák a nyelvórán. *Modern Nyelvoktatás*, 19(1–2), 36–50.
- Molnár, É. (2009). Az önszabályozás értelmezései és elméleti megközelítései. *Magyar Pedagógia*, 109(4), 343–364.
- Molnár, G. (2021). Challenges and developments in technology-based assessment: Possibilities in science education. *Europhysics News*, 52(2), 16–19. doi: 10.1051/ejn/2021202
- Molnár, G., & Csapó, B. (2018). The efficacy and development of students' problem-solving strategies during compulsory schooling: Logfile analyses. *Frontiers in Psychology*, 9, Article 302. doi: 10.3389/fpsyg.2018.00302
- Molnár, G., & Csapó, B. (2019). A diagnosztikus mérési rendszer technológiai keretei: Az eDia online platform. *Iskolakultúra*, 29(4–5), 16–32. doi: 10.14232/ISKKULT.2019.4-5.16
- Nuutila, K., Tapola, A., Tuominen, H., Molnár, G., & Niemivirta, M. (2021). Mutual relationships between the levels of and changes in interest, self-efficacy, and perceived difficulty during task engagement. *Learning and Individual Differences*, 92, Article 102090. doi: 10.1016/j.lindif.2021.102090
- Penk, C., & Richter, D. (2017). Change in test-taking motivation and its relationship to test performance in low-stakes assessments. *Educational Assessment, Evaluation and Accountability*, 29(1), 55–79. doi: 10.1007/s11092-016-9248-7
- Rios, J. A. (2021). Improving test-taking effort in low-stakes group-based educational testing: A meta-analysis of interventions. *Applied Measurement in Education*, 34(2), 1–22. doi: 10.1080/08957347.2021.1890741
- Sahin, F., & Colvin, K. F. (2020). Enhancing response time thresholds with response behaviors for detecting disengaged examinees. *Large-Scale Assessments in Education*, 8, Article 5. doi: 10.1186/s40536-020-00082-1
- Schüttelpelz-Brauns, K., Kadmon, M., Kiessling, C., Karay, Y., Gestmann, M., & Kämmer, J. E. (2018). Identifying low test-taking effort during low-stakes tests with the new Test-taking Effort Short Scale (TESS) – Development and psychometrics. *BMC Medical Education*, 18(1), Article 101. doi: 10.1186/s12909-018-1196-0
- Shi, C., Wei, B., Wei, S., Wang, W., Liu, H., & Liu, J. (2021). A quantitative discriminant method of elbow point for the optimal number of clusters in clustering algorithm. *EURASIP Journal on Wireless Communications and Networking*, 2021, Article 31. doi: 10.1186/s13638-021-01910-w
- Silm, G., Pedaste, M., & Täht, K. (2020). The relationship between performance and test-taking effort when measured with self-report or time-based instruments: A meta-analytic review. *Educational Research Review*, 31, Article 100335. doi: 10.1016/j.edurev.2020.100335
- Theilk, A., Sundre, D., Horst, S., & Finney, S. (2009). Motivation matters: Using the Student Opinion Scale to make valid inferences about student performance. *Journal of General Education*, 58(3), 129–151. doi: 10.1353/jge.0.0047
- Tóth, K., Rölke, H., Goldhammer, F., & Barkow, I. (2017). Educational process mining: New possibilities for understanding students' problem-solving skills. In B. Csapó & J. Funke (Eds.), *The nature of problem solving. Using research to inspire 21st century learning* (pp. 193–209). OECD. doi: 10.1787/9789264273955-14-en

- Vroom, V. H. (1964). *Work and Motivation*. Wiley.
- Wise, S. L., & DeMars, C. E. (2005). Low examinee effort in low-stakes assessment: Problems and potential solutions. *Educational Assessment, 10*(1), 1–17. doi: [10.1207/s15326977ea1001_1](https://doi.org/10.1207/s15326977ea1001_1)
- Wise, S. L., & DeMars, C. E. (2010). Examinee noneffort and the validity of program assessment results. *Educational Assessment, 15*(1), 27–41. doi: [10.1080/10627191003673216](https://doi.org/10.1080/10627191003673216)
- Wise, S. L., & Gao, L. (2017). A general approach to measuring test-taking effort on computer-based tests. *Applied Measurement in Education, 30*(4), 1–12. doi: [10.1080/08957347.2017.1353992](https://doi.org/10.1080/08957347.2017.1353992)
- Wise, S. L., & Kong, X. J. (2005). Response Time Effort: A new measure of examinee motivation in computer-based tests. *Applied Measurement in Education, 18*(2), 163–183. doi: [10.1207/s15324818ame1802_2](https://doi.org/10.1207/s15324818ame1802_2)
- Wise, S. L., & Ma, L. (2012, April, 14–16). *Setting response time thresholds for a CAT item pool: The normative threshold method* [Conference Presentation]. 2012 National Council on Measurement in Education (NCME) Annual Conference, Vancouver, British Columbia, Canada.
- Wise, S. L., Ma, L., & Theaker, R. A. (2014). Identifying non-effortful student behavior on adaptive tests: Implications for test fraud detection. In N Kingston & A. Clark (Eds.), *Test fraud: Statistical detection and methodology* (pp. 175–185). Taylor and Francis.
- Wolgast, A., Schmidt, N., & Ranger, J. (2020). Test-taking motivation in education students: Task battery order affected within-test-taker effort and importance. *Frontiers in Psychology, 11*, Article 559683. doi: [10.3389/fpsyg.2020.559683](https://doi.org/10.3389/fpsyg.2020.559683)

A tesztmegoldási motiváció kérdőíves és logadat alapú mérésének összehasonlító elemzése alacsony tétellel rendelkező interaktív problémamegoldó környezetben

ABSTRACT

COMPARATIVE ANALYSIS OF THE MEASUREMENT POSSIBILITIES OF TEST-TAKING MOTIVATION IN LOW-STAKES TESTS

Róbert Csányi & Gyöngyvér Molnár

Students' test results are not only determined by their level of knowledge and skills. They are influenced by various affective factors like motivation, especially test-taking motivation; and the stake of the test significantly influences students' test-taking motivation. As the stake of a test decreases, the test-taking motivation of the students may decrease proportionally, and it may affect test performance. The present study investigated an empirical research on the influential factor of test-taking motivation on students' test results. The sample of the study was drawn from 6 to 8 grade students (n=554) who had to explore and solve interactive problems in uncertain situations that did not require prior knowledge. Students' test-taking motivation was measured in two ways; using (1) a self-report questionnaire and (2) log data (time-on-tasks, number of clicks). The test-taking motivation of boys proved to be significantly lower than that of the girls' as well as higher graders' test-taking motivation was significantly lower than that of the lower graders'. Generally, students' test-taking motivation correlated significantly with their test performance, however, if we used the self-reported questionnaire data, this relationship was lower than when we used logfile data to measure test-taking motivation. The number of clicks correlated on a higher level with students' final results than time on task. We detected a significant decrease in students' test-taking motivation while moving forward in the test, i.e., test-taking motivation is not an individual, constant value, but it also depends on other factors. In several cases, we found a discrepancy between the answers to the self-report questionnaire and students' logged test-taking behavior, i.e., the real behavior did not confirm their answers. The discrepancy between the answers to the self-report questionnaire and the real test-taking behavior may be attributed to social expectations and the not completely realistic self-assessment of students. The results of the study also point to the limitations of both self-report questionnaires and time-on-task-based methods.

Magyar Pedagógia, 121(3). 281–307. (2021)
DOI: 10.17670/MPed.2021.3.281

Levelezési cím / Address for correspondence:

Csányi Róbert, Szegedi Tudományegyetem Neveléstudományi Doktori Iskola. H-6722 Szeged, Petőfi Sándor sgt. 32–34.

Molnár Gyöngyvér, Szegedi Tudományegyetem Neveléstudományi Intézet. H-6722 Szeged, Petőfi Sándor sgt. 32–34.